

Mathématiques pour la Physique

Analyse et Géométrie

Léo Bernus, Loïc Chantry

Note de cours

2019-2020

Table des matières

Table des matières	i
Table des figures	vii
Liste des tableaux	xi
Avant-propos	1
Conventions et notations	3
I Espaces vectoriels.	5
I.1 Espace vectoriel	5
I.1.1 Structure d'espace vectoriel	5
Rappel et définition sur les structures d'ensembles	5
I.1.2 Sous espace vectoriel	10
I.1.3 Dépendance et indépendance linéaire	11
I.1.4 Dimension finie	14
I.2 Applications linéaires	15
I.2.1 Généralités	15
I.2.2 Théorèmes du rang et dimension	17
I.2.3 Projecteurs	17
I.2.4 Dualité	17
I.2.5 Matrice	18
I.2.6 Determinant	20
I.3 Réduction des endomorphismes	24
I.3.1 Arithmétique des polynomes	24
I.3.2 Calculs polynomiaux sur les endomorphismes	28
I.3.3 Sous-espace stable, valeurs propre sous espace. propre	29
I.3.4 Diagonalisation et trigonalisation	30
Endomorphisme diagonalisables	30
Endomorphisme trigonalisables	30
II Topologie, continuité, produit scalaire	31
II.1 Norme et topologie	32
II.1.1 Norme et distance	32
Généralités	32
Boules et sphère	32
Suite d'un espace normé	32
Normes équivalentes	32
II.1.2 Topologie	32
Généralités	32
Adhérence et caractérisation séquentiel des fermés	32
Partie dense	32
Intérieur, Frontière	32
II.1.3 Espace complet	32
II.2 Continuité et compacité	32

II.2.1 Limite et continuité	32
II.2.2 Compacité	32
II.2.3 Connexité par arc	32
II.2.4 Formes quadratiques, produits scalaire, endomorphismes symétrique et orthogonaux	32
II.2.5 Forme bilinéaire et forme quadratique	32
II.2.6 Réduction d'une forme quadratique	32
II.2.7 Produit scalaire	32
II.2.8 Endomorphisme auto-adjoint	32
II.2.9 Auto-morphisme orthogonaux	32
II.2.10 Produit mixte et. produit vectoriel	32
III Bases du calcul différentiel	33
III.1 Introduction	33
III.2 Fonctions de plusieurs variables à valeur dans les réels	34
III.2.1 Dérivée directionnelle, ou dérivée le long d'un vecteur	34
Remarque	35
Exercices	35
Remarques	36
Exercice	36
III.2.2 Dérivée partielle	37
Remarque	37
Gradient	37
Remarques	38
Exercices	38
III.3 Différentiabilité	39
III.3.1 Différentielle	40
Représentation matricielle	40
Exemple	41
III.3.2 Propriétés	42
III.3.3 Condition suffisante	42
Exemples	42
III.4 Composition et opération sur les différentielles	43
III.4.1 Opérations élémentaires	43
III.4.2 Composition des différentielles	44
Matrices	45
III.4.3 Difféomorphisme et changement de variables	46
Règle de la chaîne	46
Exemple simple	47
Coordonnées polaires	47
III.4.4 Dérivées d'ordre supérieur	48
Théorème de Schwarz	49
Développement de Taylor	49
III.4.5 Exercices	50
Équation d'onde	50
Solution	50
Laplacien	51
III.5 Théorèmes d'inversion	52
III.5.1 Inversion locale	52
III.5.2 Inversion globale	53
Exemple important : coordonnées polaires	53
Courbe paramétrée	53
Exercice	54
III.5.3 Fonctions implicites	54
Approche heuristique	54
Théorème des fonctions implicites	55
Interprétation graphique	55
Différentielle de la fonction implicite	55

Exercices sur les fonctions implicites	56
Échauffement : le cercle	56
Théorie des perturbations	57
Thermodynamique	58
III.6 Annexe : preuves techniques	60
III.6.1 Différentiabilité	60
III.6.2 Théorèmes d'inversion	60
Théorème d'inversion locale	60
Théorème des fonctions implicites	60
IV Courbes	61
IV.1 Courbes en dimension finie	61
IV.1.1 Généralité et Définition	61
IV.1.2 Abscisse Curviligne	62
IV.1.3 Etude locale et Branche infinie	63
IV.2 Courbes planes	64
IV.2.1 Courbure	64
IV.2.2 Convexité	69
IV.3 Courbes gauches	71
IV.3.1 Base de Frenet, Courbure, Plan osculateur, Cercle osculateur et des courbes gauches	71
IV.3.2 Torsion	73
IV.3.3 Hélice	74
IV.4 Exercices	76
V Surfaces régulières	77
V.1 Surfaces sans bord	77
V.1.1 Définition	77
V.1.2 Graphes de fonctions	79
V.1.3 Submersion	82
V.1.4 Les Quadriques	84
Quadriques non dégénérées	84
Quadrique dégénérée une fois	85
Quadrique dégénérée deux fois	87
V.1.5 Plan tangent	87
V.1.6 Continuité sur une surface	88
V.1.7 Champs de vecteurs et fibré tangent	90
V.2 Première et seconde formes fondamentales	90
V.2.1 Définitions première forme fondamentale	90
V.2.2 Du problème du plongement	94
V.2.3 Application de Gauss et seconde forme fondamentale	95
V.2.4 Orientation	97
V.3 Theorema Egregium de Gauss	98
V.3.1 Courbure(s)	99
V.3.2 Dérivation et symbole de Christoffel	101
Definitions	101
Courbure géodésique et normale d'une courbe paramétrée sur une surface	102
V.3.3 Le théorème remarquable et les équations de Mainardi-Coddazzi	103
V.3.4 Équation de Gauss et de Coddazzi-Mainardi dans la base principale	105
V.4 Classification des surfaces compactes à bord	108
V.4.1 Definitions des surfaces à bord	108
V.4.2 Partition intérieur extérieur	108
V.4.3 Bord localement difféomorphe à une courbe paramétrée	109
V.4.4 Surface compacte et bord comme union de courbes paramétrées fermées simples	110
VI Variétés	115
VI.1 Qu'est ce qu'une variété...	116
VI.1.1 Qu'est-ce qu'une variété ?	116

VI.1.2	Espace tangent	117
	Construction géométrique de l'espace tangent	117
	Construction algébrique de l'espace tangent	118
	Transformation contravariante des coordonnées par changement de carte .	120
	Exercices-exemples	121
VI.1.3	Dualité : formes linéaires et espace cotangent	121
	Construction de l'espace dual	121
	Transformation covariante des coodonnées par changement de carte . . .	123
VI.1.4	Les tenseurs en toute généralité	123
	Prélude : dual du dual	123
	Fugue : construction des tenseurs	124
	Exercice	125
	Toccata : transformation contravariante et covariante des coordonnées d'un tenseur	126
VI.1.5	Le tenseur métrique	126
	Propriétés du tenseur métrique	127
	Dualité métrique entre vecteurs et formes	128
	Musique tensorielle : dualité métrique des tenseurs	128
VI.2	Algèbre extérieure	129
VI.2.1	En guise de prélude : les 2-formes	129
	Propriétés des 2-formes	129
	Le champ électromagnétique comme 2-forme	130
VI.2.2	Formes multilinéaires alternées, ou p -formes	132
VI.2.3	Le tenseur de Levi-Civita	134
	Espace des n -formes comme espace vectoriel de dimension 1	134
	Exercices - exemples	135
	Produit vectoriel de $n - 1$ vecteurs	135
	Gymnastique avec le tenseur de Levi-Civita	136
	Dualité de Hodge	137
	Quelques applications dans l'espace-temps relativiste	138
VI.2.4	Le champ électromagnétique	139
VI.2.5	Prolégomène : particule dans un champ de force pure	139
	Construction du champ électromagnétique	140
VI.2.6	Exercices	140
VI.3	Dérivation	141
VI.3.1	Connexion covariante de Koszul, de Riemann, de Levi-Civita	141
	Différentiel des scalaires sur une variété	141
	Coordonné orthogonal	142
	Connexion covariantes de Koszul	142
	Tenseur parrallèle	143
	Christoffel d'une base dans une connexion de Koszul	143
	Crochet de commutation	144
	Torsion dans une connexion de Koszul	144
	Connexion trivial	145
	Construction de la connexion trivial à l'aide d'un plongement . . .	145
	Propriétés de la connexion trivial	146
	Connexion de Riemann et de Levi-Civita	148
	Riemann	148
	Connexion de Levi-Civita	149
	Théorème fondamental de la géométrie Riemannienne	149
	Propriété de la dérivation covariante.	150
	Divergence...	151
VI.3.2	Formes différentielles et dérivation extérieur	153
	Dérivation extérieur	153
	Rotationnel	156
	Lemme de Poincaré	157
	Definition des formes exact et fermé	157
	Circulation des 1-formes	157

Lemme de Poincaré	158
Dans les variété de dimension 3	159
Exercice en rab	160
VI.4 Intégration	160
VI.4.1 Intégration sur un volume	160
VI.4.2 Intégration sur une sous variété a bords de dimension p	162
Sous variété à bords et orientation du bords	162
Intégrale d'une n-p-forme où d'une p-forme	163
Cas des surfaces de l'espace ordinaire de dimension 3	163
VI.4.3 Théoreme de Stockes	164
Formulation générale	164
Dans l'espace ordinaire de dimension 3	164
Stockes	164
Green-Ostrogradski	165
VI.5 Flots et dérivée de Lie	165
VI.5.1 Flot d'un champ de vecteurs	165
Équation différentielle dans une variété	165
Équations aux variations	166
Exercice : pendule simple.	168
Intégrales premières et submersion	168
VI.5.2 Dérivée de Lie	170
Minimum vital sur l'image réciproque du flot	170
Propriétés de la dérivée de Lie	170
Coordonnées naturelles de la dérivée de Lie	173

Table des figures

I.1	Représentation des racines sixième de l'unité dans le plan complexe (source : mathraining.be)	8
I.2	Représentation des fonctions injective, surjective et bijective (source : Wikipedia)	16
III.1	Représentation d'un ouvert E contenant un point a , un voisinage de a noté $\mathcal{V}(a)$, d'une boule ouverte de rayon r centrée sur a notée $B(a, r)$, et comment faire rentrer $a + t\mathbf{v}$ dans $B(a, r)$ et donc dans $\mathcal{V}(a)$	34
III.2	Représentation de la nappe d'équation $z = \exp(-2x^2 - y^2)$ et d'une ligne passant par cette surface, d'équation paramétrique $t \mapsto (-1+t, -1+t, z((-1+t, -1+t)))$	35
III.3	Quelques exemples de lignes de niveaux de fonctions de deux variables à valeurs dans \mathbb{R} . Dans ce cas, les lignes de niveau sont également les surfaces de niveau. À gauche : les lignes de niveau du potentiel électrostatique engendré par une charge ponctuel sont des cercles concentriques, et les lignes de champ sont les droites qui concourent en la particule. À droite : lignes de niveau d'une fonction différentiable quelconque avec une ligne de champ.	38
III.4	Lignes de niveau de la montagne d'équation $z = e^{-2x^2 - y^2}$. On voit que le vecteur \mathbf{u} est bien orthogonal à la ligne de niveau.	39
III.5	On représente la surface plongée dont le paramétrage est décrit par l'équation (III.44). Nous avons représenté les vecteurs colonnes du jacobien (eq. (III.45)), $\partial_\theta f(\theta, u)$ et $\partial_u f(\theta, u)$ en deux points de la surface. Nous voyons qu'ils sont tangents aux lignes $u = cst$ et $\theta = cst$ respectivement.	41
III.6	Représentation du graphe de la fonction définie à l'équation (III.48). On voit clairement que toutes les dérivées le long de la direction y sont nulles mais que ça n'empêche aucunement la fonction de dépendre de y à cause de la non convexité de l'ensemble de définition.	43
III.7	Représentation symbolique de la composition des différentielles. En bleu : composition de f et g . En rouge : composition de $df(a)$ et $dg(f(a))$	45
III.8	Représentation du changement de variables en coordonnées polaires.	47
III.9	À gauche on a représenté un cas où $x > 0$, à droite un cas où $x < 0$	48
III.10	Cercle polaire.	48
III.11	À gauche, un arc paramétré injectif et régulier, donc bijectif. À droite, un arc paramétré régulier, mais non injectif : on voit que $\gamma(t_2) = \gamma(t_3)$ avec $t_2 \neq t_3$	53
III.12	Graphe de la trajectoire d'une chute libre.	54
III.13	Exemple d'ensemble de départ et d'ensemble d'arrivée pour que la fonction carré soit un \mathcal{C}^1 difféomorphisme.	54
III.14	Nous représentons en haut à gauche les lignes de niveau d'une fonction de deux variables à valeurs dans \mathbb{R} . Nous avons représenté les voisins correspondant aux notations du théorème en un point où $\partial_2 f(a, b) \neq 0$. Nous avons indiqué quelques cas où $\partial_2 f(a, b) \neq 0$ donc avec une tangente verticale. En dessous, nous avons représenté deux zooms : l'un où $\partial_2 f = 0$, auquel cas la ligne n'est localement pas le graphe d'une fonction (un antécédent peut donner plusieurs images), et l'autre où $\partial_2 f \neq 0$, dans ce cas la ligne est localement le graphe d'une fonction de x . À droite nous avons tenté de faire une représentation en perspective pour montrer à quoi correspond $\mathcal{W}_c = f(\mathcal{U}_a \times \mathcal{V}_b)$	56

IV.1 Représentation d'une hélice (bleu) de rayon $a = 1$ et de pas $b = 1$ à l'intérieur d'un cylindre (rouge).	64
IV.2 Représentation des différents type de point sur une courbe géométrique	65
IV.3 Représentation de la courbe paramétrée $t \mapsto a(q-1)u(t) + apu((1-q)t)$ pour $a = 4$ $p = 3$ $q = 5$. Deux cercles osculateur sont représenté en vert et les tangente en bleu.	68
IV.4 Représentation des différents type de possibilité concernant la convexité locale. À gauche, une courbe localement convexe, est représentée autour d'un point en rouge l'ensemble convexe bordant localement le point considéré. À droite, une courbe avec un point d'inflexion pour lequel il n'existe aucun ensemble convexe qui le borde, on comprend à l'aide des segments en noir et du fait que la courbe traverse sa tangente en ce point (en bleu), que quelle que soit la manière dont on cherche un ensemble qui borde avec la condition de l'arc ouvert, cet ensemble ne sera pas convexe.	69
IV.5 Cas où la courbe géométrique comporte un morceau de droite, cette courbe est localement convexe, néanmoins sa courbure change de signe. Concernant les points sur le morceau de droite les ensembles convexe peuvent être choisis d'un côté ou de l'autre de la courbe.	70
IV.6 Construction de \tilde{A} dans la première partie de la démonstration de la proposition IV.2.10.	71
IV.7 Représentation du plan osculateur de la tangente et du cercle osculateur d'une courbe représentant la couture d'une balle de tennis $t \mapsto (a \cos(t) + b \cos(3t), a \sin(t) + b \sin(3t), c \sin(2t))$ (ou $c = 2\sqrt{ab}$ et $b = a/3$)	75
IV.8 Représentation de 3 hélice plongé dans des surfaces de révolutions, à gauche l'hélice plonger dans le cône (sa projection dans le plan correspond à une spirale logarithmique) $t \mapsto (a \exp(kt) \cos(t), a \exp(kt) \sin(t), a \cot \alpha \exp(kt))$, au centre l'hélice plonger dans le parabolôde de révolution $t \mapsto (a \tan \alpha (\cos(t) + t \sin(t)), a \tan \alpha (\sin(t) - t \cos(t)), \frac{a}{2} \tan^2 \alpha (1 + t^2))$ ou on a pris $k = 1/10$, à droite enfin l'hélice plonger dans la sphère $t \mapsto (a(k \cos(t) \cos(kt) + \sin(t) \sin(kt)), a(k \sin(t) \cos(kt) - \cos(t) \sin(kt)), a\sqrt{1 - k^2} \cos(kt))$ ou on a pris $k = 3/5$	76
V.1 Représentation d'une surface (bleu clair) et de l'homéomorphisme avec un ouvert de \mathbb{R}^2 . Le point P de la définition est en vert, mais le difféomorphisme est valable dans l'intersection de l'ouvert O (jaune) avec S et donc également pour le point courant rouge.	78
V.2 Représentation d'un Cône.	79
V.3 Représentation d'une sphère.	80
V.4 Représentation d'un graphe local. L'ouvert relatif $\tilde{O} \cap S$ est en jaune, les vecteurs $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ sont en noir, F est en bleu, P est le point représenté en noir et S en vert.	81
V.5 À gauche représentation d'un tore $a = 1$ et $b = 0.3$	83
V.6 À gauche représentation d'un ellipsoïde d'axe $a = 1$, $b = 0.7$ et $c = 0.4$, à droite un cône elliptique d'axe $a = 1$, $b = 0.7$ et $c = 2.0$	85
V.7 À gauche représentation d'un hyperbolôde à une nappe d'axe $a = 1$, $b = 0.7$ et $c = 2.0$, à droite un hyperbolôde à deux nappes d'axe $a = 1$, $b = 0.7$ et $c = 2.0$	86
V.8 À gauche représentation d'un parabolôde hyperbolique d'axe $a = 1$ et $b = 0.7$, à droite un parabolôde elliptique d'axe $a = 1$ et $b = 0.7$	86
V.9 À gauche représentation d'un cylindre elliptique d'axe $a = 1$ et $b = 0.7$, au centre une cylindre parabolique $a = 1$, à droite un cylindre hyperbolique d'axe $a = 1$ et $b = 0.7$	87
V.10 Représentation (rouge) d'un ellipsoïde d'axe $a = 3$ et $b = 2.4$ et $c = 2$, en bleu le plan tangent et en noir la base naturelle, l'espace normal est porté par le vecteur rouge, on remarque que le vecteur tangent à la courbe tracé sur la surface appartient au plan tangent.	88
V.11 Transformation isométrique de la Caténoïde en Hélicoïde	92
V.12 À gauche schématisation de la transformation de Mercator, on commence par projeter la sphère sur la surface de révolution de nappe paramétrée $(\theta, \phi) \mapsto \lambda(\theta)\mathbf{u}_z + \lambda(\theta) \tan \theta u(\phi)$ (vert), puis on projette sur le cylindre droit en noir. Au centre, la transformation de Mercator en rouge d'un petit chemin inclus dans la sphère en bleu. À droite, représentation de la carte terrestre dans la transformation de Mercator.	94
V.13 Évolution du plongement dans \mathbb{R}^3 d'une section $t = \text{Cst}$ d'horizon d'un trou noir à mesure que le spin $a = 0, \frac{\sqrt{7}}{8}, \frac{\sqrt{7}}{4}, 1.05 \frac{\sqrt{7}}{8}$ grandit, on obtient la valeur limite pour la troisième figure en partant de la gauche dans ce cas les pôles sont des points planaires (voir plus loin). Pour des valeurs supérieures seules des parties de l'horizon sont plongeables.	95

V.14 Représentation d'un ruban de Möbius	98
V.15 Surface d'Enneper	100
V.16 Représentation de la courbure normale et géodésique. Le vecteur tangent \mathbf{T} est en rouge, tandis que le vecteur courbure globale $\boldsymbol{\kappa}$ est en vert. La courbure géodésique $\nabla_{\mathbf{T}} \mathbf{T}(\gamma(s))$ est en noir et la courbure normale $II_P(\mathbf{T}, \mathbf{T})\mathbf{n}(P)$ en bleu.	103
V.17 Représentation des courbes paramétrées construites, on a P en noir, $C_1(\epsilon_1)$ en vert, $\tilde{C}_2(\epsilon_2)$ en mauve, $C_2(\epsilon_1)$ en rouge et $\tilde{C}_1(\epsilon_2)$ en bleu.	106
V.18 Représentation de l'homéomorphisme f construit lors de la preuve. Le demi-disque et le disque sont en rose pâle quand \mathcal{U}_1 est en jaune et $\mathcal{U}_{1/2}$ en vert. La frontière de \mathcal{U}_1 à été représentée en noir. Les points $(0, 0)$ sont les petits points noirs. Les flèches noires symbolisent le passage homéomorphe entre l'espace en jaune et l'espace en vert.	109
V.19 Représentation de l'homéomorphisme h (flèche rouge) et h_t (flèche bleu) construit lors de la preuve. Le demi-disque est en rose pâle $(t, 0)$ et $\gamma(t)$ sont représentés par des points rouge cerclés de noir. Les ensembles W_t et \tilde{D}_t sont en bleu. La surface S est représentée en vert.	110
V.20 Surface comprise entre les bras spiraux définis par l'équation polaire $\rho : \theta \mapsto \theta/\sqrt{1 + \theta^2}$. Posons $\gamma : \theta \mapsto \rho(\theta)\mathbf{u}(\theta)$. La surface est l'image de l'application $\varphi : (t, \theta)[0, 1] \times [-\pi, +\infty[\mapsto t\gamma(\theta) - (1-t)\gamma(\theta - \pi)$, on comprend que le bord est constitué d'une courbe paramétrée bornée de longueur infinie. L'ensemble $\bar{S} - S$ est constitué du bord et du cercle unité qui n'appartient pas à la surface : celle-ci n'est pas fermée donc pas compacte.	111
V.21 Représentation de la configuration de la démonstration, en mauve de plus en plus prononcé l'ensemble $\mathcal{U}_{+, \epsilon}$ à mesure que ϵ décroît. Le cercle \mathcal{C} est ici représenté en bleu, quand les termes de x_n sont en rouge. On remarque que pour $\epsilon < m$ l'ensemble $\mathcal{U}_{+, \epsilon}$ se retrouve forcément partitionné entre ses points à l'intérieur et ses points à l'extérieur du cercle. Il ne peut plus être connexe par arc.	113
VI.1 A droite, application de recollement entre $\phi_i(\mathcal{V}_i \cap \mathcal{V}_j)$ et $\phi_j(\mathcal{V}_i \cap \mathcal{V}_j)$. A gauche, Représentation d'un atlas de \mathcal{M}	116
VI.2 Nous avons fait une représentation plus métaphorique pour illustrer le choix de ce nom. L'idée est de disposer d'une liste finie exhaustive de cartes qui recouvrent toute la variété. Si le nombre de cartes est infini mais dénombrables, il faut simplement imaginer un livre avec un nombre infini de pages.	117
VI.3 Représentation de la base naturelle. On voit que les vecteurs de la base naturelle sont tangents aux lignes des coordonnées.	119
VI.4 Représentation de l'espace tangent en un point comme classe d'équivalence de courbes régulières admettant une tangente en un point.	119
VI.5 Coordonnées sphériques. En rouge, les vecteurs de la base naturelle. On voit qu'ils ne sont pas égaux aux vecteurs de la base orthonormée locale.	122
VI.6 Action des formes coordonnées sur la variation locale d'un arc paramétré. Au vecteur $\overrightarrow{d}P$, elles associent les variations induites sur chaque coordonnée de la carte choisie.	123
VI.7 Coordonnées covariantes et contravariantes dans la base naturelle quand la métrique est définie positive. Quand celle-ci n'est pas orthonormée, les coordonnées covariantes et contravariantes ne sont pas égales. Quand la métrique est de signature $(-, +, +, +)$, l'orthonormalité de la base n'assure même pas l'égalité des coordonnées covariantes et contravariantes en raison de la géométrie hyperbolique qui lie l'espace et le temps (voir tout bon livre de relativité restreinte).	128
VI.8 Illustration de l'action du flot sur un point de \mathcal{M} . Le flot d'un champ de vecteurs transporte un point le long des lignes de champ de ce vecteur.	165
VI.9 Représentation de D_{+t} et D_{-t} pour un flot (ou un champ de vecteurs) donné. Pour simplifier la représentation, on a choisi \mathcal{M} comme étant un ouvert de \mathbb{R}^2	166
VI.10 Action du flot sur l'espace tangent. Il faut prendre sa différentielle.	167
VI.11 Nous représentons le portrait de phase du pendule simple (pour cela il suffit de tracer les différentes courbes des niveaux d'énergie, à savoir $J^2/2 - \cos \theta = \text{cst}$). nous faisons un zoom sur les deux points d'équilibre. Nous y retrouvons les hyperboles pour le point hyperbolique (à droite) et les cercles pour le point elliptique (à gauche). Si nous avions pris des unités "physiques", il y aurait une affinité sur chacun des axes, ce qui justifierait la terminologie elliptique du point. La dynamique des séparatrices et de leur destruction lorsqu'on ajoute une perturbation non intégrable sera étudiée en M2 DSG.	169

Table des figures

Liste des tableaux

Avant-propos

Ce cours a été rédigé dans le but d'offrir aux étudiants de la première année du Master Astronomie et Astrophysique un support leur permettant d'étudier, de travailler et de réviser à la maison. Il ne dispense cependant pas d'assister au cours ..

L'objet de ce cours sera d'apporter aux étudiants de l'observatoire des bases solides et *rigoureuse* concernant les mathématiques qu'ils utilisent et utiliserons dans la poursuite de leurs études. Les auteurs pensent en particuliers à la géométrie différentiel et du calcul sur les variétés, dont la compréhension du cadre théoriques et des outils est un véritables enjeu concernant l'apprehension géométrique de la relativité général qu'ils auront à étudier l'an prochain, mais également beaucoup aux équations différentiel et aux calculs variationnel souvent abordé de façon intuitive et peu rigoureuse par les physiciens.

Les étudiants de l'observatoire provenant d'horizons différents, ils nous semblent approprié d'unir les connaissances en mathématiques des étudiants du M1. Cette unification procédera en plusieurs temps, nous commencerons par rappeler les élément de base concernant l'algèbre, la topologie des espaces normés et du calcul différentiel. Un certain nombre de rappelles et d'approfondissement seront également donné sur les courbes régulières, ces approfondissement sur les courbes servirons de prolégomènes à ce qui suivra sur les surfaces régulières. Nous verrons ensuite un certain nombres d'éléments concernant l'algèbre et le calcul extérieur. Enfin nous aborderons les équations différentiel pour finir avec le calculs variationnel.

Nous encourageons les étudiants à contacter les auteurs de ce texte s'ils y trouvent coquille et/ou approximations, mais également s'ils souhaite y apporter des améliorations. Nos adresse mails sont : leo.bernu@obspm.fr et loic.chantry@obspm.fr. Nous serons très heureux d'améliorer ainsi la qualité du présent document.

Conventions et notations

Nous noterons les scalaires, les nombres à l'aide de symboles latins a, b, c, \dots ou grecs $\alpha, \beta, \gamma, \dots$ majuscules et minuscules. Les vecteurs (appartenant principalement à \mathbb{R}^3) seront notés à l'aide de lettres, de symboles latins ou grecs mis en gras ($\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \dots$).

On notera \mathcal{E}_3 l'espace affine à 3 dimensions ordinaires, que nous munirons du produit scalaire ordinaire $\mathbf{u} \cdot \mathbf{v}$. Considérons $(\mathbf{e}_i)_{i=1,2,3}$ une base orthonormale de \mathcal{E}_3 , alors nous noterons $(a^j)_{j=1,1,2,3}$ les coordonnées dans la base $(\mathbf{e}_i)_{i=1,2,3}$ (le choix de la base dans laquelle sont écrites les composantes utilisées sera toujours précisé s'il y a une ambiguïté). Les vecteurs de base sont indexés "en bas" tandis que les composantes des vecteurs seront indexés "en haut". Ce choix de notation est justifié par la notion de covariance qui ne sera pas abordée ici. Nous adopterons également la convention d'Einstein, c'est-à-dire que nousoublierons le symbole de somation dès qu'apparaîtra un indice "en haut" et "en bas",

Convention ordinaire	Convention d'Einstein
$\mathbf{v} = \sum_{j=1}^3 v^j \mathbf{e}_j$	$\mathbf{v} = v^j \mathbf{e}_j$

(1)

on parle également de contraction des indices. Nous noterons $\partial_i = \frac{\partial}{\partial x^i}$ la dérivation partielle par rapport à la i -ème coordonnée.

On notera $\delta_{ij} = \delta_j^i = \delta^{ij} = \delta_i^j$ le symbole de Kronecker valant 1 si $i = j$ et 0 sinon. On notera ϵ_{ijk} le symbole de Levi-Civita valant 0 si (i, j, k) n'est pas une permutation de $(1, 2, 3)$, valant 1 pour une permutation paire et -1 pour une permutation impaire.

Une égalité valant pour une définition sera noté $\hat{=}$.

Rewrite this.

CHAPITRE I

Espaces vectoriels.

Sommaire

I.1	Espace vectoriel	5
I.2	Applications linéaires	15
I.3	Réduction des endomorphismes	24

L'objet de ce chapitre est d'introduire les notions essentielles d'espace vectoriel et affine, de norme et de distance, ainsi que quelques rappels concernant la diagonalisation des endomorphismes. Ces notions de base serviront de cadre et d'outils pour la suite du cours.

Nous souhaitons remercier VOEDTS 2002 dont les documents de cours nous ont beaucoup aidé lors de la rédaction de ce chapitre.

I.1 Espace vectoriel

Les espaces vectoriels sont des ensembles d'éléments munis d'une structure opératoire simple permettant de faire des *combinaison linéaire*. Faire une combinaison linéaire sur les éléments d'un espace vectoriel signifie construire à partir d'élément de l'espace vectoriel et d'élément d'un *corps commutatif* de référence, un autre élément de l'ensemble. Le corps commutatif en question \mathbb{K} est attaché à la structure de l'espace vectoriel, on parle de \mathbb{K} -espace vectoriel (en abrégé \mathbb{K} -EV). Ici il s'agira principalement du corps des nombres réels \mathbb{R} ou des nombres complexe \mathbb{C} , mais cela pourrait tout aussi bien être les nombres rationnel \mathbb{Q} ou encore les corps finis de type $\mathbb{Z}/p\mathbb{Z}$. Néanmoins l'utilisation des corps finis peut poser problème pour certaine démonstration en particulier pour $\mathbb{Z}/2\mathbb{Z}$. Sauf si cela est précisé explicitement, le corps utilisé sera $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

I.1.1 Structure d'espace vectoriel

Rappel et définition sur les structures d'ensembles

Commençons par rappeler la définition de groupe (abélien).

Définition I.1.1 (Groupes abélien (G, \star)). *La donnée d'un ensemble d'élément G et d'une loi de composition \star forme un groupe si ils vérifient :*

- \star est interne : $\forall a, b \in G, a \star b \in G$,
- \star est associative : $\forall a, b, c \in G, (a \star b) \star c = a \star (b \star c)$,
- G possède un élément neutre pour \star : $\exists e \in G$ tel que $\forall a \in G, a \star e = e \star a = a$,
- l'existence d'un élément symétrique dans G pour chaque élément de G : $\forall a \in G, \exists b$ qu'on appelle le symétrique de a par \star dans G tel que $a \star b = b \star a = e$,
- (Si la loi \star est commutative, alors (G, \star) est abélien : $\forall a, b \in G, a \star b = b \star a$.)

Un sous ensemble H de G est un sous-groupe si (H, \star) est un groupe.

Exemple I.1.2. L'ensemble des entiers relatifs \mathbb{Z} munie de l'addition $+$ forme un groupe. Les fractions rationnel \mathbb{Q}_* muni de la multiplication ordinaire \times est un groupe.

Remarque I.1.3.

- L'élément neutre et le symétrique sont unique. En effet, supposons qu'il y ai e_1 et e_2 , deux éléments neutre alors $\forall a \in G, a * e_1 = a = a * e_2$, l'existence de l'opposé et l'associativité de a implique directement $e_1 = e_2$. Pour l'unicité de l'opposé, supposons à nouveau qu'il existe $b_1, b_2 \in G$ opposé de $a \in G$, alors $b_1 = b_1 * e = b_1 * (a * b_2) = (b_1 * a) * b_2 = e * b_2 = b_2$.
- Lorsque la loi de composition est "multiplicative", on parlera d'inverse pour le symétrique, on le notera $b = a^{-1}$, le neutre sera noté 1_G . Lorsqu'elle est "additive" on parlera d'opposée pour le symétrique, on le notera $b = -a$, le neutre sera noté 0_G .
- Tout sous groupe $(H, *)$ de $(G, *)$ contient l'élément neutre.
- Pas besoin de vérifier l'associativité pour vérifier que $(H, *)$ sous groupe de $(G, *)$, on n'a uniquement besoin de vérifier, l'internalité et l'existence de l'inverse dans l'ensemble.

Exercice 1. — Groupe et sous groupes finis.

On admettra les théorèmes suivant :

La division euclidienne.

Soit $a, b \in \mathbb{Z} \times \mathbb{Z}_*$ alors $\exists ! q, r \in \mathbb{Z} \times \llbracket 0, |b| \rrbracket$ tel que $a = qb + r$.

Le théorème de Bezout :

Les entiers a_1, a_2, \dots, a_n sont premiers entre eux si et seulement s'il existe des entiers relatifs x_1, x_2, \dots, x_n tels que $1 = x_1 a_1 + \dots + x_n a_n$.

Le calculs des PGCD et PPCM en fonction de la décomposition en facteur premier :

$\forall a \in \mathbb{N}$ on note $a = \prod_{p \in \mathcal{P}} p^{\alpha_p(a)}$ la décomposition en facteur premier de a . La suite $p \rightarrow \alpha_p(0)$ est identiquement nul à partir d'un rang. Soit $(a_i)_{i=1,\dots,n}$ une famille finie d'entier. Alors :

$$\left\{ \begin{array}{l} \bigwedge_{i=1 \dots n} a_i = \prod_{p \in \mathcal{P}} p^{\min_i \{\alpha_p(a_i)\}} \\ \bigvee_{i=1 \dots n} a_i = \prod_{p \in \mathcal{P}} p^{\max_i \{\alpha_p(a_i)\}} \end{array} \right. \quad (\text{I.1})$$

- a. Le but de cette question est de montrer que les sous-groupes de $(\mathbb{Z}, +)$ sont de la forme $(m\mathbb{Z}, +)$ où $m \in \mathbb{Z}$.
 - (i) Montrer que $(\mathbb{Z}, +)$ est un groupe
 - (ii) Montrer que $\forall m \in \mathbb{N}$, $(m\mathbb{Z}, +)$ est un sous-groupe de $(\mathbb{Z}, +)$
 - (iii) Soit $(G, +)$ un sous groupe de $(\mathbb{Z}, +)$, montrer qu'il existe $m \in \mathbb{Z}$ tel que $m\mathbb{Z} \subset G$ (on posera s'il existe $m = \min(\mathbb{N}_* \cap G)$)
 - (iv) Soit $g \in G$, poser la division euclidienne de g par m , conclure.
- b. Soit (G_1, \star) et (G_2, \odot) deux groupes, un morphisme de groupe entre G_1 et G_2 est une application $\varphi : G_1 \rightarrow G_2$ vérifiant $\forall x, y \in G_1 \quad \varphi(x \star y) = \varphi(x) \odot \varphi(y)$. On appelle noyau de φ l'image réciproque de $\{1_{G_2}\}$, on le note $\ker \varphi = \varphi^{-1}(\{1_{G_2}\})$. Le but de cette question est de prouver que le noyau d'un morphisme est un sous groupe de l'espace de départ
 - (i) Soit $\varphi : G_1 \rightarrow G_2$ un morphisme, montrer que $\varphi(1_{G_1}) = 1_{G_2}$
 - (ii) Montrer que l'image d'un sous-groupe H_1 de G_1 est un sous-groupe H_2 de G_2 . En particulier que $\text{im } \varphi$ est un sous-groupe.
 - (iii) Montrer que l'image réciproque d'un sous groupe H_2 de G_2 est un sous groupe H_1 de G_1 . En particulier que $\ker \varphi$ est un sous-groupe.
- c. Il s'agit ici de montrer que les sous-groupes finis de (\mathbb{C}_*, \times) sont les sous-groupes racine de l'unité. Pour $z \in \mathbb{C}$, on note $|z|$ le module de z .
 - (i) Soit $\mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\}$, on le muni de la multiplication ordinaire sur \mathbb{C} . Montrer que (\mathbb{U}, \times) est un sous-groupe.
 - (ii) Soit (G, \times) un sous groupe *fini* montrer que $G \subset \mathbb{U}$.
 - (iii) Pour $n \in \mathbb{N}_*$, on note $\mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}$ les racines n -ième de l'unité, montrer qu'il s'agit d'un sous-groupe de \mathbb{U} . Montrer que $\mathbb{U}_n = \{\exp\left(\frac{2ik\pi}{n}\right) \mid k \in \llbracket 0, n \rrbracket\}$.
 - (iv) Soit (G, \times) un sous groupe finis de \mathbb{U} , soit pour tout $z \in G$, posons $\varphi_z : \mathbb{Z} \rightarrow G, k \mapsto z^k$. Montrer que φ_z est un morphisme.

- (v) Soit G sous-groupes finis de (\mathbb{U}, \times) . On pose $G = \{z_1, z_2, \dots, z_n\}$, montrer que $\exists (n_i) i = 1 \dots n \in \mathbb{N}_*$ tel que $\text{im } \varphi_{z_i} = \mathbb{U}_{n_i}$.
- (vi) Posons $m = \bigvee_{i=1}^n n n_i$, En déduire que $G \subset \mathbb{U}_m$.
- (vii) Montrer que les $a_i = m/n_i$ sont premiers entre eux. En déduire que $\exists (b_k)_{k=1 \dots n}$ tel que $1 = \sum_{k=0}^n a_i b_i$
- (viii) En déduire $G = \mathbb{U}_m$

Correction de l'exercice 1.
a.

- (i) On vérifie rapidement que $+$ est interne, associative, inversible et commutative dans \mathbb{Z}
- (ii) La seule petite difficulté ici est de vérifier que $+$ est interne dans $m\mathbb{Z}$, soit $a, b \in m\mathbb{Z}$ alors $\exists \tilde{a}, \tilde{b} \in \mathbb{Z}$ tel que $a = m\tilde{a}$ et $b = m\tilde{b}$ et donc $a + b = m\tilde{a} + m\tilde{b} = m(\tilde{a} + \tilde{b}) \in m\mathbb{Z}$.
- (iii) $(G, +)$ est un sous groupe, si $G = \{0\}$ alors il suffit de prendre $m = 0$, sinon on peut poser $m = \min(\mathbb{N}_* \cap G)$. Par construction $m \in G \setminus \{0\}$, donc $\forall k \in \mathbb{Z} km = \underbrace{m + m + \dots + m}_{k \text{ fois}} \in G$, on a donc $m\mathbb{Z} \subset G$.
- (iv) On pose la division euclidienne, $g = qm + r$ (ou $q \in \mathbb{Z}$ et $r \in \llbracket 0, m \rrbracket$) ce qui se réécrit $r = g - qm$ et donc $r \in G \cap \llbracket 0, m \rrbracket = \{0\}$ et donc $r = 0$. Ce qui permet d'écrire $g = qm$ et donc $G \subset m\mathbb{Z}$. Ce qui permet de conclure : $G = m\mathbb{Z}$

b.

- (i) φ est un morphisme on a donc, $\varphi(1_{G_1}) = \varphi(1_{G_1} \star 1_{G_1}) = \varphi(1_{G_1}) \odot \varphi(1_{G_1})$ et donc $\varphi(1_{G_1}) = 1_{G_2}$, puisque G_2 est un groupe.
 - (ii) Posons $H_2 = \varphi(H_1)$, on a $\forall a, b \in H_2 \exists \tilde{a}, \tilde{b} \in H_1$ tel que $a = \varphi(\tilde{a})$ et $b = \varphi(\tilde{b})$,
 - interne $a \odot b = \varphi(\tilde{a}) \odot \varphi(\tilde{b}) = \varphi(\tilde{a} \star \tilde{b}) \in H_2$, puisque H_1 sous-groupe
 - inverse H_1 sous-groupe et donc $\tilde{a}^{-1} \in H_1$ et donc $\varphi(\tilde{a}^{-1}) \in H_2$ de plus, $a \odot \varphi(\tilde{a}^{-1}) = \varphi(\tilde{a} \star \tilde{a}^{-1}) = \varphi(1_{G_1}) = 1_{G_2}$, d'où l'existence de l'inverse.
 - (iii) Posons $H_1 = \varphi^{-1}(H_2) = \{a \in | \varphi(a) \in H_2\}$. $\forall a, b \in H_1$
 - interne : $\varphi(a \star b) = \varphi(a) \odot \varphi(b) \in H_2$ (car H_2 sous groupe de G_2) et donc $a \star b \in H_1$
 - inverse : Puisque (G_1, \star) est un groupe, $\exists a^{-1}$ tel que $a \star a^{-1} = 1_{G_1}$ en composant cette identité par φ et en utilisant le résultat de la précédente question, on obtient $\varphi(a) \odot \varphi(a^{-1}) = 1_{G_2}$ et donc $\varphi(a^{-1}) = \varphi(a)^{-1} \in H_2$ car $\varphi(a) \in H_2$ et H_2 sous groupe et donc $a^{-1} \in H_1$.
- Ainsi (H_1, \star) sous groupe. Puisque $\ker \varphi = \varphi^{-1}(\{1_{G_2}\})$, $\ker \varphi$ est une sous groupe de G_1 .

c.

- (i) Seul l'internalité et l'inverse ne sont pas trivial, elles découlent du fait que $\forall z, z' \in \mathbb{C}$ on a $|z \times z'| = |z||z'|$
- (ii) Soit $z \in G$, on suppose $|z| \neq 1$ alors $H = \{z^k | k \in \mathbb{Z}\}$ est un sous-groupe contenu dans G . Les éléments de H sont deux à deux distincts, donc G est de cardinal infini \Rightarrow contradiction. Donc $|z| = 1$ et donc $G \subset \mathbb{U}$.
- (iii) $\forall z \in \mathbb{U}_n$, on a $z^n = 1$ et donc $|z^n| = |z|^n = 1$ et donc $|z| = 1$, on a donc $\mathbb{U}_n \subset \mathbb{U}$. Montrons que c'est un sous-groupe. L'internalité et l'inverse sont conséquence du fait que \times est commutative. Posons $U_n = \{\exp\left(\frac{2ik\pi}{n}\right) | k \in \llbracket 0, n \rrbracket\}$, les propriétés de l'exponentiel complexe montre rapidement que $U_n \subset \mathbb{U}_n$. Soit $z \in \mathbb{U}_n$, alors puisque $\mathbb{U}_n \subset \mathbb{U}$ $\exists \alpha \in [0, 2\pi[$ tel que $z = \exp(i\alpha)$ et donc $n\alpha = 0[2\pi]$ et donc $\exists k \in \mathbb{Z}$ tel que $\alpha = 2k\pi/n$. Posons la division Euclidienne de k par n , $\exists q, r$ tel que $k = qn + r$ et $z = \exp(i2k\pi/n) = \exp(i2(qn+r)\pi/n) = \exp(i2r\pi/n) \in U_n$ puisque $r \in \llbracket 0, n \rrbracket$. D'où $\mathbb{U}_n \subset U_n$ ce qui permet d'écrire $\mathbb{U}_n = U_n$.
- (iv) $\forall z \in G, \forall i, j \in \mathbb{Z}, \varphi_z(i+j) = z^{i+j} = z^i z^j = \varphi_z(i)\varphi_z(j)$, ce qui montre que φ_z est un morphisme de $(\mathbb{Z}, +)$ vers (G, \times) .
- (v) $\forall i \in \llbracket 1, n \rrbracket$, φ_{z_i} est un morphisme, donc $\ker \varphi_{z_i}$ sous-groupe de $(\mathbb{Z}, +)$ et donc $\exists n_i \in \mathbb{N}$ tel que $\ker \varphi_{z_i} = n_i \mathbb{Z}$. Supposons $\exists n_i = 0$, alors $\ker \varphi_{z_i} = \{0\}$ et donc $\{\varphi_{z_i}(j) | j \in \mathbb{Z}\} = \text{im } \varphi_{z_i} \subset G$ sont deux à deux distincts, ce qui est impossible puisque G est finis. Donc $\forall i \in \llbracket 1, n \rrbracket, n_i \in \mathbb{N}_*$. Par construction $\varphi_{z_i}(n_i) = z_i^{n_i} = 1$ et donc $z_i \in \mathbb{U}_{n_i}$.

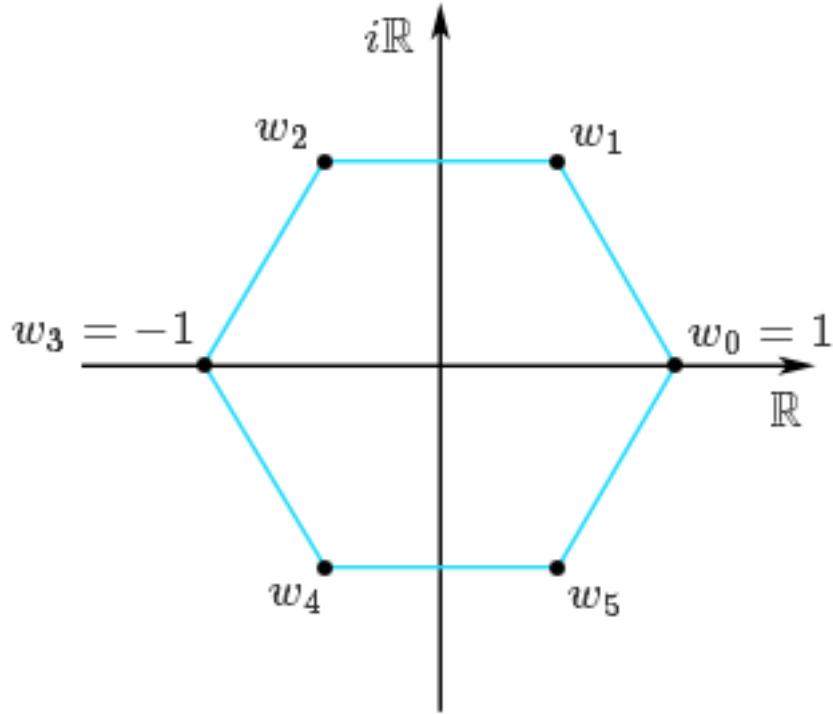


FIGURE I.1 – Représentation des racines sixième de l'unité dans le plan complexe (source : mathraining.be)

Ce qui implique également, $\text{im } \varphi_{z_i} \subset \mathbb{U}_{n_i}$, or $\forall k, l \in [\![0, n_i]\!]$, avec $k \neq l$ on a $0 < |k-l| < n_i$ et donc $z_i^{k-l} \neq 1$ ainsi $\varphi_{z_i}(k) \neq \varphi_{z_i}(l)$ ce qui implique que $\text{Card im } \varphi_{z_i} = n_i$, finalement $\text{im } \varphi_{z_i} = \mathbb{U}_{n_i}$.

- (vi) Simple conséquence du fait que si $k|l$ $\mathbb{U}_k \subset \mathbb{U}_l$, en effet $\forall z \in G$, $\exists ! i$ tel que $z = z_i$ et (par construction) $n_i|m$ et donc $z = z_i \in \mathbb{U}_{n_i} \subset \mathbb{U}_m$ et donc $G \subset \mathbb{U}_m$.
- (vii) On a $a_i = m/n_i = \prod_{p \in \mathcal{P}} p^{\max_k \{\alpha_p(a_k)\} - \alpha_p(n_i)}$ et donc $\bigwedge_{i=1}^n = \prod_{p \in \mathcal{P}} p^{\min_i \{\max_k \{\alpha_p(a_k)\} - \alpha_p(n_i)\}} = \prod_{p \in \mathcal{P}} p^{\max_k \{\alpha_p(a_k)\} - \max_i \{\alpha_p(n_i)\}} = 1$. Les a_i sont premiers entre eux. On peut donc leurs appliquer le théorème de Bezout : $\exists (b_k)_{k=1..n}$ tel que $1 = \sum_{k=0}^n a_i b_i$.
- (viii) Le dernier résultat implique $z_m = \exp(2i\pi \frac{1}{m}) = \exp\left(2i\pi \frac{\sum_{k=0}^n a_i b_i}{m}\right) = \exp\left(2i\pi \sum_{k=0}^n \frac{b_i}{n_i}\right) = \prod_{i=1}^n \underbrace{\exp\left(2i\pi \frac{b_i}{n_i}\right)}_{\in \mathbb{U}_{n_i} \subset G} \in G = \prod_{i=1}^n z_{n_i}^{b_i}$ ce qui permet de rapidement montrer que $\mathbb{U}_m \subset G$, ce qui permet de conclure : $\mathbb{U}_m = G$, on a de plus $m = n$.

Rappelons également la définition d'un corps commutatif.

Définition I.1.4 (Corps commutatif $(K, +, \times)$). *Un corps commutatif est un ensemble K muni de deux lois internes notées $+$ et \times vérifiant les conditions suivantes :*

- $(K, +)$ forme un groupe abélien, dont l'élément neutre est noté 0 ,
- (K_*, \times) forme un groupe abélien multiplicatif, dont l'élément neutre est 1 ,
- la multiplication est distributive pour l'addition (à gauche comme à droite) c'est-à-dire que : $\forall (a, b, c) \in K^3$, $a \times (b + c) = a \times b + a \times c$ et $(b + c) \times a = b \times a + c \times a$,

Exemple I.1.5. Les nombres réels muni de la multiplication et de l'addition ordinaire $(\mathbb{R}, +, \times)$ forme un corps. De même pour l'ensemble des nombres complexes $(\mathbb{C}, +, \times)$.

Exercice 2. — Corps fini $\mathbb{Z}/p\mathbb{Z}$.

Soit $n \in \mathbb{N}_*$ et $a, b \in \mathbb{Z}$, on définit la congruence modulo n la relation binaire défini sur \mathbb{Z} par

$$a \equiv b \pmod{n} \iff a - b \in n\mathbb{Z} \iff n \text{ divise } a - b \quad (\text{I.2})$$

- a. Montrer que la congruence modulo n est une relation d'équivalence
- b. On définit $\bar{a} = \{b \in \mathbb{Z} | b \equiv a \pmod{n}\}$ la classe d'équivalence de a , on note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes d'équivalences de la congruence modulo n . Montrer que $\text{Card}[\mathbb{Z}/n\mathbb{Z}] = n$.
- c. Montrer qu'on peut définir \times sur $\mathbb{Z}/n\mathbb{Z}$ a partir de la multiplication sur \mathbb{Z} par $\forall a, b \in \mathbb{Z}$, $\bar{a} \times \bar{b} \hat{=} a \times b$.
- d. Quel est la condition sur n pour que $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ soit un corps ?

Correction de l'exercice 2.

- a. Une relation d'équivalence vérifie :

- Symétrique : $\forall a, b \in \mathbb{Z}$ tel que $a \equiv b \pmod{n}$ alors $a - b \in n\mathbb{Z} \Rightarrow b - a \in n\mathbb{Z}$ ($(n\mathbb{Z}, +)$ est un groupe) et donc $b \equiv a \pmod{n}$.
- Reflexive : $\forall x \in \mathbb{Z}$ $x - x = 0 \in n\mathbb{Z}$ et donc $x \equiv x \pmod{n}$
- Transitive : Soit $a, b, c \in \mathbb{Z}$ vérifiant :

$$\begin{cases} a \equiv b \pmod{n} \\ b \equiv c \pmod{n} \end{cases} \implies \begin{cases} a - b \in n\mathbb{Z} \\ b - c \in n\mathbb{Z} \end{cases} \implies a - c \in n\mathbb{Z} \implies a \equiv c \pmod{n}$$

La congruence modulo n est donc bien une relation d'équivalence.

- b. Posons $\varphi : \llbracket 0, n \rrbracket \longrightarrow \mathbb{Z}/n\mathbb{Z}$ $i \mapsto \bar{i}$, il s'agit de montrer que φ est bijective. Montrons tout d'abords qu'elle est surjective, soit $K \in \mathbb{Z}/n\mathbb{Z}$, alors prenons $b \in K$ et effectuons la division euclidienne b par n , $\exists! q, r \in \mathbb{Z}$ $\times \llbracket 0, n \rrbracket$ tel que $b = qn + r$, et donc $b \in \bar{r}$, on a donc $K = \bar{b} = \bar{r} = \varphi(r)$. Pour l'injectivité considérons $i, j \in \llbracket 0, n \rrbracket$ tel que $\varphi(i) = \varphi(j) \Rightarrow i \equiv j \pmod{n} \Rightarrow i - j \in n\mathbb{Z} \Rightarrow \exists q \in \mathbb{Z}$ tel que $i = j + qn$, or $i, j \in \llbracket 0, n \rrbracket$ donc $q = 0$ et $i = j$. Ainsi φ est un bijection et donc $\text{Card}(\llbracket 0, n \rrbracket) = \text{Card}(\mathbb{Z}/n\mathbb{Z}) = n$.
- c. La difficulté consiste ici à palier à l'ambiguïté de la définition, en effet soit $a \neq c \in \mathbb{Z}$ tel que $\bar{a} = \bar{c}$, a priori la définition du produit ne garantit pas, $\forall b \in \mathbb{Z}$ que $\bar{a}b = \bar{c}b$. Il faut donc la vérifier pour montrer que cette définition de comporte pas d'ambiguïté, la symétrie du produit dans \mathbb{Z} complétera la démonstration. On a donc $\bar{a} = \bar{c}$, cela implique $c = a + qn$ et donc $cb = ab + qcn$ et donc $\bar{a}b = \bar{c}b$. Ce qui garantit l'existence de l'opération \times sur les classes d'équivalence de la congruence modulo n .
- d. Pour que $(\mathbb{Z}/n\mathbb{Z}, +, \times)$, il faut que $(\mathbb{Z}/n\mathbb{Z}_*, \times)$ soit un groupe. Ici l'internalité, l'associativité et l'existence de l'élément neutre sont simplement conséquence des propriétés de \times sur \mathbb{Z} , par contre l'existence de l'inverse est plus difficile. Supposons que la classe de $i \in \llbracket 0, n \rrbracket$ soit inversible, il existe donc $j \in \llbracket 0, n \rrbracket$ tel que $\bar{i} \times \bar{j} = \bar{1}$ et donc $\bar{ij} = \bar{1}$ ce qui est équivalent à ce qu'il existe $q \in \mathbb{Z}$ tel que $ij = qn + 1$ ce qui est une identité de Bezout et donc l'inversibilité de \bar{i} est équivalente à $i \wedge n = 1$. Donc pour que $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ soit un corps il faut et il suffit que $\forall i \in \llbracket 0, n \rrbracket$ on ai $i \wedge n = 1$, ce qui est équivalent à ce que n soit un nombre premier.

Théorème I.1.6 (Corps fini $\mathbb{Z}/p\mathbb{Z}$). $\mathbb{Z}/p\mathbb{Z}$ est un corps si et seulement si p est un nombre premier

Dans la suite, sauf si cela est précisé, le corps $\mathbb{K} = \mathbb{R}$. Passons enfin à la définition d'un \mathbb{K} -espace-vectoriel.

Définition I.1.7 (\mathbb{K} -espace-vectoriel $(E, +, \cdot)$). $(K, +, \cdot)$ un corps commutatif dont on notera 0_K et 1_K les élément neutre pour l'addition et la multiplication. Alors un ensemble E a une structure de \mathbb{K} -espace-vectoriel si on peut définir

I.1.7.1. Une loi de composition interne $+$, tel que $(E, +)$ soit un groupe abélien, on notera 0_E son élément neutre.

I.1.7.2. Une loi de composition externe $\cdot : \mathbb{K} \times E \longrightarrow E$, $(\lambda, x) \mapsto \lambda \cdot x$ appelé multiplication externe et vérifiant les propriétés :

I. Espaces vectoriels.

- Distributive sur le corps \mathbb{K} : $\forall \alpha, \beta \in \mathbb{K}, \forall x \in E$ on a $(\alpha + \beta).x = \alpha.x + \beta.x$.
- Distributive sur l'espace vectoriel E : $\forall \alpha \in \mathbb{K}, \forall x, y \in E$ on a $\alpha.(x + y) = \alpha.x + \alpha.y$
- $1_{\mathbb{K}}$ est neutre pour la multiplication externe : $\forall x \in E$ $1_{\mathbb{K}}.x = x$
- Vérifie une associativité mixte : $\forall \alpha, \beta \in \mathbb{K}, \forall x \in E$ on a $\alpha.(\beta.x) = (\alpha \times \beta).x$

On appellera vecteurs les éléments d'un espace vectoriel et scalaire le corps \mathbb{K} .

Exemple I.1.8. Soit $(\mathbb{L}, +, \times)$ un corps et $(\mathbb{K}, +, \times)$ un sous corps (On peut imaginer \mathbb{C} et \mathbb{R}) alors \mathbb{L} est un \mathbb{K} -espace vectoriel. Il suffit simplement de construire la loi de composition externe en restreignant la multiplication \times sur $\mathbb{L} \times \mathbb{L}$ à celles sur $\mathbb{K} \times \mathbb{L}$. Et donc \mathbb{C} peut être considéré comme un \mathbb{R} -espace vectoriel, on montrera qu'il est de dimension 2. On peut également construire sur \mathbb{R} une structure de \mathbb{Q} -espace vectoriel mais nous n'utiliserons pas de tel monstre dans la suite du cours.

Exemple I.1.9. L'espace des polynômes $\mathbb{K}[X]$ est un \mathbb{K} -ev.

Exemple I.1.10. Un produit fini de type \mathbb{K}^n est un \mathbb{K} -espace vectoriel.

Exemple I.1.11. L'espace des fonctions $\mathcal{F}(\mathbb{K}, \mathbb{K})$ est un \mathbb{K} -espace vectoriel.

Dans la suite dés que nous désignerons par E un \mathbb{K} -espace vectoriel. On notera en abrégé : E est un \mathbb{K} -ev. Dans la suite I désignera une famille d'indice.

I.1.2 Sous espace vectoriel

Définition I.1.12 (Combinaison linéaire). Soit $x_{i \in I}$ (ici I doit être une famille finie) une famille de vecteurs de E , un vecteur $x \in E$ est dis combinaison linéaire (CL en abrégé) si il existe une famille de scalaire $\alpha_{i \in I}$ tel que

$$x = \sum_{i \in I} \alpha_i x_i \quad (\text{I.3})$$

Puisque la combinaison linéaire, par définition n'implique que des combinaisons finies, et afin de ne pas surcharger l'écriture en introduisant en permanence des sous famille d'indice finie, on précisera souvent "presque tous nul" à propos des scalaire d'une combinaison linéaire. Cela signifiera qu'un nombre fini d'entre eux est non nuls.

Définition I.1.13 (Sous-espace vectoriel). Une partie $F \subset E$ est un sous espace vectoriel si et seulement si F est non vide et est stable par combinaison linéaire d'élément de F :

$$\forall x, y \in F, \forall \lambda \in \mathbb{K} x + \lambda y \in F$$

On notera SEV en abrégé.

Théorème I.1.14. Soit $(F_i)_{i \in I}$ une famille non vide de SEV de E alors $\bigcap_{i \in I} F_i$ est un SEV de E .

La démonstration est laissé à titre d'exercice. Ce théorème est la base de la notion de SEV engendrer par une famille de vecteur.

Théorème I.1.15. Soit $\mathcal{F} = (x_i)_{i \in I}$ un famille de vecteurs d'un \mathbb{K} -EV E , alors il existe un plus petit (au sens de l'inclusion) sous-espace vectoriel $F \subset E$ contenant tout les éléments de la famille \mathcal{F} , on l'appelle sous espace vectoriel engendrer par \mathcal{F} et on l'appelle $F = \text{vect}(\mathcal{F})$

Preuve. On considère $\mathbb{X}(\mathcal{F})$ la famille des sous-espace vectoriel qui contienne contenant la famille \mathcal{F} , cette famille est non vide car $E \in \mathbb{X}(\mathcal{F})$, l'ensemble

$$F = \bigcap_{G \in \mathbb{X}(\mathcal{F})} G$$

est un sous espace vectoriel qui contient tout les vecteurs de \mathcal{F} . De plus il est évidemment le plus petit du point de vue de l'inclusion. ■

Théorème I.1.16. Soit $\mathcal{F} = (x_i)_{i \in I}$ un famille non vide de vecteurs d'un \mathbb{K} -EV E , alors $\text{vect}(\mathcal{F})$ est l'ensemble des combinaison linéaire des vecteur de \mathcal{F}

Preuve. Posons $F = \{x \in E \mid x \text{ est combinaison linéaire de } \mathcal{F}\}$, de façon évidente F est une sous-espace vectoriel de E . En effet par définition, $\forall \lambda \in \mathbb{K}, \forall x, y \in F, \exists (\alpha_i)_{i \in I}, (\beta_i)_{i \in I} \in \mathbb{K}^I$ (ou les α_i et les β_i sont presque tous nul) tel que $x = \sum \alpha_i x_i$ et $y = \sum \beta_i y_i$ et donc $x + \lambda y = \sum (\alpha_i + \lambda \beta_i) x_i \in F$. De plus F contient \mathcal{F} , or puisque tout sous espace vectoriel est stable par combinaison linéaire, tout sous espace vectoriel qui contient \mathcal{F} , contient les combinaisons linéaire des éléments de \mathcal{F} , c'est à dire qu'il contient F . F est donc le plus petit sous espace vectoriel (au sens de l'inclusion) contenant \mathcal{F} , car il est contenue dans tout les autres. ■

Définition I.1.17 (Famille génératrice, Famille équivalente). Une famille \mathcal{F} de vecteur de E est dite génératrice de E si et seulement si,

$$E = \text{vect}(\mathcal{F}).$$

Pareillement, un sous-espace vectoriel F étant lui même un espace vectoriel, on dis que \mathcal{F} est génératrice de F si et seulement si $F = \text{vect}(\mathcal{F})$.

Deux familles de vecteur \mathcal{F} et \mathcal{G} sont dites équivalentes si elle engendre le même sous espace vectoriel,

$$\text{vect}(\mathcal{F}) = \text{vect}(\mathcal{G})$$

Définition I.1.18 (Somme de SEV). Soit F_1, \dots, F_p des sous espace vectoriel de E , on appelle somme de des sous espace vectoriel $(F_i)_{i=1 \dots p}$ le sous espace engendrer par $\bigcup_{i=1}^p F_i$, on note,

$$\sum_i F_i = \text{vect} \left(\bigcup_{i=1}^p F_i \right).$$

Proposition I.1.19. La somme $\sum_i F_i$ est exactement l'ensemble des vecteur x de E admettant au moins une écriture de la forme,

$$x = \sum_i x_i \text{ avec } \forall i x_i \in F_i$$

La démonstration est laissé a titre d'exercice.

On va maintenant s'intéresser à l'unité ou non de la décomposition de la proposition précédentes, cela nous conduira a introduire la notion de famille libre ou lié.

I.1.3 Dépendance et indépendance linéaire

Quand on retire un élément d'une famille de vecteur, alors deux possibilité emerge concernant l'espace engendrer par la famille. Soit l'espace engendrer par la famille amputée est strictement inclus dans l'espace engendrer par la famille complète, soit les sous espace sont égaux. Si la moindre amputation, conduit a un espace plus petit alors la famille est dite libre. A contrario si il existe une amputation qui ne réduit pas la taille du sous espace engendrer alors on dis que la famille est lié, il est alors facile de comprendre que l'élément amputée est combinaison linéaire des éléments de la famille amputée.

Définition I.1.20 (Famille lié).

- Une famille $\mathcal{F} = (x_i)_{i \in I}$ de vecteur est lié si et seulement si elle est réduite au vecteur nul ou si l'un des vecteur est combinaison linéaire des autres, dans ce cas, il existe une sous famille stricte de \mathcal{F} qui engendre $\text{vect}(\mathcal{F})$.
- Une famille \mathcal{F} est libre dans le cas contraire, et donc si elle est génératrice minimal de $\text{vect}(\mathcal{F})$.

Proposition I.1.21. La famille $\mathcal{F} = (x_i)_{i \in I}$ est lié s'il existe une famille de scalaire $(\alpha_i)_{i \in I}$ non tous nul tel que,

$$\sum_{i \in I} \alpha_i x_i = 0_E$$

I. Espaces vectoriels.

Corrolaire I.1.22. Toutes sous famille d'une famille libres est libres. Toute famille contenant une sous famille lié est lié.

On laissera les preuves à titre d'exercice.

Exercice 3. —

- a. Montrer que $f_n(x) = xn$ est libre dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
- b. Montrer que la famille de fonction $g_{n,\alpha}(x) = x^n \exp(\alpha x)$ où $n \in \mathbb{N}$ et $\alpha \in \mathbb{R}$ est libre dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
- c. Montrer que la famille de fonction $h_\alpha(x) = |x - \alpha|$ est libre dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Correction de l'exercice 3.

- a. On considère l'existence d'une relation de liaison,

$$\exists (\lambda_i)_{i \in I} \in \mathbb{R} \quad f = \sum_{j \in I} \lambda_j f_j = 0$$

Alors la fonction f est identiquement nul, de plus en tant que combinaison linéaire de fonction infiniment dérivable elle est également dérivable, on considère alors $n = \min(I)$ et on dérive n fois la relation précédentes, on évalue ensuite en 0 la fonction on obtient $\lambda_n = 0$. On achève la démonstration en effectuant une récurrence.

- b. Considérons une sous famille finie et on suppose qu'il existe une relation de liaison entre ces vecteurs,

$$\exists (\lambda_i)_{i \in I} \in \mathbb{R} \quad g = \sum_i \lambda_i g_{n_i, \alpha_i} = 0$$

Ou on a choisi $i \mapsto (n_i, \alpha_i)$ injective. Alors en notant $\alpha = \max_j \alpha_j$, on note I_0 la sous famille de I tel que $\alpha(i) = \alpha$ (non vide par construction), on multiplie alors la relation de liaison en x par $\exp(-\alpha x)$ et on étudie le comportement à l'infini. On obtient alors $g(x) \exp(-\alpha x) \underset{+\infty}{\equiv} \sum_{i \in I_0} \lambda_i x^i$. Or g est identiquement nul, et donc cela implique qu'à l'infini le polynôme $\sum_{i \in I_0} \lambda_i x^i \rightarrow 0$ ce qu'il n'est possible que pour le polynôme nul. On en déduit que $(\lambda_i)_{i \in I_0} = 0$. Et on recommence avec $I - I_0$.

- c. $\forall a \in \mathbb{R}$ l'espace des fonctions de $\mathbb{R} \mapsto \mathbb{R}$ dérivable en a est un espace vectoriel. On suppose à nouveau qu'il existe une relation de liaison entre ces les h_α , on a donc,

$$\exists (\lambda_i)_{i \in I} \in \mathbb{R} \quad h = \sum_i \lambda_i h_{\alpha_i} = 0 \implies \lambda_0 h_{\alpha_0} = - \sum_{i \neq 0} \lambda_i h_{\alpha_i}$$

A nouveau la famille $i \mapsto \alpha_i$ est choisi injective. Alors la dernière équation implique que $x \mapsto \lambda_0 |x - \alpha_0|$ est dérivable en α_0 , ce qui implique que $\lambda_0 = 0$. Un récurrence immédiate permet de conclure.

Définition I.1.23 (Base). Une famille $\mathcal{B} = (u_i)_{i \in I}$ de vecteur de E est une base de E si et seulement si \mathcal{B} est libre et génératrice de E

Théorème I.1.24. Une famille $\mathcal{B} = (u_i)_{i \in I}$ de vecteur de E est une base de E si et seulement si \mathcal{B} est une famille libre maximal (ou ce qui reviens au même \mathcal{B} est une famille génératrice minimal).

Corrolaire I.1.25. Une famille $\mathcal{B} = (u_i)_{i \in I}$ est une base de E . Alors,

$$\forall x \in E, \exists! (\alpha_i)_{i \in I} \in \mathbb{K}^I \quad (\text{presque tous nul}) \text{ tel que } x = \sum_{i \in I} \alpha_i u_i$$

la famille $(\alpha_i)_{i \in I}$ est appelé famille de coordonnée de x dans \mathcal{B} . En notant $\mathbb{K}^{(I)}$, l'ensemble des famille de scalaire indéxié sur I qui sont presque tous nul. Alors l'application,

$$\begin{array}{ccc} K^{(I)} & \rightarrow & E \\ (\alpha_i)_{i \in I} & \mapsto & \sum_{i \in I} \alpha_i u_i \end{array}$$

est une bijection. C'est un *isomorphisme d'espace vectoriel*.

Exemple I.1.26. La famille X^n est une base de $\mathbb{K}[X]$ des polynômes. Elle est appelé base canonique de $\mathbb{K}[X]$.

Exemple I.1.27. Dans \mathbb{K}^n , la famille $u_{i=1 \dots n} = (0, \dots, i, \dots, 0)$ est également une base appelé base canonique de \mathbb{K}^n .

Théorème I.1.28. Soit F_1, \dots, F_p est une famille de SEV possédant chacune respectivement sa base $\mathcal{B}_1, \dots, \mathcal{B}_p$, ces sous-espaces sont indépendant si et seulement si la famille $\mathcal{B} = \bigcup \mathcal{B}_i$ est libre. Si \mathcal{B} est libre alors elle est une base de $= \bigoplus_{i=1 \dots p} F_i$.

Définition I.1.29 (SEV indépendant). Si F_1, \dots, F_p sont p SEV, alors ils sont indépendant ssi l'application

$$\begin{aligned}\varphi : \quad \prod_{i=1}^p F_i &\rightarrow \sum_{i=1}^p F_i \\ (x_1, \dots, x_p) &\mapsto \sum_{i=1}^p x_i\end{aligned}$$

est bijective (la surjectivité est toujours assurée par construction), cela veut dire que son caractère bijectif est équivalent à son caractère injectif. Et donc l'indépendance est équivalente à montrer que tout vecteur de $x \in \sum_{i=1}^p F_i$ admet une unique décomposition sous la forme,

$$x = x_1 + \dots + x_p \quad \text{avec } \forall i x_i \in F_i$$

On dis aussi que la somme des sous-espaces vectoriel F_1, \dots, F_p est somme direct. On note,

$$\sum_{i=1}^p F_i = \bigoplus_{i=1}^p F_i$$

Avec les même notation, $\forall x \in \bigoplus_{i=1}^p F_i$, x_i est dite composante de $x \in \bigoplus_{i=1}^p F_i$ selon F_i et parallèlement au $(F_k)_{k \neq i}$.

Remarque I.1.30.

- L'application φ de la définition précédente est toujours, linéaire on montre que c'est un isomorphisme d'espace vectoriel.
- Il est équivalent de parler de l'indépendance de la famille de vecteur $(u_i)_{i=1 \dots p}$ que de parler de l'indépendance des la famille des droites vectoriel $D_i = \text{vect}(u_i)$ associées.

Théorème I.1.31. Les sous espace F_1, \dots, F_p sont linéairement indépendant si et seulement si,

$$\forall j \in \{1, \dots, p\} \quad F_j \cap \left(\sum_{i \neq j} F_i \right) = \{0_E\}$$

Preuve. Supposons que les sous-espace soit indépendant, puisque $+$ est commutative dans E on se ramène au cas où $j = p$. Prenons $x \in F_p \cap \left(\sum_{i \neq p} F_i \right)$, alors $x \in \left(\sum_{i \neq p} F_i \right)$ et donc $x = x_1 + x_2 + \dots + x_{p-1}$. On a donc,

$$x = x_1 + x_2 + \dots + x_{p-1} + 0_E = 0_E + \dots + 0_E + x$$

L'indépendance des sous-espaces vectoriel implique l'unicité de la décomposition et donc $x = 0_E$. Réciproquement, supposons $\forall j \in \{1, \dots, p\} \quad F_j \cap \left(\sum_{i \neq j} F_i \right) = \{0_E\}$ et. supposons également qu'on puisse décomposer $x \in \sum_{i=1}^p F_i$ deux. deux manière différentes,

$$x = \sum_{i=1}^p x_i = \sum_{i=1}^p y_i \quad \text{avec pour tout } i \quad x_i \text{ et } y_i \in F_i$$

alors

$$y_k - x_k = \sum_{i=1, i \neq k}^p x_i - y_i \in F_k \cap \left(\sum_{i \neq k} F_i \right) = \{0_E\}$$

et donc $\forall k, x_k = y_k$ et prouve donc l'indépendance des sous espaces. ■

I.1.4 Dimension finie

Définition I.1.32 (Dimension finie). *On dis qu'un espace vectoriel est de dimension finie si d'une famille génératrice on peut extraire une famille génératrice finie.*

Théorème I.1.33. *Dans un espace vectoriel engendré par une famille de cardinal p , toute famille de cardinal strictement supérieur à p est liée.*

Preuve. Elle se fait par récurrence sur p . Commençons donc par $p = 1$, et notons e_1 la "famille" de vecteur engendrant E . Alors tout vecteur u_1, u_2 de E est combinaison linéaire de e_1 . Et donc $\exists \lambda, \mu$ tel que $u_1 = \lambda e_1$ et $u_2 = \mu e_1$. On a donc $\mu u_1 - \lambda u_2 = 0_E$, la famille est liée.

Supposons maintenant le résultat démontré au rang p , soit une famille de $p + 1$ vecteur génératrice de $E_{p+1} = \text{vect}(e_{1 \leq i \leq p+1})$. Soit une famille quelconque de $p + 2$ vecteur $u_{1 \leq i \leq p+2}$ de E_{p+1} , il existe donc des scalaires $(\alpha_{i,j})_{1 \leq i \leq p+1, 1 \leq j \leq p+2}$ tel que $u_i = \sum_{j=1}^{p+1} \alpha_{i,j} e_j$, on construit donc les $p + 1$ vecteurs $(v_j)_{j=2, \dots, p+2}$ en se débarrassant de e_{p+1} dans les CLs.

$$v_j = \alpha_{p+1,1} u_j - \alpha_{p+1,j} u_1 = \sum_{i=2}^{p+1} (\alpha_{p+1,1} \alpha_{i,j} - \alpha_{p+1,j} \alpha_{i,1}) e_i \in \text{vect}(e_i)_{2 \leq i \leq p+1}$$

On pose $F_p = \text{vect}(e_i)_{2 \leq i \leq p+1}$, F_p est engendré par p vecteur, on a $(v_j)_{j=2, \dots, p+2}$ un famille de cardinal $p + 1$ dans F_p , on peut donc appliquer l'hypothèse de récurrence, il existe donc $(\beta_k)_{k=2, \dots, p+2}$ tel que $\sum_{j=2}^{p+2} \beta_j v_j = 0_E$ ce qui implique par construction :

$$\sum_{j=2}^{p+2} \beta_j \alpha_{p+1,1} u_j - \left(\sum_{j=2}^{p+2} \beta_j \alpha_{p+1,j} \right) u_1 = 0_E$$

Ce qui montrer que la famille des $u_{1 \leq i \leq p+2}$ de E_{p+1} est liée. ■

Théorème I.1.34 (Base incomplète). *Si \mathcal{G} est une famille génératrice d'un espace vectoriel E de dimension finie (non réduite à $\{0_E\}$) et si \mathcal{L} une famille libre de E , il est possible de compléter \mathcal{L} à l'aide de vecteurs de G pour obtenir une base de E .*

Preuve. E est de dimension finie on peut donc extraire de \mathcal{G} un famille $\mathcal{G}_F = \{e_1, \dots, e_p\}$ fini engendrant E . On pose $\mathcal{L}_n \hat{=} \{u_1, \dots, u_n\} = \mathcal{L}$, d'après le théorème précédent on a $n \leq p$. Si \mathcal{L}_n n'est pas libre, il existe $x_{n+1} \in E$ tel que x ne soit pas CL de \mathcal{L}_n et la famille $\mathcal{L}_{n+1} \hat{=} \mathcal{L}_n \cup e_{n+1}$ est également libre. Soit elle engendre, E soit on répète l'opération. Il faut un nombre fini d'étape, car la cardinal d'un famille libre est forcément inférieur à p d'après le théorème précédent. ■

Corrolaire I.1.35 (Dimension). *Un espace vectoriel de dimension finie, non réduit à $\{0_E\}$ possède des bases qui sont toutes finies et de même cardinal. On appelle ce cardinal dimension de E , on le note $\dim E$.*

Preuve. On suppose \mathcal{B}_1 et \mathcal{B}_2 deux bases de E , de cardinaux respectifs n_1 et n_2 . Comme \mathcal{B}_1 est génératrice et \mathcal{B}_2 libre le théorème I.1.33 implique $n_2 \leq n_1$, et l'inverse est vrai également. Donc $n_1 = n_2$. ■

Définition I.1.36 (Rang famille de vecteur). *Soit \mathcal{F} une famille de vecteur de E , on dit que \mathcal{F} est de rang fini si et seulement si $\text{vect}\mathcal{F}$ est un sous-espace vectoriel de dimension finie dans ce cas, le rang de \mathcal{A} est égal à la dimension de $\text{vect}\mathcal{F}$, on note $\text{rang}(\mathcal{A}) = \dim \text{vect}\mathcal{A}$*

Théorème I.1.37. *Dans un espace vectoriel de dimension finie, tout sous espace vectoriel possède au moins un supplémentaire.*

Théorème I.1.38. *Soit F et G deux SEV d'un espace E , alors $F + G$ est de dimension fini ssi F et G le sont et alors*

$$\dim(F + G) = \dim E + \dim F - \dim F \cap G$$

Preuve. Si F ou G n'étais de dimension il serai simple de montrer que $F + G$ ne serai pas de dimension fini. Si c'est le cas alors la dimension de $F + G$ serai facile a calculer s'il étais indépendant (Cf : I.1.28). Sinon puisque $F \cap G$ est un sous espace vectoriel de F , on pose F' un supplémentaire (existence I.1.37). On a alors $F + G = (F' + F \cap G) + G = F' + (F \cap G + G) = F' + G = F' \oplus G$. En effet $F' \cap G = (F' \cap F) \cap G = F' \cap (F \cap G) = \{0_E\}$. Donc $\dim F + G = \dim F' + \dim G$ de plus puisque $F = F' \oplus (F \cap G)$, on a $\dim F = \dim F' + \dim(F \cap G)$, ce qui permet de conclure. ■

Exercice 4. —

Soit $\mathbb{L} \subset \mathbb{K}$ un sous corps de \mathbb{K} , et E un \mathbb{K} -ev de dimension finie. Montrer que si \mathbb{K} est un \mathbb{L} -ev de dimension finie alors E est un \mathbb{L} -ev de dimension finie. Et,

$$\dim_{\mathbb{L}} E = \dim_{\mathbb{K}} E \times \dim_{\mathbb{L}} \mathbb{K}$$

Correction de l'exercice 4. Si \mathbb{K} est un \mathbb{L} -ev de dimension finie, on construit $(k_j)_{j=1..p}$ une base du \mathbb{L} -ev K . On note $(e_i)_{i=1..n}$ une base du \mathbb{K} -ev E , on a donc $\forall x \in E, \exists!(\alpha_i)_{i \leq n} \in \mathbb{K}^n$ tel que $x = \sum_{i=1}^n \alpha_i e_i$. Or puisque $\forall i \alpha_i \in \mathbb{K}, \exists!(\gamma_j)_{j=1..p} \in \mathbb{L}^p$ tel que $\alpha_i = \sum_{j=1}^p \gamma_{ij} k_j$ et donc, $\forall x \in E, \exists!(\gamma_{i,j})_{i=1..n, j=1..p} \in \mathbb{L}^{np}$ tel que $x = \sum_{i=1}^n \sum_{j=1}^p \gamma_{ij} k_j e_i$. Et donc $(k_j e_i)_{i=1..n, j=1..p}$ base de E en tant que \mathbb{L} -ev. ■

I.2 Applications linéaires

Les applications linéaires sont des fonctions transportant linéairement les objets d'un espaces vectoriel a un autres ils permettent de transporter des propriété, mais ont également énormement d'application en physique. En effet, ces objets très généraux se rencontre très. souvent, connaitre leurs propriétés et ce qu'on peut faire avec eux est donc très utiles pour le physiciens. L'objet de cette sections sera de présenter tout d'abords les propriétés générales des ces objets puis de voir comment on peut les simplifier.

I.2.1 Généralités

Définition I.2.1 (Fonction, Injection, Surjection). Soit A et B deux ensemble. Un fonction f est un objet associant a chaque élément de $x \in A$ un élément $y = f(x) \in B$. On note

$$\begin{array}{rccc} f & : & A & \rightarrow & B \\ & & x & \mapsto & f(x) \end{array}$$

L'ensemble des fonctions de A vers B est noté $\mathcal{F}(A, B)$.

- Une fonction $f \in \mathcal{F}(A, B)$ est dite injective si et seulement si chaque élément de l'espace d'arrivée est atteint au maximum une fois, on a

$$\forall x, x' \in A \text{ tel que } f(x) = f(x') \implies x = x'$$

- Une fonction $f \in \mathcal{F}(A, B)$ est dite surjective si et seulement si chaque élément de l'espace d'arrivée est atteint au moins une fois, on a

$$\forall y \in B \exists x \in A \text{ tel que } y = f(x)$$

Une fonction $f \in \mathcal{F}(A, B)$ est dite bijective si et seulement si chaque élément de l'espace d'arrivée est une fois et une seule, on a

$$\forall y \in B \exists!x \in A \text{ tel que } y = f(x)$$

Remarque I.2.2. On parlera indifféremment de fonction ou d'application

Une application linéaire est une fonction entre deux \mathbb{K} -espace vectoriel E et F vérifiant des propriétés de linéarité

Définition I.2.3 (Application linéaire). Soit E et F deux \mathbb{K} -espace vectoriel une application $u : E \rightarrow F$ est dites linéaire si et seulement si elle vérifie :

$$\forall \lambda \in \mathbb{K} \forall x, y \in E \quad f(x + \lambda y) = f(x) + \lambda f(y)$$

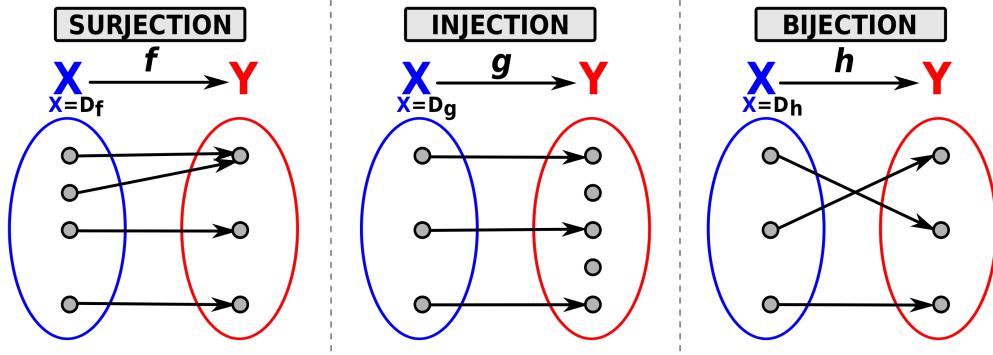


FIGURE I.2 – Représentation des fonctions injective, surjective et bijective (source : Wikipedia)

- Si u est bijection alors on parle d'un isomorphisme, si une telle application existe entre E et F on dit que E et F sont isomorphes.
- Lorsque $E = F$ on parle d'endomorphisme.
- Pour un endomorphisme bijectif on parle d'automorphisme. On note l'ensemble des automorphismes sur E par $\mathcal{GL}(E)$.

L'espace des applications linéaires de E vers F est noté $\mathcal{L}(E, F)$.

Remarque I.2.4. Une application linéaire est un morphisme de groupe de $(E, +)$ vers $(F, +)$, on parlera indifféremment de morphisme ou d'application linéaire.

Théorème I.2.5. Soit $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$ alors l'application composée est notée $u \circ v \in \mathcal{L}(E, G)$. Si u et v sont des isomorphismes alors $u \circ v$ et u^{-1} le sont aussi. En particulier les automorphismes munis de la composition forment un groupe.

Quand les espaces vectoriels de départ et d'arrivée possèdent des bases, alors une application linéaire est déterminée par un ensemble de scalaires.

Théorème I.2.6. Soit $\mathcal{B}_E = (e_i)_{i \in I}$ une base de E et $\mathcal{B}_F = (f_i)_{i \in I}$ une famille de vecteurs de F , alors il existe un unique morphisme $u : E \rightarrow F$ vérifiant

$$\forall i \in I \quad u(e_i) = y_i$$

Un morphisme est entièrement déterminé par la donnée de l'image des vecteurs d'une base de l'espace de départ.

Théorème I.2.7. Soit $E = E_1 \oplus E_2$ alors si on se donne $u_1 \in \mathcal{L}(E_1, F)$ et $u_2 \in \mathcal{L}(E_2, F)$, alors il existe une unique $u \in \mathcal{L}(E, F)$ tel que $u|_{E_1} = u_1$ et $u|_{E_2} = u_2$

Théorème I.2.8. Soit une application linéaire $u : E \rightarrow F$, l'image direct d'un sous espace vectoriel de E par u est un sous espace vectoriel de F et l'image réciproque d'un sous espace vectoriel de F est un sous espace vectoriel de E .

Définition I.2.9 (Image et Noyau). Soit une application linéaire $u : E \rightarrow F$

- On appelle noyau de u , l'espace vectoriel de E qui est image réciproque du sous-espace vectoriel $\{0_F\}$, on le note $\ker u = u^{-1}(0_E) = \{x \in E | u(x) = 0_F\}$
- On appelle image de u le sous espace vectoriel de F image direct de E par u , on le note $\text{im } u = \{y \in F | \exists x \in E, u(x) = y\}$, si elle existe on appelle rang de u la dimension de $\text{im } u$, on note $\text{rang } u = \dim \text{im } u$.

Théorème I.2.10. Soit une application linéaire $u : E \rightarrow F$, alors

- u est injective si et seulement si $\ker u = \{0_E\}$
- u est surjective si et seulement si $\text{im } u = F$

I.2.2 Théorèmes du rang et dimension

Lemme I.2.11 (Lemme du rang). Soit une application linéaire $u : E \rightarrow F$ et S un supplémentaire de $\ker u$ dans E , la restriction de u à S induit un isomorphisme $u|_S$ de S vers $\text{im } u$.

Preuve. On a de façon évidente $\ker u|_S = \ker u \cap S = \{0_E\}$ puisque $\ker u \oplus S = E$ et donc $u|_S$ est injective en tant qu'application de S vers $\text{im } u$. De plus, $\forall y \in \text{im } u$, $\exists x \in E$ tel que $u(x) = y$, or $\exists x_S, x_K \in S \times \ker u$ tel que $x = x_K + x_S$ (puisque $\ker u \oplus S = E$) et donc $y = u(x) = u(x_K + x_S) = u(x_K) + u(x_S) = 0_F + u(x_S) = u(x_S)$ et donc il existe $x_S \in S$ tel que $y = u(x_S)$ ce qui prouve la surjectivité de $u|_S$. ■

Théorème I.2.12 (du rang). E un \mathbb{K} -ev de dimension fini. Soit u un endomorphisme de E , u est de rang fini et $\dim_{\mathbb{K}} E = \text{rang } u + \dim_{\mathbb{K}} \ker u$

Théorème I.2.13. Soit E et F deux \mathbb{K} -ev de dimension finie, alors $\mathcal{L}(E, F)$ est un \mathbb{K} -ev de dimension finie et

$$\dim \mathcal{L}(E, F) = \dim E \times \dim F$$

Si l'on note $n = \dim E$ et $p = \dim F$, $\mathcal{B}_E = (e_i)_{i=1 \dots n}$ une base de E et $\mathcal{B}_F = (f_i)_{i=1 \dots p}$ une base de F alors $\forall (i, j) \in [\![1, n]\!] \times [\![1, p]\!]$,

$$u_{i,j}(e_k) = \begin{cases} 0_F & k \neq j \\ f_i & k = j \end{cases}$$

forme une base de $\mathcal{L}(E, F)$.

I.2.3 Projecteurs

Définition I.2.14 (Projecteur). Soit $E = E_1 \oplus E_2$ alors $\forall x \in E$, $\exists x_1, x_2 \in E_1 \times E_2$ tel que $x = x_1 + x_2$, alors l'application $p_1 : x \mapsto x_1$ est le projecteur sur E_1 parallèlement à E_2 . De même on définit $p_2 : x \mapsto x_2$ le projecteur sur E_2 parallèlement à E_1 .

Remarque I.2.15. Soit $E = E_1 \oplus E_2$, la composition d'un projecteur avec lui-même redonne le projecteur en question : $\forall x \in E$, $y = p_1(x) = p_1(x) \in E_1$ soit $y = y + 0_E$ et donc $p_1(y) = y$ soit encore $p_1^2(x) = x$, et encore $p_1^2 = p_1$.

Théorème I.2.16. Si $E = \bigoplus_{i=1}^p E_i$, alors l'application

$$\begin{aligned} & : \prod_{i=1}^p E_i & \rightarrow & E \\ & (x_1, \dots, x_p) & \mapsto & \sum_i x_i \end{aligned}$$

Est un isomorphisme d'espace vectoriel. Si p_i est la famille de projecteur attaché à la décomposition $E = \bigoplus_{i=1}^p E_i$ alors

$$u^{-1} : x \mapsto (p_i(x))_{i=1 \dots p}$$

Théorème I.2.17. Soit u un endomorphisme de E , u est un projecteur si et seulement si $u^2 = u$. Dans ce cas, $\text{im } u \oplus \ker u = E$ et u est le projecteur sur $\text{im } u$ parallèlement à $\ker u$.

I.2.4 Dualité

Définition I.2.18 (Forme linéaire). Soit E un \mathbb{K} -ev alors $E^* = \mathcal{L}(E, \mathbb{K})$ est appelé espace dual de E , ses éléments sont appelé formes linéaires.

Théorème I.2.19. En dimension fini, soit $\mathcal{B} = (e_i)_{i=1 \dots n}$ (où $n = \dim E$), $\forall u \in E^*$, $\exists (\alpha_i)_{i=1 \dots n} \in \mathbb{K}^n$ tel que,

$$\forall x = \sum_{i=1}^n x_i e_i \quad u(x) = \sum_{i=1}^n \alpha_i x_i$$

Théorème I.2.20 (Hyperplan). Soit E un \mathbb{K} -ev de dimension n alors $\forall u \in E^*$ non identiquement nul $\ker u$ est de dimension $n - 1$. Réciproquement si H est un SEV de dimension $n - 1$, alors il existe $\varphi \in E^*$ tel que $\ker \varphi = H$.

Si φ et ψ deux forme linéaire alors ,

$$\exists \lambda \in \mathbb{K} \text{ tel que } \psi = \lambda \varphi \iff \ker \psi = \ker \varphi$$

Définition I.2.21 (Rang famille de forme linéaire). Soit $\mathcal{F} = (\psi_i)_{i=1 \dots p} \in (E^*)^p$ une famille de forme, on appelle rang de \mathcal{F} le rang de l'application linéaire,

$$\begin{aligned} \Psi : E &\rightarrow \mathbb{K}^p \\ x &\mapsto (\psi_1(x), \dots, \psi_p(x)) \end{aligned}$$

Théorème I.2.22. La famille $\mathcal{F} = (\psi_i)_{i=1 \dots p} \in (E^*)^p$ est libre si et seulement si l'application,

$$\begin{aligned} \Psi : E &\rightarrow \mathbb{K}^p \\ x &\mapsto (\psi_1(x), \dots, \psi_p(x)) \end{aligned}$$

est surjective.

Théorème I.2.23 (Base dual). Soit E un \mathbb{K} -ev de dimension n alors E^* est de dimension n et $\forall \mathcal{B} = (e_i)_{i=1 \dots n}$, on appelle base dual, alors les applications $e_i^* : x \mapsto x_i$ coordonnée sont des formes linéaires constituant une base de E^* . Cette base dual est unique, elle vérifie $e_i^*(e_j) = \delta_{i,j}$.

I.2.5 Matrice

En dimension finie :

- Le théorème I.1.34 implique l'existence d'une base de E , et donc chaque élément de E est entièrement déterminer par $\dim E$ scalaire.
- Le théorème I.2.13 implique que chaque élément de $\mathcal{L}(E, F)$ est entièrement déterminé par $\dim E \dim F$ scalaires.

Ces scalaires sont enregistrer dans un objet en forme de tableau qu'on appelle matrice.

Définition I.2.24 (Matrice). On appelle matrice à p lignes et n colonne à coefficient dans \mathbb{K} le tableau rectangulaire de scalaire de \mathbb{K} à p lignes et n colonne. On note

$$M = (a_{ij})_{i=1 \dots p, j=1 \dots n}$$

On appelle $\mathcal{M}_{p,n}(\mathbb{K})$ l'ensemble des matrices à p lignes et n colonne à coefficient dans \mathbb{K} . Lorsque $p = n$ on parle de matrice carré et on note $\mathcal{M}_{n,n}(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$.

Théorème I.2.25. On a $\dim \mathcal{M}_{p,n}(\mathbb{K}) = np$. On note $E_{i,j} = (\delta_{i,k}\delta_{j,l})_{k=1\dots p, l=1\dots p}$ une famille de np matrice formant ce qu'on appelle la base canonique de $\mathcal{M}_{p,n}(\mathbb{K})$. Soit $\mathcal{B}_E = (e_i)_{i=1\dots n}$ une base de E , $\forall x \in E$ il existe $(\alpha_i)_{i=1\dots n}$ tel que $x = \sum_{i=1}^n \alpha_i e_i$, on note

$$X = \mathbf{Mat}_{\mathcal{B}_E}(x) = (\alpha_i)_{i=1\dots n} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix}$$

Soit $(x_j)_{j=1\dots k}$ une famille de k vecteur de E de composante $x_j = \sum_{i=1}^n \beta_{ij} e_i$, on range ces composantes dans les colonnes d'une matrice à n ligne et k colonne, on note

$$F = \mathbf{Mat}_{\mathcal{B}_E}(x_1, \dots, x_k) = (\beta_{ij})_{i=1\dots n, j=1\dots k} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1i} & \cdots & \beta_{1k} \\ \vdots & \ddots & & \ddots & \vdots \\ \beta_{i1} & & \beta_{ij} & & \gamma_{ik} \\ \vdots & \ddots & & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{ni} & \cdots & \alpha_{nk} \end{pmatrix}$$

Soit $u \in \mathcal{L}(E, F)$ et $\mathcal{B}_F = (f_i)_{i=1\dots n}$ une base de F alors I.2.13 implique qu'il $\exists (\gamma_{i,j})_{i=1\dots p, j=1\dots n}$ tel que $u = \sum_{i,j} \gamma_{i,j} u_{i,j}$, on note,

$$M = \mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u) = (\gamma_{ij})_{i=1\dots p, j=1\dots n} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1j} & \cdots & \gamma_{1n} \\ \vdots & \ddots & & \ddots & \vdots \\ \gamma_{i1} & & \gamma_{ij} & & \gamma_{in} \\ \vdots & \ddots & & \ddots & \vdots \\ \gamma_{p1} & \cdots & \gamma_{pj} & \cdots & \gamma_{pn} \end{pmatrix}$$

la matrice de u dans la base de départ \mathcal{B}_E et la base d'arrivée \mathcal{B}_F . Le choix des base \mathcal{B}_E et \mathcal{B}_F détermine un isomorphisme,

$$\varphi_{\mathcal{B}_E, \mathcal{B}_F} : u \mapsto \mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u)$$

Remarque I.2.26. Les colonnes du $\mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u)$ contiennent les coordonné dans la base \mathcal{B}_F de l'image des vecteurs de la base \mathcal{B}_E .

Que se passe t'il quand on compose des applications linéaires dans des espace de dimension finie. Prenons E, F et G trois \mathbb{K} -ev de dimension finie, 3 bases $\mathcal{B}_E = (e_i)_{i=1\dots n}, \mathcal{B}_F = (f_i)_{i=1\dots p}$ et $\mathcal{B}_G = (g_i)_{i=1\dots m}$, et 2 application linéaire $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$ tel que $\mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u) = (u_{i,j})_{i=1\dots p, j=1\dots n}$ et $\mathbf{Mat}_{\mathcal{B}_F, \mathcal{B}_G}(v) = (v_{i,j})_{i=1\dots m, j=1\dots p}$, alors,

$$\forall i = 1\dots n \quad v \circ u(e_i) = v\left(\sum_{j=1}^p u_{i,j} f_j\right) = \sum_{j=1}^p u_{i,j} v(f_j) = \sum_{j=1}^p u_{i,j} \sum_{k=1}^m v_{j,k} g_k = \sum_{k=1}^m \left(\sum_{j=1}^p u_{i,j} v_{j,k} \right) g_k$$

Ainsi $\mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_G}(v \circ u) = \left(\sum_{j=1}^p u_{i,j} v_{j,k} \right)_{i=1\dots m, j=1\dots p}$. Ce qui permet d'introduire le produit matriciel.

Définition I.2.27 (Produit Matriciel). Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B \in \mathcal{M}_{m,p}(\mathbb{K})$, on appelle produit AB la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ dont le coefficient de la i -ligne et j -colonne est le produit de la i -ligne de B avec la j -colonne de A .

Les calculs précédent assurent avec les mêmes notation que

Théorème I.2.28.

$$\mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_G}(v \circ u) = \mathbf{Mat}_{\mathcal{B}_F, \mathcal{B}_G}(v) \mathbf{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u)$$

Que ce passe t'il en terme matricielle quand on change de base ? Prenons dans E une base $\mathcal{B} = (e_i)_{i=1\dots n}$ et une autre base $\mathcal{B}' = (u_i)_{i=1\dots n}$. On note $P = \mathbf{Mat}_{\mathcal{B}}(\mathcal{B}')$ la matrice de $\mathcal{M}_n(\mathbb{K})$

I. Espaces vectoriels.

contenant les composantes de la base \mathcal{B}' dans la base \mathcal{B} . Soit $x = \sum_{i=k}^n x'_k u_k = \sum_{i=1}^n x_i e_i$, notons $X' = \text{Mat}_{\mathcal{B}'}(x)$ et $X = \text{Mat}_{\mathcal{B}}(x)$. On a $u_k = \sum_{i=1}^n P_{ik} e_i$ et donc $x = \sum_{k=1}^n x'_k u_k = \sum_{k=1}^n x'_k (\sum_{i=1}^n P_{ik} e_i) = \sum_{i=1}^n (\sum_{k=1}^n P_{ik} x'_k) e_i = \sum_{i=1}^n (P X')_i e_i$ soit encore $X = P X'$.

Théorème I.2.29. Soit E un \mathbb{K} -ev de dimension n et $\mathcal{B}_E, \mathcal{B}'_E$ deux base de E . Alors la matrice $\text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E)$ est inversible et $\text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E)^{-1} = \text{Mat}_{\mathcal{B}'_E}(\mathcal{B}_E)$. On a $\forall x \in E$ la relation de passage,

$$\text{Mat}_{\mathcal{B}_E}(x) = \text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E) \text{Mat}_{\mathcal{B}'_E}(x)$$

De même pour toute famille $(x_k)_{k=1 \dots m} \in E$ de vecteur

$$\text{Mat}_{\mathcal{B}_E}(x_1, \dots, x_m) = \text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E) \text{Mat}_{\mathcal{B}'_E}(x_1, \dots, x_m)$$

Soit F un \mathbb{K} -ev de dimension p et $\mathcal{B}_F, \mathcal{B}'_F$, deux base de F . Alors $\forall u \in \mathcal{L}(E, F)$ on a,

$$\text{Mat}_{\mathcal{B}'_E \mathcal{B}'_F}(u) = \text{Mat}_{\mathcal{B}_F}(\mathcal{B}'_F) \text{Mat}_{\mathcal{B}_E \mathcal{B}_F}(u) \text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E)$$

Ce qui pour un endomorphisme u de $\mathcal{L}(E)$ donne :

$$\text{Mat}_{\mathcal{B}'_E \mathcal{B}'_E}(u) = \text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E) \text{Mat}_{\mathcal{B}_E \mathcal{B}_E}(u) \text{Mat}_{\mathcal{B}_E}(\mathcal{B}'_E)$$

Définition I.2.30 (Transposé). Pour tout matrice $A = (a_{ij})_{i=1 \dots p, j=1 \dots n} \in \mathcal{M}_{np}(\mathbb{K})$, on appelle transposé de A , la matrice ${}^t A = (a_{ji})_{j=1 \dots n, i=1 \dots n} \in \mathcal{M}_{pn}(\mathbb{K})$.

Définition I.2.31 (Trace). Si $A \in \text{Mat}_n(\mathbb{K})$, on appelle trace A la somme de ses élément diagonaux,

$$\text{trace } A = \sum_{i=1}^n A_{ii}$$

c'est une forme linéaire sur $\text{Mat}_n(\mathbb{K})$.

Théorème I.2.32. Soit $A, B \in \text{Mat}_n(\mathbb{K})$ alors $\text{trace } AB = \text{trace } BA$, on en déduit qu'on peut définir la trace d'un endomorphisme.

Soit E un \mathbb{K} -ev de dimension fini n et soit \mathcal{B} une base de E , on définit la trace d'un endomorphisme par $\mathcal{L}(E)$ par $\text{trace } u = \text{trace } \text{Mat}_{\mathcal{B}}(u)$ (On vérifie aisément que cette formule est indépendante du choix de la base \mathcal{B}). On dis que la trace est un invariant de similitude.

I.2.6 Determinant

Le déterminant est un outils extrêmement utiles pour caractérisé les matrice carré. Il permet de déterminer celles qui sont inversibles. En effet nous montrerons que les matrice non-inversibles sont caractérisé par $\det A = 0$. Le déterminant est un polynome à n^2 variable constitué de monôme de degré n . Cela implique par exemple que si on identifie $\mathcal{M}_n(\mathbb{K})$ à \mathbb{K}^{n^2} , alors l'espace des matrices non inversibles est une variété plongé de dimension $n^2 - 1$ (ce vocabulaire sera précisé dans la suite du cours). Ainsi "la plupart" des matrices sont inversible, on montrera en particulier que $\mathcal{GL}_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$.

Définition I.2.33 (Application multilinéaires). Soit $(E_i)_{i=1 \dots p}$ et F sont p \mathbb{K} -ev, alors une application $\Phi : \prod_{i=1}^p E_i \longrightarrow F$ est dites p -linéaire si et seulement si elle est linéaire par rapport à chacun de ses arguments. On dis que $\Phi \in \mathcal{L}_p(\prod_{i=1}^p E_i, F)$.

Pour un application $\Phi : \prod_{i=1}^p E \longrightarrow F$ est dites

- symétrique si et seulement si la valeur de $\Phi(x_1, \dots, x_p)$ est invariant par transposition de deux vecteur dans le p -uplet
- antisymétrique si et seulement si la valeur de $\Phi(x_1, \dots, x_p)$ est changé en son opposé par transposition de deux vecteurs dans le p -uplet.

Théorème (et définition) I.2.34 (Permutation et Signature). L'ensemble des bijections de $\{1, \dots, p\}$, muni de la composition \circ est appelé groupe des permutations (S_p, \circ) , il est de cardinal $!p$. Une transposition est une permutation qui échange uniquement deux éléments i et j distincts, $\forall \sigma \in S_p$, il existe une famille de transpositions $(\tau_i)_{i=1 \dots k}$ décomposant $\sigma = \tau_k \circ \dots \circ \tau_1$, la parité du cardinal de la décomposition est indépendante de la décomposition choisie, on note

$$\begin{array}{rcl} \epsilon & S_p & \rightarrow & \{-1, 1\} \\ & \sigma & \mapsto & \begin{cases} -1 & \text{si le cardinal de la décomposition est impair} \\ 1 & \text{si le cardinal de la décomposition est paire} \end{cases} \end{array}$$

Par définition ϵ est un morphisme de groupe $\epsilon(\sigma \circ \tau \sigma) = \epsilon(\sigma)\epsilon(\tau)$. On montre également que

$$\begin{array}{rcl} \epsilon & S_p & \rightarrow & \{-1, 1\} \\ & \sigma & \mapsto & \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i} \end{array}$$

Proposition I.2.35. Soit $\Phi \in \mathcal{L}_p(\prod_{i=1}^p E, F)$, alors $\forall \sigma \in S_p$ et $\forall (x_i)_{i=1 \dots p} \in E^p$, alors

- Φ symétrique $\implies \Phi(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = \Phi(x_1, \dots, x_p)$
- Φ antisymétrique $\implies \Phi(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = \epsilon(\sigma)\Phi(x_1, \dots, x_p)$

Théorème (et définition) I.2.36 (Application alterné). Une application $\Phi \in \mathcal{L}_p(\prod_{i=1}^p E, F)$ est dite *alterné* si et seulement si $\forall (x_i)_{i=1 \dots p} \in E^p$ tel que $\exists i \neq j$ avec $x_i = x_j$ implique $\Phi(x_1, \dots, x_p) = 0_F$.

On montre que toute application alterné est antisymétrique, la réciproque étant vrai si la corps de base a un cardinal différent de 2.

On a également $\forall (x_i)_{i=1 \dots p} \in E^p$,

$$(x_i)_{i=1 \dots p} \text{ lié } \implies \Phi(x_1, \dots, x_p) = 0_F$$

Φ est donc invariant par ajout à un vecteur de la famille, une combinaison linéaire des autres vecteurs.

On note $M = (a_{ij}) = \text{Mat}_{\mathcal{B}}(x_1, \dots, x_n) \in \mathcal{M}(\mathbb{K})$.

Définition I.2.37 (Determinant). Soit E un \mathbb{K} -ev de dimension n . Avec les notations précédentes, on appelle déterminant dans la base \mathcal{B} l'application,

$$\begin{array}{rcl} \det_{\mathcal{B}} & E^n & \rightarrow & \mathbb{K} \\ & (x_1, \dots, x_n) & \mapsto & \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^p a_{i\sigma(i)} \end{array}$$

Pour toute matrice carré de taille n le déterminant est égal au déterminant des colonnes de la matrice dans la base canonique de \mathbb{K}^n .

Théorème I.2.38. Pour toute matrice carré M de taille n on a,

$$\det M = \det {}^t M$$

Remarque I.2.39.

- Par construction $\det_{\mathcal{B}}$ est une application n -linéaire alterné
- On a $\det_{\mathcal{B}}(\mathcal{B})$
- Une famille \mathcal{F} est une base, si dans une base \mathcal{B} on a $\det_{\mathcal{B}}(\mathcal{F}) \neq 0$
- Une matrice carré est inversible si et seulement si son déterminant n'est pas nul
- On ne change pas la valeurs du déterminant d'une matrice en rajoutant à une colonne (resp. une ligne) une combinaison linéaire des autres colonnes (resp. lignes).
- On change un déterminant en son opposé en permettant deux lignes ou deux colonnes

I. Espaces vectoriels.

Preuve. Soit $\sigma \in S_n$ alors puisque σ est bijective $\forall i, \exists! k$ tel que $\sigma(k) = i$ et donc $a_{i\sigma(i)} = a_{\sigma^{-1}(k)k}$ et donc puisque la multiplication est commutative dans \mathbb{K} on a $\prod_{i=1}^n a_{i\sigma(i)} = \prod_{i=1}^n a_{\sigma^{-1}(i)i}$. De plus $\epsilon(\sigma) = \epsilon\sigma^{-1}$ ce qui conduit à,

$$\det M = \sum_{\sigma \in S_n} \prod_{i=1}^n a_{\sigma^{-1}(i)i}$$

La démonstration se conclut en remarquant que $\sigma \rightarrow \sigma^{-1}$ est bijective dans S_n ■

On note E_n un \mathbb{K} -ev de dimension n et $\mathcal{A}_n(E_n)$ l'espace des applications n -linéaire alterné sur E_n .

Théorème I.2.40 (Fondamental des déterminants). **L'espace $\mathcal{A}_n(E_n)$ est un espace vectoriel de dimension 1, plus précisément $\forall \mathcal{B}$ base de E_n et $\forall \Phi \in \mathcal{A}(E_n)$, on a,**

$$\Phi = \Phi(\mathcal{B}) \det_{\mathcal{B}}$$

Preuve. Déjà $\mathcal{A}_n(E_n)$, n'est pas nul puisque $\det_{\mathcal{B}} \neq 0_{\mathcal{A}_n(E_n)}$. Choisissent une base $\mathcal{B} = (e_i)_{i=1 \dots n}$ et une famille de vecteur $(x_i)_{i=1 \dots n}$ qu'on décompose $x_i = \sum_{j=1}^n a_{ji} e_j$. On commence le calcul

$$\begin{aligned} \Phi(x_1, \dots, x_n) &= \Phi\left(\sum_{i_1=1}^n a_{i_1 1} e_{i_1}, \dots, \sum_{i_n=1}^n a_{i_n 1} e_{i_n}\right) \\ &= \sum_{i_1=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} \dots a_{i_n 1} \Phi(e_{i_1}, \dots, e_{i_n}) \end{aligned}$$

Or Φ est alterné il ne reste donc dans la somme de la dernière équation uniquement les termes pour lesquels $\{i_1, \dots, i_n\}$ est une permutation, il existe donc une unique permutation $\sigma \in S_n$ tel que pour tout i on a $k_i = \sigma(i)$. On peut donc écrire,

$$\begin{aligned} \Phi(x_1, \dots, x_n) &= \sum_{\sigma \in S_n} \prod_{i=1}^n a_{\sigma(i), i} \Phi(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^n a_{\sigma(i), i} \Phi(e_1, \dots, e_n) \\ &= \Phi(\mathcal{B}) \det_{\mathcal{B}}(x_1, \dots, x_n) \end{aligned}$$

Où le passage à la seconde ligne utilise l'antisymétrie des formes alternées. ■

Théorème (et définition) I.2.41 (Cofacteur). Soit $A \in \mathcal{M}_n(\mathbb{K})$ alors $\forall i, j$ on appelle cofacteur C_{ij} le déterminant de la matrice X_{ij} obtenue en remplaçant j -em colonne par la i -em vecteur de la base canonique de \mathbb{K}^n . Par $n - i$ transposition on peut faire descendre la ligne i en dernière position on effectue ensuite $n - j$ permutation de colonne pour amener la j -em colonne en dernière position. On obtient donc

$$C_{ij} = (-1)^{i+j} \begin{vmatrix} & & 0 \\ A_{i,j} & \vdots & 0 \\ & \cdots & 1 \end{vmatrix} = (-1)^{i+j} \det A_{ij}$$

Où A_{ij} est la matrice de taille n_1 obtenue en rayant la i -em ligne et la j -em colonne.
Si l'on note C_k , les colonnes de A , on a,

$$\det A = \det(C_1, \dots, C_j, \dots, C_n) = \sum_{i=1}^n a_{ij} \det(C_1, \dots, E_i, \dots, C_n) = \sum_{i=1}^n a_{ij} (\text{com}A)_{ij}$$

On appelle comatrice $\text{com}A$ la matrice des cofacteurs, alors on montre que

$$A^t \text{com}A = {}^t \text{com}A A = \det A I_n$$

Preuve. Posons $B = A^t \text{com}A$, alors

$$b_{ii} = \sum_{j=1}^n a_{ij}(^t \text{com}A)_{ji} = \sum_{j=1}^n a_{ij}(\text{com}A)_{ij} = \det A$$

et

$$b_{ij} = \sum_{k=1}^n a_{ik}(^t \text{com}A)_{kj} = 0$$

Posons $M = (C_1, \dots, \underset{\text{i-ème position}}{C_i}, \dots, \underset{\text{j-ème position}}{C_i}, \dots, C_n)$, $\det M = 0$ par construction et développons à partir de la j -ème colonne. Par construction $\forall k$ les cofacteurs de la j -ème colonne de M sont égaux à ceux de A .

$$0 = \sum_{k=1}^n a_{ki}(\text{com}M)_{kj} = \sum_{k=1}^n a_{ki}(\text{com}A)_{kj} = \sum_{k=1}^n (^t \text{com}A)_{jk} a_{ki} = (^t \text{com}A A)_{ji} = (A^t \text{com}A)_{ij}$$

Théorème (et définition) I.2.42. Pour toute endomorphisme $u \in \mathcal{L}(E_n)$, le déterminant de u est le scalaire tel que $\forall (x_k)_{k=1 \dots n} \in (E_n)^n$ famille de vecteur, pour toute base \mathcal{B}

$$\det_B(u(x_1), \dots, u(x_n)) = \det u \det(x_1, \dots, x_n)$$

Preuve. Il s'agit surtout de montrer que la définition est indépendante de la base considéré. En effet $\det_B \circ u \in \mathcal{A}_n(\mathbb{K})$ et donc puisque cette ensemble est de dimension 1, il est proportionnel au déterminant dans la base \mathcal{B} , il y a donc existence de $\det_{\mathcal{B}} u$ tel que $\det_B \circ u = \det_{\mathcal{B}} u \det_{\mathcal{B}}$. Soit \mathcal{B} et \mathcal{B}' deux base de E_n . Alors en utilisant les résultats du théorème I.2.40

$$\begin{aligned} \det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) &= \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(u(x_1), \dots, u(x_n)) \\ \det_{\mathcal{B}} u \det_{\mathcal{B}}(x_1, \dots, x_n) &= \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'} u \det_{\mathcal{B}'}(x_1, \dots, x_n) \\ \det_{\mathcal{B}} u \det_{\mathcal{B}}(x_1, \dots, x_n) &= \det_{\mathcal{B}'} u \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(x_1, \dots, x_n) \\ \det_{\mathcal{B}} u \det_{\mathcal{B}}(x_1, \dots, x_n) &= \det_{\mathcal{B}'} u \det_{\mathcal{B}}(x_1, \dots, x_n) \end{aligned}$$

Ce qui conduit au résultat puisque ceci est valable pour tout famille de vecteur $(x_k)_{k=1 \dots n}$ ■

Remarque I.2.43.

- En posant $y_k = v(x_k)$ la formule précédente implique $\forall u, v \in \mathcal{L}(E_n)$,

$$\begin{aligned} \det_B(u \circ v(x_1), \dots, u \circ v(x_n)) &= (\det u \circ v) \det(x_1, \dots, x_n) \\ &= \det_B(u \circ v(x_1), \dots, u \circ v(x_n)) \\ &= \det_B(u(y_1), \dots, u(y_n)) \\ &= \det u \det_{\mathcal{B}}(y_1, \dots, y_n) \\ &= \det u \det_{\mathcal{B}}(v(x_1), \dots, v(x_n)) \\ &= \det u \det v \det_{\mathcal{B}}(x_1, \dots, x_n) \\ \det u \circ v &= \det u \det v \end{aligned}$$

- De même pour toute matrice $A, B \in \mathcal{M}_n(\mathbb{K})$, on a $\det AB = \det BA = \det A \det B$
- On interprète le déterminant d'un endomorphisme comme la facteur de proportionnalité entre le volume de l'image d'un famille de vecteur et la volume de la famille

$$\det u = \frac{\text{Vol}(u(x_1), \dots, u(x_n))}{\text{Vol}(x_1, \dots, x_n)}$$

Exercice 5. —

Calculer le déterminant de la matrice

$$\begin{vmatrix} \alpha & 0 & \cdots & \cdots & 0 & \beta \\ 0 & \ddots & 0 & 0 & \ddots & 0 \\ \vdots & 0 & \alpha & \beta & \vdots & \\ \vdots & 0 & \beta & \alpha & \vdots & \\ 0 & \ddots & 0 & 0 & \ddots & 0 \\ \beta & 0 & \cdots & \cdots & 0 & \alpha \end{vmatrix} = 0$$

Exercice 6. —

Montrer que $\det(\text{com}A) = (\det A)^{n-1}$

Exercice 7. —

Montrer que $\forall x$ on a

$$M = \begin{pmatrix} x^2 & (x+1)^2 & (x+2)^2 & (x+3)^2 \\ (x+1)^2 & (x+2)^2 & (x+3)^2 & (x+4)^2 \\ (x+2)^2 & (x+3)^2 & (x+4)^2 & (x+5)^2 \\ (x+3)^2 & (x+4)^2 & (x+5)^2 & (x+6)^2 \end{pmatrix}$$

Proposition I.2.44. Soit $(x_k)_{k=1\dots n} \in \mathbb{K}^n$, on appelle déterminant de Van-dermonde le déterminant :

$$V(x_1, \dots, x_n) = M = \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ x_1 & & x_j & & x_n \\ x_1^2 & & x_j^2 & & x_n^2 \\ \vdots & & \vdots & & \vdots \\ x_1^{n-1} & & x_j^{n-1} & & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq i \leq j \leq n} (x_j - x_i)$$

Preuve. Tout d'abords. on remarque que la formule est évidente si $\exists i \neq j$ tel que $x_i = x_j$. Pour tout polynôme $P = X^{n-1} + \sum_{k=0}^{n-2} \alpha_k X^k$ de degré $n-1$ et de coefficient dominant 1, il existe une combinaison linéaire des lignes $L_n \leftrightarrow L_n + \sum_{i=1}^{n-1} \alpha_{k-1} L_k$, tel que la dernière ligne deviens égale à $(P(\alpha_1), \dots, P(\alpha_n))$, on choisit alors $P = \prod_{i=1}^{n-1} (X - \alpha_i)$. Et donc on fais apparaitre $n-1$ 0 sur la dernière ligne. On a,

$$V(x_1, \dots, x_n) = P(x_n)V(x_1, \dots, x_{n-1})$$

Ce qui permet de. conclure par une récurrence immédiate. ■

Exercice 8. —

Soit $A : t \mapsto A(t) = (C_1(t), \dots, C_n(t))$ une matrice dont tout les coefficients sont des fonctions dérivable sur \mathbb{R} ,

a. Montrer que $\varphi : t \mapsto \det A(t)$ a pour dérivé $t \mapsto \sum_{i=1}^n \det(C_1(t), \dots, C_{i-1}(t), C'_i(t), C_{i+1}(t), \dots, C_{n(t)}(t))$

b. Calculer

$$D_n(x) = \begin{pmatrix} x & 1 & 0 & \cdots & 0 \\ \frac{x^2}{2!} & x & 1 & \ddots & \vdots \\ \vdots & \frac{x^2}{2!} & x & 1 & 0 \\ \frac{x^{n-1}}{(n-1)!} & \vdots & \frac{x^2}{2!} & x & 1 \\ \frac{x^n}{n!} & \frac{x^{n-1}}{(n-1)!} & \vdots & \frac{x^2}{2!} & x \end{pmatrix}$$

I.3 Réduction des endomorphismes

On peut, avec les éléments de $(\mathbb{K}[X], +, \times)$ construire un certain nombre de théorème et de propriété semblable à certaines propriétés sur $(\mathbb{Z}, +, \times)$. Cela est dut à l'existence d'une notion de division Euclidienne dans chacun de ses anneaux intègre. On rappelle qu'un anneau intègre est un ensemble $(\mathcal{A}, +, \times)$ pour lequel $(\mathcal{A}, +)$ est un groupe abélien, et la multiplication est associative, distributive par rapport a l'addition, possède un élément neutre et $\forall a, b \in \mathcal{A}_*, ab \neq 0$. Les résultats que nous établirons sur les idéaux de $(\mathbb{K}[X], +, \times)$, va nous permettre de simplifier l'expression des endomorphismes d'espace vectoriel. Les résultats établit en sous-section I.3.1 pourraient être généraliser au anneaux en toute généralité, néanmoins par souci de simplicité et de finalité nous nous concentrerons sur $\mathbb{K}[X]$.

I.3.1 Arithmétique des polynomes

Définition I.3.1 (Divisibilité). Soit $P, Q \in \mathbb{K}[X]$, on dis que $Q|P$ si et seulement si $\exists R \in \mathbb{K}[X]$ tel que $P = QR$.

Théorème I.3.2. Soit $P, Q \in \mathbb{K}[X]$ tel que $Q|P$ et $P|Q$ alors $\exists \alpha \in \mathbb{K}_*$ tel que $P = \alpha Q$.

Théorème I.3.3 (Division Euclidienne). Soit $A, B \in \mathbb{K}[X] \times \mathbb{K}[X]_*$ alors $\exists !(Q, R) \in \mathbb{K}[X] \times \mathbb{K}_{\deg B-1}[X]$ tel que $A = BQ + R$.

Preuve. On ne perd pas de généralité à choisir A, B de coefficient de plus haut degré égaux à 1, on posera donc $A = X^n + \sum_{k=0}^{n-1} a_k X^k$ et $B = X^p + \sum_{k=0}^{p-1} b_k X^k$. Trois cas se présentent à nous,

- cas 1 si $p > n$ alors $A = 0_{\mathbb{K}[X]} B + A$ convient
- cas 2 si $p = n$ alors $Q = 1$ et $R = \sum_{k=0}^{n-1} (a_k - b_k) X^k$
- cas 3 si $p < n$ alors on note $m = n - p > 0$ et on pose $A_1 = \sum_{k=0}^{n-1} (b_{k-m} - a_k) X^k$ on a $A = X^m B + A_1 = A$, on a $\deg A_1 \leq n-1$, si $\deg A_1 > n$ on réitère le processus avec Q_1 ($A_1 = \alpha_1 Q_1$ le terme de plus haut degré de Q_1 est égal à 1). On construit comme cela une suite de polynômes $(A_i)_{i \in I}$ de degré strictement décroissant le terme de plus haut degré, on arrive comme cela au bout d'un nombre fini d'étape aux cas 1 ou 2. Si on est ramener au cas 1 alors on pose $Q = X^m + \sum_{i \in I} (\prod_{k=1}^i \alpha_k) X^{\deg(Q_i)-p}$ et $R = (\prod_{i \in I} \alpha_i) Q_{\max I}$ fonctionne. Si on est ramener au cas 2 alors on construit un A_i supplémentaire.

Pour l'unicité, considérons $(Q, R) \in \mathbb{K}[X] \times \mathbb{K}_{\deg B-1}[X]$ et $(Q', R') \in \mathbb{K}[X] \times \mathbb{K}_{\deg B-1}[X]$ deux couples satisfaisant la division euclidienne alors on a $B(Q - Q') = R' - R$, supposons que $Q \neq Q'$ alors $\deg((Q - Q')B) \geq \deg B$ ce qui entraîne en contradiction avec $\deg(R - R') < \deg B$. ■

Théorème I.3.4 (D'Alembert). $\forall P \in \mathbb{C}[X]$ de degré supérieur ou égal à 1, alors, $\exists \lambda \in \mathbb{C}$ tel que $P(\lambda) = 0$.

Théorème (et définition) I.3.5 (Polynôme irréductible). Un polynôme $P \in \mathbb{K}[X]$ est irréductible si $\deg P \geq 1$ et si $\forall R \in \mathbb{K}[X]_*$ tel que $R|P$, on a soit

- $\exists a \in \mathbb{K}_*$ tq $R = aP$
- $\deg R = 0$

Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1 quand les polynômes irréductibles de \mathbb{R} sont les polynômes de degré 1 auxquels et les polynômes de degré 2 de déterminant strictement négatif.

Preuve. Commençons par montrer que $\forall P \in \mathbb{K}[X] \forall a \in \mathbb{K}$

$$(X - a)|P \iff P(a) = 0$$

Le sens direct est évident, pour la réciproque c'est une simple conséquence de la division euclidienne, en effet, supposons que $a \in \mathbb{K}$ soit racine de P , effectuons la division euclidienne de P par $X - a$, alors $\exists Q \in \mathbb{K}[X]$ et R de degré 0 tel que $P = Q(X - a) + R$ évaluons cette division en a on obtient $R(a) = 0$ et donc $R = 0$ (R de degré 0).

Le théorème de D'Alembert I.3.4 assure l'existence d'une racine pour tout polynôme de \mathbb{C} , et donc tout polynôme de degré supérieur ou égal à 2 est divisible par un polynôme de degré 1. Ce qui implique que les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.

Pour les polynômes irréductibles de $P \in \mathbb{R}[X]$,

- Soit ce polynômes ont une racine a dans \mathbb{R} et donc est divisible par $X - a$ et donc $\exists a \in \mathbb{R} - \{0\}$ tel que $P = a(X - a)$.
- Soit on se ramène au cas précédent en notant que $\mathbb{R}[X] \subset \mathbb{C}[X]$ et donc qu'il existe $\lambda \in \mathbb{C}$ tel que $P(\lambda) = 0$, puisque $P \in \mathbb{R}[X]$ on a $P(z) = P(\bar{z})$ et donc $(X - z)(X - \bar{z}) \in \mathbb{R}[X]$ divise P et donc $\exists a \in \mathbb{R}$ tel que $P = a(X^2 - 2\operatorname{Re}(z)X + |z|^2)$ dont le discriminant est égal à $-4(\operatorname{im} z)^2 < 0$

■

Définition I.3.6 (Idéal). Un idéal \mathcal{I} de $\mathbb{K}[X]$ est une partie de l'ensemble des polynômes vérifiant

- $(\mathcal{I}, +)$ est un groupe
- est stable par la multiplication : $\forall P \in \mathcal{I}, \forall Q \in \mathbb{K}[X]$ on a $PQ = QP \in \mathcal{I}$

$\forall P \in \mathbb{K}[X]$, on note $P\mathbb{K}[X] = \{PQ \mid Q \in \mathbb{K}[X]\}$ l'idéal engendré par P .

Lemme I.3.7 (Inclusion des idéaux et divisibilité). Si $Q|P \in \mathbb{K}[X]$ alors $P\mathbb{K}[X] \subset Q\mathbb{K}[X]$

Théorème I.3.8 (Idéaux de polynômes). **Tout idéal de $\mathbb{K}[X]$ est engendré par un élément.**

Preuve. Soit $\mathcal{I} \subset \mathbb{K}[X]$ un idéal. Si $\mathcal{I} = \{0\}$ alors il est engendré par le polynôme nul. Sinon $\{\deg P | P \in \mathbb{K}[X]\} \subset \mathbb{N}$ est non vide et possède un plus petit élément $d \geq 0$ et il existe $A \in \mathcal{I}$ tel que $\deg A = d$. On a évidemment $A\mathbb{K}[X] \subset \mathcal{I}$. Soit $P \in I$ on fait la division euclidienne de P par A , on a donc l'existence de Q, R avec $\deg R < d$ tel que $P = QA + R$, or $R = P - QA \in \mathcal{I}$ et $\deg R < d$ donc $R = 0$ et donc $P \in A\mathbb{K}[X]$. Soit encore $\mathcal{I} = A\mathbb{K}[X]$. ■

Théorème (et définition) I.3.9 (PGCD a Bezout). **Soit deux polynômes P et Q , alors,**

$$P\mathbb{K}[X] + Q\mathbb{K}[X] = \{AX + BQ | A, B \in \mathbb{K}[X]\}$$

Est un idéal il est donc généré pas un élément D

$$P\mathbb{K}[X] + Q\mathbb{K}[X] = D\mathbb{K}[X]$$

appelé PGCD de P et Q , on le note $D \hat{=} P \wedge Q$. On a alors

$$\begin{cases} D \text{ divise } P \text{ et } Q \\ \exists A, B \in \mathbb{K}[X] \text{ tel que } D = AP + BQ \end{cases}$$

Deux éléments sont dis premier entre eux si $P\mathbb{K}[X] + Q\mathbb{K}[X] = \mathbb{K}[X]$, on note $P \wedge Q = 1$.
La proposition précédente deviens,

$$1 = P \wedge Q \begin{cases} D \text{ divise } P \text{ et } Q \\ \exists A, B \in \mathbb{K}[X] \text{ tel que } 1 = AP + BQ \end{cases}$$

c'est ce qu'on appelle le *théorème de Bezout*.

Théorème I.3.10 (Gauss). **Soit $P, Q, R \in \mathbb{K}[X]$ tel que $P|QR$ et $P \wedge Q = 1$ alors $P|R$**

Remarque I.3.11. Les résultats du théorème précédent peuvent facilement être étendus à des familles de polynômes.

Théorème (et définition) I.3.12 (Réduction en facteur irréductibles des polynômes). **Tout polynômes normalisé de degré supérieur ou égale à 1 de $\mathbb{K}[X]$ admet un unique décomposition en facteur irréductible normalisé.**

$$\forall P \in \mathbb{K}[X] \exists (P_i)_{i=1 \dots k} \text{ irréductible et } \exists (\alpha_i)_{i=1 \dots k} \in \mathbb{N}_*^k \text{ et } \lambda \in \mathbb{K} \text{ tel que } P = \lambda \prod_{i=1}^k P_i^{\alpha_i}$$

On dis que P est scindé si cette décomposition ne fais intervenir que des polynômes irréductible de degré 1.

Un outils extrêmement puissant sur le polynômes est ce qu'on appelle résultant. Cette outils permet de caractérisé analytiquement et continument les polynômes premier entre eux, mais également les polynômes scindé. Ce caractere analytique et continue du résultant va permettre de démontrer des résultats extrêmement puissant comme le fais que l'ensemble des matrice diagonalisable sur \mathbb{C} est dense sur $\mathcal{M}_n(\mathbb{C})$.

Définition I.3.13 (Application de Sylvester et Résultant). *On considère deux polynômes $A = \sum_{i=0}^n a_i X^i \in \mathbb{K}_n[X]$ et $B = \sum_{i=0}^m b_i X^i \in \mathbb{K}_m[X]$ (où $a_n \neq 0$ et $b_m \neq 0$), alors on définit l'application linéaire de Sylvester :*

$$\begin{aligned} Syl(A, B) : \quad & \mathbb{K}_{m-1}[X] \times \mathbb{K}_{n-1}[X] & \rightarrow & \mathbb{K}_{m+n-1}[X] \\ & (U, V) & \mapsto & AU + BV \end{aligned}$$

La matrice de cette application linéaire dans les bases canoniques de départ et d'arrivée est une matrice carré de taille $(m + n)$ s'écrit,

$$M = \begin{pmatrix} m \text{ colonnes} & & n \text{ colonnes} \\ \begin{matrix} a_n & 0 & \cdots & 0 & b_m & 0 & \cdots & 0 \\ a_{n-1} & a_n & \ddots & \vdots & \vdots & b_m & \ddots & \vdots \\ \vdots & a_{n-1} & \ddots & 0 & \vdots & & \ddots & 0 \\ \vdots & \vdots & \ddots & a_n & b_1 & & & b_m \\ a_0 & & & a_{n-1} & b_0 & \ddots & \vdots & \vdots \\ 0 & \ddots & & & \vdots & 0 & \ddots & b_1 \\ \vdots & \ddots & a_0 & \vdots & \vdots & \vdots & \ddots & b_0 \\ 0 & \cdots & 0 & a_0 & 0 & \cdots & 0 & b_0 \end{matrix} \end{pmatrix}$$

On appelle Résultant de A et B le déterminant de cette matrice. On le note $\text{Res}(A, B) = \det \text{Syl}(A, B)$

Propriété I.3.14.

- $\forall \lambda, \mu \in \mathbb{K}$ on a $\text{Res}(\lambda P, \mu Q) = \lambda^m \mu^n \text{Res}(P, Q)$
- $\text{Res}(P, Q) = (-1)^{mn} \text{Res}(Q, P)$

Théorème I.3.15. Si $A \in \mathbb{K}_n[X]$ et $B \in \mathbb{K}_m[X]$ non nuls, ne sont pas premier entre eux si et seulement si le morphisme $\text{Syl}(A, B)$ n'est pas injectif et donc $\text{Res}(A, B) = 0$

Preuve. Si A et B ne sont pas premier entre eux, en notant $C = A \wedge B \in \mathbb{K}_{\min(m,n)}[X]$ on a $\deg(C) > 1$, on a l'existence de $\tilde{A} \in \mathbb{K}_{n-\deg(C)}[X] \subset \mathbb{K}_{n-1}[X]$ et $\tilde{B} \in \mathbb{K}_{m-\deg(C)}[X] \subset \mathbb{K}_{m-1}[X]$ tel que $A = C\tilde{A}$ et $B = C\tilde{B}$ alors $\text{Syl}(A, B)(\tilde{B}, -\tilde{A}) = A\tilde{B} - B\tilde{A} = C\tilde{A}\tilde{B} - C\tilde{B}\tilde{A} = 0$. Or $\tilde{B}, -\tilde{A}$ n'est pas nul dans $\mathbb{K}_{m-1}[X] \times \mathbb{K}_{n-1}[X]$ et donc $\text{Syl}(A, B)$ n'est pas injectif soit encore $\text{Res}(A, B) = 0$

Réciproquement si $\text{Syl}(A, B)$ n'est pas injectif n'est pas injectif considérons un couple de polynôme $(P, Q) \in \ker \text{Syl}(A, B) \subset \mathbb{K}_{m-1}[X] \times \mathbb{K}_{n-1}[X]$ non nuls vérifiant $AP + QB = 0$, alors $A|BQ$. On suppose que $A \wedge B = 1$ et donc en appliquant le théorème de Gauss I.3.10, on a $A|Q$ ce qui est impossible puisque $\deg(A) = m$ et $\deg(Q) = m - 1$. ■

Nous allons caractériser le résultats de polynômes scindés dans \mathbb{K} en fonction de la liste des racines.

Théorème I.3.16 (Poisson). On suppose que les polynômes $A \in \mathbb{K}_n[X]$ et $B \in \mathbb{K}_m[X]$ sont scindé, alors $\exists (\alpha_i)_{i=1 \dots n} \in \mathbb{K}^n$ et $\exists (\beta_i)_{i=1 \dots m} \in \mathbb{K}^m$ tel que $A = a_n \prod_{i=1}^n (X - \alpha_i)$ et $B = b_m \prod_{i=1}^m (X - \beta_i)$ alors

$$\text{Res}(A, B) = a_n^m b_m^n \prod_{1 \leq i \leq n, 1 \leq j \leq m} (\alpha_i - \beta_j)$$

Preuve. On commence pas considéré les racines distincte entre elle. Les coefficient de A appartiennent à $\mathbb{K}_1[\alpha_1 \dots \alpha_n, a]$ et les coefficient de B appartiennent à $\mathbb{K}_1[\beta_1 \dots \beta_m, b]$ chaque terme du déterminant est un produit d'un élément de chaque colonne et donc $\text{Res}(A, B)$ est un polynôme de $\mathbb{K}[\alpha_1 \dots \alpha_n, \beta_1 \dots \beta_m, a, b]$ ou en tant que polynômes en α_i ($\forall i = 1 \dots n$), $\text{Res}(A, B)$ est de degré m (m colonne font intervenir les coefficient de A), de façon symétrique en tant que polynômes en β_j ($\forall j = 1 \dots m$), $\text{Res}(A, B)$ est de degré n (n colonnes font intervenir les coefficients de B). Considérer comme polynômes en α_1 , $\text{Res}(A, B)$ est de degré m est à pour racine la liste des $(\beta_j)_{j=1 \dots m}$ tel que $\alpha_1 = \beta_j$ (en effet, dans ce cas la A et B en tant que polynôme sur X ne sont pas premier entre eux et alors $\text{Res}(A, B) = 0$), et donc $\prod_{j=1}^m \alpha_1 - \beta_j$ divise $\text{Res}(A, B)$, des considérations de degré implique qu'il existe $R_1 \in \mathbb{K}[\alpha_2 \dots \alpha_n, \beta_1, \dots, \beta_m, a, b]$ tel que $\text{Res}(A, B) = \prod_{j=1}^m (\alpha_1 - \beta_j) R_1$. Des considération de degré implique qu'en tant que polynôme en $\alpha_{i=2 \dots n}$, R_1 est toujours de degré m or en tant que polynômes de $\beta_{j=1 \dots m}$ R_1 est de degré $n - 1$, de plus en tant que polynôme en α_2 , les $\beta_{j=1 \dots m}$ sont à nouveau racines car $\prod_{j=1}^m (\alpha_1 - \beta_j)$ est de degré 0 en α_2 . On procède donc par une récurrence de n étapes en éliminant étapes par étapes l'une des variables α_i et on obtient donc l'existence

de $R_n \in \mathbb{K}[\beta_1, \dots, \beta_m, a, b]$ où chaque variables est de degré 0 et donc $R_n \in \mathbb{K}[a, b]$. On a donc $\text{Res}(A, B) = R_n \prod_{1 \leq i \leq n, 1 \leq j \leq m} (\alpha_i - \beta_j)$. Pour déterminer R_n , il suffit de remarquer que chaque coefficients des m premiers colonnes est proportionnel à a et chaque coefficient des n dernières colonnes est proportionnel à b , il existe donc $c \in \mathbb{K}$ tel que $R_n = a^m b^n c$. Pour déterminer c il suffit de prendre $A = X^n$ alors le déterminant de la matrice de Sylvester se réduit au déterminant de

$$\begin{vmatrix} 1 & 0 & \cdots & 0 & b_m & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & b_m & \ddots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots & & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 & b_1 & & & b_m \\ 0 & & & & b_0 & \ddots & \vdots & \vdots \\ 0 & \ddots & & \vdots & 0 & \ddots & b_1 & \vdots \\ \vdots & \ddots & 0 & \vdots & \vdots & \ddots & b_0 & b_1 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_0 \end{vmatrix} = b_0^n$$

Ce qui permet de déduire $c = 1$. ■

Corrolaire I.3.17.

- Pour tout polynômes $A, B, C \in \mathbb{K}[X]$ on a $\text{Res}(AB, C) = \text{Res}(A, C)\text{Res}(B, C)$.
- Soit R le rest de la division euclidienne de A par B , alors si le coefficient dominant de B est b_m alors $\text{Res}(A, B) = ((-1)^m b_m)^{n-\deg R} R(R, B)$

Preuve. Les preuve sont laisser à titre d'exercices, dans \mathbb{C} il suffit d'utiliser le théorème précédent (dans \mathbb{C} tout les polynômes sont scindés). Sur \mathbb{R} on utilisera le fait que les polynômes réel sont des polynômes sur \mathbb{C} vérifiant quelques symétries. ■

I.3.2 Calculs polynomiaux sur les endomorphismes

Dans ce qui suivra, nous travaillerons sur l'espace $\mathbb{K}[u]$ des polynômes d'un endomorphisme $u \in \mathcal{L}(E)$.

Théorème (et définition) I.3.18 (Polynômes et idéal annulateur). Soit $u \in \mathcal{L}(E)$, un polynôme $P \in \mathbb{K}[X]_*$ est dit **annulateur** si et seulement si $P(u) = 0$ dans $\mathcal{L}(E)$.

L'application $\varphi_u : P \in \mathbb{K}[X] \mapsto P(u) \in \mathbb{K}[u]$ est un morphisme dont le noyau est un idéal de $\mathbb{K}[X]$. On l'appelle **idéal annulateur de u** .

Remarque I.3.19. On sait Thm.I.3.8 que les idéaux annulateur de $\mathbb{K}[X]$ sont engendré par un élément alors soit

- $\ker \varphi_u = \{0\}$ et donc il n'y a pas de polynômes annulateur de u
- $\ker u$ est non réduit au polynômes nul, il donc engendré par un élément non nul, on l'appelle *polynôme minimal annulateur de u* .

Théorème I.3.20. $u \in \mathcal{L}(E)$ alors $\mathbb{K}[u]$ possède un polynôme minimal annulateur π_u si et seulement si $\mathbb{K}[u]$ est de dimension finie égale au degré de π_u et $\{I_E, u, u^2 \dots u^{\deg \pi_u - 1}\}$ est une base de $\mathbb{K}[u]$.

Preuve. Si u n'a pas de polynôme annulateur alors $\mathbb{K}[u]$ est isomorphe à $\mathbb{K}[X]$ et est donc de dimension infinie

Si u possède un polynôme annulateur minimal (notons d son degré) alors il est évident que $\{I_E, u, u^2 \dots u^{\deg \pi_u - 1}\}$ est libre sinon on aurait un polynôme annulateur de degré inférieur. De plus $\forall w \in \mathbb{K}[u]$ il existe $P \in \mathbb{K}[X]$ tel que $w = P(u)$. On effectue la division euclidienne de P par π_u , on obtient $Q, R \in \mathbb{K}[X]$ tel que $P = Q\pi_u + R$ (avec $\deg R < \deg \pi_u$), en évaluant en u , on obtient $w = R(u)$ et donc $\{I_E, u, u^2 \dots u^{\deg \pi_u - 1}\}$ est génératrice. ■

Exercice 9. —

Soit $A \in \mathcal{M}_n(\mathbb{R})$ tel que $A^2 + 2I_n = 3A$ expliciter A^{2020} comme combinaison linéaire de A et I_n .

Correction de l'exercice 5. On sait que $X^2 - 3X + 2$ est annulateur, il se factorise $(X - 1)(X - 2)$, on cherche le rest de la division euclidienne de X^{2020} par $(X - 1)(X - 2)$. Ce reste est de degré inférieur ou égal à 1. On a $R = aX + b$ et $X^{2020} = Q(X^2 - 3X + 2) + R$, pour trouver R il suffit simplement d'évaluer ce polynôme en sur les racine de $X^2 - 3X + 2$, soit en 1 et en 2. On a donc $a + b = 1$ et $2^{2020} = 2a + b$, on a donc $a = 2^{2020} - 1$ et $b = 2 - 2^{2020}$. On a donc $u^{2020} = (2^{2020} - 1)u + (2 - 2^{2020})I_n$

Théorème I.3.21. Soit $u \in \mathcal{L}(E)$ possédant un polynôme minimal, un élément $P(u) \in \mathbb{K}[u]$ est inversible si et seulement si $P \wedge \pi_u = 1$. De plus l'inverse de $P(u)$ est également dans $\mathbb{K}[u]$.

Preuve. Si P est premier avec π_u le théorème de Bezout I.3.9 implique qu'il existe $U, V \in \mathbb{K}[X]$ tel que $UP + V\pi_u = 1$, évalué en u cette identité conduit à $P(u) \circ U(u) = U(u) \circ P(u) = I_n$. Ce qui prouve l'inversibilité de $P(u)$ et le fait que l'inverse soit également un polynôme en u .

Si P n'est pas premier avec π_u alors en posant $D = P \wedge \pi_u$ (degré $D > 1$) on peut écrire $P = DP_1$ et $\pi_u = D\pi_1$, alors $\pi_1 P(u) = \pi_1 DP_1(u) = \pi_u P_1(u) = P_1 \pi_u(u) = 0_{\mathcal{L}(E)}$ ce qui implique la non inversibilité de $P(u)$ ■

I.3.3 Sous-espace stable, valeurs propre sous espace. propre

Nous sommes maintenant outillés pour travailler à réduire les endomorphismes.

Définition I.3.22 (Sous espace stable). Soit F un sous espace vectoriel de E on dit que F est stable par u si et seulement si l'espace vectoriel $u(F) \subset F$.

Théorème I.3.23. Si $\forall u, v \in \mathcal{L}(E)$ commute alors $\ker v$ et $\text{im } v$ sont stable pas u .

La démonstration est laissée à titre d'exercice.

Théorème I.3.24 (Décomposition des noyeaux). Soit P et Q deux polynômes de $\mathbb{K}[X]$ premier entre eux et $u \in \mathcal{L}(E)$, alors on a

$$\ker PQ(u) = \ker P(u) \circ Q(u) = \ker P(u) \oplus \ker Q(u)$$

Preuve. Comme $P(u)$ et $Q(u)$ commutent il est clair que $\ker P(u)$ et $\ker Q(u)$ sont inclus dans $\ker PQ(u)$. Il s'agit de montrer qu'ils sont indépendants et que leur somme directe est égale à $\ker PQ(u)$. Le théorème de Bezout implique l'existence de $U, V \in \mathbb{K}[X]$ tel que $UP + VQ = 1$, on en déduit que $\forall x \in E$, $x = U(u) \circ P(u)(x) + V(u) \circ Q(u)(x)$, si $x \in \ker P(u) \cap \ker Q(u)$, alors l'identité précédente implique $x = U(u) \circ P(u)(x) + V(u) \circ Q(u)(x) = 0_E + 0_E = 0_E$, les espaces $\ker P(u)$ et $\ker Q(u)$ sont donc indépendants.

Soit $x \in \ker PQ(u)$, alors $x = U(u) \circ P(u)(x) + V(u) \circ Q(u)(x)$ ce qui achève la démonstration. ■

Corollaire I.3.25. Si $P = P_1 \dots P_m$ un polynôme annulateur de u et $i \neq j \Rightarrow P_i \wedge P_j = 1$ alors

$$E = \bigoplus_{i=1}^m \ker P_i(u)$$

Définition I.3.26 (Valeurs, Vecteur et sous espace propre). Soit $u \in \mathcal{L}(E)$, soit $\lambda \in \mathbb{K}$ et $x \in E$ tel que $u(x) = \lambda x$ alors on dit que

- λ est une valeur propre on note $u = \{\lambda | \lambda \text{ valeur propre de } u\}$
- x est un vecteur propre
- $E_u(\lambda) = \ker(u - \lambda I_E) \neq \{0_E\}$ est un sous espace propre

Théorème I.3.27. Soit $(\lambda_i)_{i=1 \dots p}$ des valeurs propres deux à deux distinctes de u alors les sous espaces propres sont indépendants.

Théorème I.3.28. Soit $u \in \mathcal{L}(E)$ possédant un polynôme minimal annulateur π_u alors,

$$\text{sp}(u) = \{\alpha \in \mathbb{K} | \pi_u(\alpha) = 0\}$$

Preuve. L'inclusion $\text{sp}(u) \subset \{\alpha \in \mathbb{K} | \pi_u(\alpha) = 0\}$ est évidente. Réciproquement, si α une racine du polynôme minimal annulateur de multiplicité p , alors $\pi_u = (X - \alpha)^p R$ et $(X - \alpha)^p \wedge R = 1$ et donc le théorème de décomposition des noyeaux I.3.24 implique que $E = \ker(u - \alpha I_E)^p \oplus \ker R(u)$. Si α n'était pas valeur propre, alors $(u - \alpha I_E)$ serait injectif et donc $(u - \lambda I_E)^p$ également ce qui impliquerait $E = \ker R(u)$ ce qui entraîne que R est annulateur, or cela est impossible puisque π_u est annulateur minimal. ■

Définition I.3.29 (Polynôme Caractéristique). Soit $u \in \mathcal{L}(E_n)$, alors on appelle polynôme caractéristique le polynôme

$$\chi_A(X) = \det(u - XI_{E_n}) = (-X)^n + (-1)^{n-1}(\text{trace } u)X^{n-1} + \dots + \det u$$

On généralise immédiatement pour les matrices carrées. On a de façon évidente $u = \{x \in \mathbb{K} | \chi_u(x) = 0\}$

Définition I.3.30 (Multiplicité de valeurs propres). Soit $\forall \lambda \in u$ on appelle multiplicité de la valeur propre λ sa multiplicité comme racine du polynôme caractéristique.

Théorème I.3.31 (Caley-Hamilton). Si u est un endomorphisme de E alors le polynôme caractéristique est annulateur pour u . Ce résultat se généralise aux matrices carrées.

Preuve. On considère,

$$\begin{aligned} \varphi_A &: \mathbb{K} \rightarrow \mathcal{M}_n \mathbb{K} \\ s &\mapsto {}^t \text{com}(A - sI_n) \end{aligned}$$

Les règles de calcul des cofacteurs montrent que chacun d'entre eux est un polynôme de degré $n-1$, il existe donc une unique famille de matrices $(A_i)_{i=0,\dots,n-1} \in \mathcal{M}_n \mathbb{K}$ tel que

$$\varphi_A(s) = \sum_{i=0}^{n-1} A_i s^i$$

On a $(A - sI_n) {}^t \text{com}(A - sI_n) = \chi_A(s)I_n$ et donc $AA_0 + s(AA_1 - A_0) + \dots + s^{n-1}(AA_{n-1} - A_{n-2}) - s^n A_{n-1} = \chi_A(s)I_n$, on identifie coefficient par coefficient, on note $\chi_A(s) = \sum_{i=0}^n \alpha_i s^i$, on obtient le système :

$$\left\{ \begin{array}{l} AA_0 = \alpha_0 I_n \\ AA_1 - A_0 = \alpha_1 I_n \\ \vdots \\ AA_i - A_{i-1} = \alpha_i I_n \\ \vdots \\ -A_{n-1} = \alpha_n I_n \end{array} \right.$$

En multipliant la première équation par I_n , la seconde par A , la troisième par A^2 et ainsi de suite, puis en additionnant on obtient $\chi_A(A)I_n = 0$ ce qui donne le résultat escompté. ■

Corollaire I.3.32. Le polynôme caractéristique d'un endomorphisme en dimension finie étant de degré n multiple du polynôme minimal annulateur qui est donc de degré inférieur ou égal à n . En particulier $\dim \mathbb{K}[u] \leq n$.

I.3.4 Diagonalisation et trigonalisation

Les endomorphismes sont, dans leurs sous espaces propres extrêmement facile à calculer, dans ce qui suit nous verrons quelques conditions permettant une simplification de l'endomorphisme étudié basé sur les principes de complémentarité des sous espaces propres.

Endomorphisme diagonalisables

Endomorphisme trigonalisables

CHAPITRE II

Topologie, continuité, produit scalaire

Sommaire

II.1 Norme et topologie 32

II.2 Continuité et compacité 32

Nous souhaitons remercier VOEDTS 2002 dont les documents de cours nous ont beaucoup aidé lors de la rédaction de ce chapitre.

II.1 Norme et topologie

II.1.1 Norme et distance

Généralités

Boules et sphère

Suite d'un espace normé

Normes équivalentes

II.1.2 Topologie

Généralités

Adhérence et caractérisation séquentielle des fermés

Partie dense

Intérieur, Frontière

II.1.3 Espace complet

II.2 Continuité et compacité

II.2.1 Limite et continuité

II.2.2 Compacité

II.2.3 Connexité par arc

II.2.4 Formes quadratiques, produits scalaires, endomorphismes symétriques et orthogonaux

II.2.5 Forme bilinéaire et forme quadratique

II.2.6 Réduction d'une forme quadratique

II.2.7 Produit scalaire

II.2.8 Endomorphisme auto-adjoint

II.2.9 Auto-morphisme orthogonaux

II.2.10 Produit mixte et. produit vectoriel

CHAPITRE III

Bases du calcul différentiel

Sommaire

III.1 Introduction	33
III.2 Fonctions de plusieurs variables à valeur dans les réels	34
III.3 Différentiabilité	39
III.4 Composition et opération sur les différentielles	43
III.5 Théorèmes d'inversion	52
III.6 Annexe : preuves techniques	60

III.1 Introduction

Le calcul différentiel est la base absolue de toutes les techniques de calculs utilisées en physique. La grande idée du calcul différentiel est de ramener l'étude du courbe à l'étude du droit, c'est-à-dire de ramener les variations d'objets complexes à l'étude d'applications linéaires au voisinage d'un point. La grande idée du calcul différentiel peut se ramener à la formule suivante. Si f est différentiable en a , alors il existe un ouvert centré sur 0 tel que pour tout vecteur \mathbf{h} dans cet ouvert, on a :

$$f(a + \mathbf{h}) = f(a) + df(a)\mathbf{h} + o(\|\mathbf{h}\|), \quad (\text{III.1})$$

soit :

$$\text{variations au voisinage d'un point} = \text{algèbre linéaire} + \text{majoration de normes}. \quad (\text{III.2})$$

Autrement dit, il va s'agir de montrer que sous certaines hypothèses de régularité qui seront précisées dans le cours, sur un voisinage suffisamment petit, toute fonction peut être approchée par une application linéaire, notée ici $f(a)$, sous réserve que l'on soit capable de montrer que le reste, noté $o(\|\mathbf{h}\|)$, soit suffisamment majoré en norme pour que l'approximation linéaire soit vraie.

Un point délicat invoqué ici est la notion de voisinage. Pour rappel :

Définition III.1.1. Soit E un espace topologique et $a \in E$. On appelle voisinage de a et on note $\mathcal{V}(a)$, tout ensemble qui contient un ouvert qui contient a .

Cela implique que pour parler de différentiabilité, il faut qu'il existe un voisinage de a , donc au moins un ouvert contenant a , dans lequel l'équation (III.1) est valable. Nous verrons plus tard que c'est une hypothèse forte. Cependant, nous pouvons montrer qu'il est toujours possible de parler de variation au voisinage d'un point en y ajoutant un vecteur (ici noté \mathbf{h}). En effet, si $\mathcal{V}(a)$ est un voisinage de a , alors il existe r réel strictement positif tel que la boule ouverte de centre a et de rayon r notée $B(a, r)$ soit incluse dans $\mathcal{V}(a)$. Dans ce cas, pour tout vecteur non nul \mathbf{h} , il est facile de vérifier que $\forall t \in]-\frac{r}{\|\mathbf{h}\|}, \frac{r}{\|\mathbf{h}\|}[$, $a + t\mathbf{h} \in B(a, r) \subset \mathcal{V}(a)$. Nous avons représenté cette opération sur la figure III.1.

Tout ce chapitre va consister à tourner autour de l'équation (III.1). La définition autour de laquelle va tourner tout le cours sera la suivante :

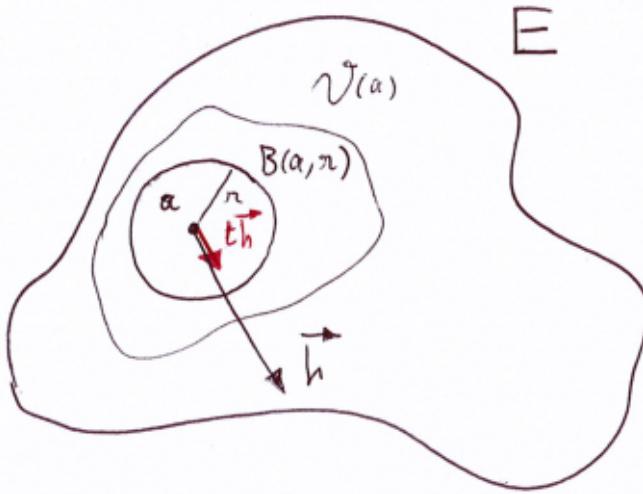


FIGURE III.1 – Représentation d'un ouvert E contenant un point a , un voisinage de a noté $\mathcal{V}(a)$, d'une boule ouverte de rayon r centrée sur a notée $B(a, r)$, et comment faire rentrer $a + t\mathbf{v}$ dans $B(a, r)$ et donc dans $\mathcal{V}(a)$.

Définition III.1.2. Soit une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n$, soit $a \in \Omega$. On dit que f est différentiable en a s'il existe une application linéaire L_a , \mathcal{V}_a un voisinage de a , tels que pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $a + \mathbf{h} \in \mathcal{V}_a$,

$$f(a + \mathbf{h}) = f(a) + L_a(\mathbf{h}) + o(\|\mathbf{h}\|) \quad (\text{III.3})$$

À titre de rappel :

Définition III.1.3. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$.

$$f(\mathbf{h}) = o(\|\mathbf{h}\|) \iff \exists \varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}^p / f(\mathbf{h}) = \|\mathbf{h}\| \varepsilon(\mathbf{h}), \quad \|\varepsilon(\mathbf{h})\| \xrightarrow[\|\mathbf{h} \rightarrow 0\|]{} 0. \quad (\text{III.4})$$

Dans ce cas on dit que f est négligeable devant \mathbf{h} .

$$f(\mathbf{h}) = \mathcal{O}(\|\mathbf{h}\|) \iff \exists C > 0, \eta > 0 / \|\mathbf{h}\| < \eta \Rightarrow \|f(\mathbf{h})\| < C\|\mathbf{h}\| \quad (\text{III.5})$$

Dans ce cas on dit que f est dominée par \mathbf{h} .

Si vous êtes perdus en calcul différentiel, rappelez-vous qu'il s'agit toujours de la même question fondamentale : comment se comporte cet objet (souvent une fonction) au voisinage d'un point ? Autrement dit : comment trouver une approximation linéaire de $f(a + \mathbf{h})$ du moment que la norme de \mathbf{h} est suffisamment petite ? Les questions plus subtiles et difficiles pour démontrer les théorèmes important consisteront à déterminer la petitesse nécessaire de la norme de \mathbf{h} pour que les conditions de réalisation du théorème en question soient valides.

III.2 Fonctions de plusieurs variables à valeur dans les réels

Dans cette section nous allons construire pas à pas la notion de dérivée partielle, sans nous soucier pour le moment de la différentiabilité des fonctions. Nous allons justement montrer quelques contre-exemples où les dérivées partielles peuvent être définies sans que la fonction ne soit différentiable.

III.2.1 Dérivée directionnelle, ou dérivée le long d'un vecteur

Définition III.2.1. Soit $f : E \rightarrow \mathbb{R}$, où $E \subset \mathbb{R}^n$ est un ouvert. Soit $a \in E$ et $\mathcal{V}(a)$ un voisinage de a . Soit $\mathbf{v} \in \mathbb{R}^n$. Soit I un intervalle ouvert de \mathbb{R} contenant 0 tel que $\forall t \in I, a + t\mathbf{v} \in \mathcal{V}(a)$.

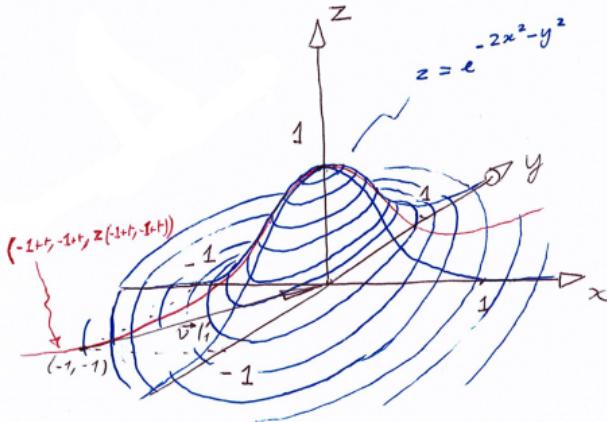


FIGURE III.2 – Représentation de la nappe d'équation $z = \exp(-2x^2 - y^2)$ et d'une ligne passant par cette surface, d'équation paramétrique $t \mapsto (-1 + t, -1 + t, z((-1 + t, -1 + t)))$.

On dit que f admet en a une dérivée le long du vecteur \mathbf{v} si la fonction d'une variable $\phi : I \rightarrow \mathbb{R}, t \mapsto f(a + t\mathbf{v})$, est dérivable en 0.

Si elle existe, on note

$$D_{\mathbf{v}}f(a) = \phi'(0) = \lim_{t \rightarrow 0} \frac{f(a + t\mathbf{v}) - f(a)}{t} \quad (\text{III.6})$$

la dérivée de f le long de \mathbf{v} en a .

Remarque Si f , définie sur un ouvert $E \subset \mathbb{R}^n$, admet une dérivée le long de \mathbf{v} en $a \in E$, alors

$$\exists \mathcal{V}(a), \exists r > 0 / \forall t \in]-r, r[, a + t\mathbf{v} \in \mathcal{V}(a), f(a + t\mathbf{v}) = f(a) + t D_{\mathbf{v}}f(a) + o(t) \quad (\text{III.7})$$

Cela veut dire qu'on peut écrire un développement limité de f le long de \mathbf{v} , mais cela n'implique pas nécessairement que f admette un développement limité sur un voisinage de a .

Exercices

- Soit une montagne d'équation $z = \exp(-2x^2 - y^2)$ (fig. III.2). Un promeneur se situe en $(-1, -1, e^{-3})$ et marche à une vitesse dont la projection sur le plan (xOy) a pour coordonnées $(1, 1)$. Calculer la composante verticale de sa vitesse. On admet que la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto \exp(-2x^2 - y^2)$ est différentiable en $(-1, -1)$ pour le moment.

Solution. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto \exp(-2x^2 - y^2)$. Dire que le promeneur se situe en $(-1, -1)$ et a une vitesse horizontale de composantes $(1, 1)$ revient à dire que, dans un voisinage de $(-1, -1)$, si on note $(x(t), y(t))$ ses coordonnées horizontales à la date t et qu'on choisit par convention $(x(0), y(0)) = (-1, -1)$, on peut trouver un intervalle I ouvert contenant 0 et une fonction $\varepsilon : I \rightarrow \mathbb{R}^2$ qui vérifie $\|\varepsilon(t)\| \xrightarrow[t \rightarrow 0]{} 0$ telle que l'on ait

$$\forall t \in I, (x(t), y(t)) = (-1, -1) + t(1, 1) + t\varepsilon(t) \quad (\text{III.8})$$

L'altitude du promeneur à la date t vaut

$$z(t) = f(x(t), y(t)) = f(-1 + t + t\varepsilon_x(t), -1 + t + t\varepsilon_y(t)) = \exp(-2(-1 + t + t\varepsilon_x(t))^2 - (-1 + t + t\varepsilon_y(t))^2) \quad (\text{III.9})$$

Comme on a admis que f était différentiable en $(-1, -1)$, on peut admettre l'existence d'une application linéaire L_a telle que

$$f((-1, -1) + t(1, 1) + t\varepsilon(t)) = f(-1, -1) + L_a((t, t) + t\varepsilon(t)) + o(((t, t) + t\varepsilon(t))) \quad (\text{III.10})$$

mais comme L_a est linéaire et qu'en dimension finie toutes les applications linéaires sont continues, on est assurés que¹ $L_a(t\varepsilon(t)) = o(t)$. Par ailleurs il est évident que $o(t\varepsilon(t)) = o(t)$. Ainsi on peut écrire

$$z(t) = f(-1, -1) + L_a(t, t) + o(t). \quad (\text{III.11})$$

Par unicité du développement limité, on est assuré que $L_a(t, t)$ sera bien égal à $D_{(1,1)}f(-1, -1)$. Ainsi il suffit de calculer $\phi'(0)$ où $\phi(t) = f(-1+t, -1+t)$. Par unicité du développement limité, on aura évidemment $z'(0) = \phi'(0)$. On calcule aisément que

$$\phi(t) = \exp(-2t^2 + 6t - 3) \quad \phi'(t) = (-4t + 6) \exp(-2t^2 + 6t - 3), \quad (\text{III.12})$$

et donc $\phi'(0) = 6e^{-3} = D_{(1,1)}f(-1, -1) = z'(0)$.

Bien entendu il ne sera pas nécessaire d'être aussi immondement laborieux à chaque fois, cette proposition de correction était surtout pour montrer comment la majoration de norme permet d'identifier la fonction et son approximation linéaire lorsque l'on dérive.

2. Le but de cet exercice est de montrer que certaines fonctions peuvent n'être dérivables directionnellement presque nulle part.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto |2x - y|$. Soit $\mathbf{w} = (u, v) \in \mathbb{R}^2$ avec $\|\mathbf{w}\| = 1$. À quelles conditions sur \mathbf{w} f est-elle dérivable le long de \mathbf{w} en $(0, 0)$?

Solution. Soit $t \neq 0$. La fonction f est continue sur \mathbb{R}^2 et pour tout t réel non nul, on a

$$\frac{f(t\mathbf{w}) - f(0, 0)}{t} = \frac{|2tu - tv|}{t} = \frac{|t|}{t} |2u - v| = \text{signe}(t) |2u - v| \quad (\text{III.13})$$

Si $2u \neq v$, cette expression n'admet pas de limite en 0 car elles diffèrent selon que t est positif ou négatif. Si $2u = v$, alors la dérivée le long de \mathbf{w} existe et vaut 0. Comme $\|\mathbf{w}\|^2 = 1 = u^2 + v^2$, cela ne laisse que deux choix possibles : $u = 1/\sqrt{5}$, $v = 2/\sqrt{5}$, ou $u = -1/\sqrt{5}$ et $v = -2/\sqrt{5}$.

Remarques

- De cet exercice on voit qu'en un point, une fonction peut admettre des dérivées directionnelles selon certains vecteurs et non d'autres.
- Une fonction peut, en un point, être dérivable le long de tout vecteur, sans même être continue ! Cela illustre que la continuité est une hypothèse forte. Pour rappel, une fonction f définie sur un ouvert E est continue en a si et seulement si

$$\forall \varepsilon > 0, \exists r > 0 / \|b - a\| < r \Rightarrow b \in E, \|f(b) - f(a)\| < \varepsilon \quad (\text{III.14})$$

La continuité en un point demande d'être proche de ce point en norme quelque soit le chemin par lequel on s'approche de ce point. Nous allons voir que les droites qui convergent vers un point ne suffisent pas.

Exercice

Soit

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$(x, y) \mapsto \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon.} \end{cases} \quad (\text{III.15})$$

Montrer que f est dérivable en 0 le long de tout vecteur mais qu'elle n'est pas continue en 0.

Solution. Soit $\mathbf{v} = (u, v) \in \mathbb{R}^2$. On calcule

$$\forall t \in \mathbb{R}^*, \frac{f(t\mathbf{v}) - f(0)}{t} = \frac{uv^2}{u^2 + t^2 v^4} \xrightarrow[t \rightarrow 0]{} \begin{cases} \frac{v^2}{u} & \text{si } u \neq 0 \in E \\ 0 & \text{si } u = 0 \end{cases} \quad (\text{III.16})$$

Donc $D_{\mathbf{v}}f(0, 0)$ existe pour tout \mathbf{v} . D'autre part, on constate que d'une part

$$f(t, t) = \frac{t^3}{t^2 + t^4} \xrightarrow[t \rightarrow 0]{} 0 \quad (\text{III.17})$$

1. En effet, si on pose, pour toute application linéaire u sur un espace vectoriel H $\|u\| = \sup_{x \in H} \frac{\|u(x)\|}{\|x\|}$, un résultat classique d'algèbre montre que ce nombre existe et est fini pour toute application linéaire sur H . Dans ce cas, on a $\|L_a(t\varepsilon(t))\| \leq \|L_a\| |t| \|\varepsilon(t)\| = o(t)$.

et d'autre part

$$f(t^2, t) = \frac{1}{2} \xrightarrow[t \rightarrow 0]{} \frac{1}{2} \neq 0 \quad (\text{III.18})$$

ce qui montre que l'on atteint une autre valeur si on approche $(0, 0)$ par une parabole et non une droite. La fonction f n'est donc pas continue. Elle n'est donc pas différentiable non plus, puisque la différentiabilité implique la continuité.

III.2.2 Dérivée partielle

Définition III.2.2. Soit $f : \mathcal{U} \rightarrow \mathbb{R}$ où $\mathcal{U} \subset \mathbb{R}^n$ est un ouvert. Soit $(\mathbf{e}_i)_{i \in [1, n]}$ la base canonique de \mathbb{R}^n . Soit $a \in \mathcal{U}$. On appelle *dérivée partielle de f en a selon la i -ième coordonnée, ou i -ième dérivée partielle de f en a* , si elle est définie, le nombre $D_{\mathbf{e}_i} f(a)$, que l'on note (selon les différents ouvrages) :

$$D_{\mathbf{e}_i} f(a) = \frac{\partial f}{\partial x_i}(a) = \partial_i f(a) = D_i f(a) = \partial_{x_i} f(a) = \dots \quad (\text{III.19})$$

Dans ce cours nous utiliserons principalement les notations $\partial f / \partial x_i$ lorsque la spécification des variables sera nécessaire, et $\partial_i f$ lorsqu'aucune ambiguïté ne sera possible.

Remarque On a donc, lorsqu'elle existe

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a_1, \dots, a_n)}{t} \quad (\text{III.20})$$

Gradient Supposons que $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ soit différentiable en a : d'après la définition III.1.2, il existe un voisinage de a , une application linéaire L_a tel que f admette le développement limité suivant

$$f(a + \mathbf{h}) = f(a) + L_a \mathbf{h} + o(\|\mathbf{h}\|) \quad (\text{III.21})$$

Soit $\mathcal{V}(a)$ un voisinage de a dans l'ensemble de définition de f tel que le développement limité de la différentielle soit valide. Pour tout vecteur \mathbf{v} , on se donne $r > 0$ tel que $\forall t \in]-r, r[, a + t\mathbf{v} \in \mathcal{V}(a)$. Soit $(\mathbf{e}_i)_i$ la base canonique de \mathbb{R}^n . Si $\mathbf{v} = \mathbf{e}_i$, alors

$$f(a + t\mathbf{e}_i) = f(a) + t \frac{\partial f}{\partial x_i}(a) + o(t) \quad (\text{III.22})$$

Si $\mathbf{v} = \sum_i v_i \mathbf{e}_i$, alors

$$f(a + t\mathbf{v}) = f(a) + tL_a \mathbf{v} + o(t) = f(a) + t \left(\sum_i v_i \mathbf{e}_i \right) + o(t) = f(a) + t \sum_i v_i L_a(\mathbf{e}_i) + o(t) \quad (\text{III.23})$$

Mais on sait que $L_a(\mathbf{e}_i) = \partial_i f(a)$, d'où

$$f(a + t\mathbf{v}) = f(a) + t \sum_i \frac{\partial f}{\partial x_i}(a) v_i = +o(t) \quad (\text{III.24})$$

Donc ici, $df(a)$ apparaît comme une forme linéaire définie de la façon suivante

$$\begin{aligned} \mathfrak{f}(a) : \quad & \mathbb{R}^n \rightarrow \mathbb{R} \\ \mathbf{v} = \sum_i v_i \mathbf{e}_i \mapsto & \sum_i \frac{\partial f}{\partial x_i}(a) v_i \end{aligned} \quad (\text{III.25})$$

Comme à toute forme linéaire de \mathbb{R}^n correspond un unique vecteur, on peut appeler gradient de f en a l'unique vecteur noté $\nabla f(a)$ tel que $\mathfrak{f}(a) = \langle \nabla f(a), \cdot \rangle$, et les coordonnées de $\nabla f(a)$ sont égales aux dérivées partielles de f en a .

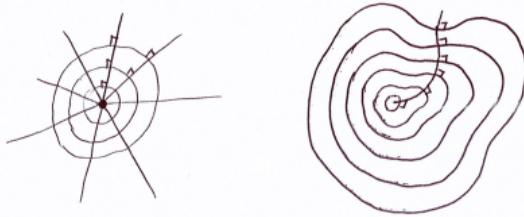


FIGURE III.3 – Quelques exemples de lignes de niveaux de fonctions de deux variables à valeurs dans \mathbb{R} . Dans ce cas, les lignes de niveau sont également les surfaces de niveau. À gauche : les lignes de niveau du potentiel électrostatique engendré par une charge ponctuel sont des cercles concentriques, et les lignes de champ sont les droites qui concourent en la particule. À droite : lignes de niveau d'une fonction différentiable quelconque avec une ligne de champ.

Remarques

1. Matriciellement, si f est à variables dans \mathbb{R} , dans la base canonique on a $\nabla f = {}^t \mathbf{f}$. En effet

$$\forall \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n, \mathbf{f}(a)\mathbf{v} = (\partial_1 f(a), \dots, \partial_n f(a)) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = {}^t \nabla f(a) \mathbf{v} \quad (\text{III.26})$$

2. Certains ouvrages notent $\mathbf{grad} f$ au lieu de ∇f pour éviter la tentation de faire des raccourcis hâtifs en notant $\nabla = {}^t(\partial_1, \dots, \partial_n)$, égalité fausse dans un système de coordonnées quelconque. Cette relation n'est vraie que dans la base canonique d'un espace euclidien².

Définition III.2.3. Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, telle que $\forall x \in \Omega, \nabla f(x) \neq \mathbf{0}$.

- On appelle *lignes de niveau de f* l'ensemble des courbes régulières de \mathbb{R}^n telles que f ne varie pas le long de ces lignes.
- On appelle *surfaces de niveau si $n = 3$ ou hypersurface de niveau si $n > 3$* , l'ensemble des sous-variétés de \mathbb{R}^n de dimension $n - 1$ telles que f ne varie pas sur ces sous-variétés.
- On appelle *lignes de champ* l'ensemble des courbes qui sont tangentées à ∇f en tout point.

Nous représentons quelques exemples de lignes de niveau sur la figure III.3. Un exemple connu d'utilisation des lignes de niveau sont les cartes de randonnées.

Exercices Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ telle que $\forall x \in \Omega, \nabla f(x) \neq \mathbf{0}$.

1. Montrer qu'en tout point de Ω , le gradient de f est orthogonal aux lignes de niveau.
2. Montrer que le gradient de f correspond localement à la direction d'accroissement maximal de f .
3. On reprend la montagne d'équation $z = e^{-2x^2-y^2}$ et son promeneur situé en $(-1, -1)$. Dans quelle direction devra aller le promeneur pour monter localement le plus vite possible ?

Solution.

1. Soit $a \in \Omega$, soit \mathcal{L} une ligne de niveau qui passe par a . Si \mathbf{v} est un vecteur directeur de \mathcal{L} en a , alors il existe un paramétrage de \mathcal{L} , c'est-à-dire un intervalle I ouvert connexe de \mathbb{R} , et une fonction $\gamma : I \rightarrow \Omega$ avec $\gamma(I) = \mathcal{L}$, on posera $\gamma(0) = a$, et il existe une fonction ϵ qui tend (en norme) vers 0 en 0 telle que

$$\gamma(t) = a + t\mathbf{v} + t\epsilon(t) \quad (\text{III.27})$$

Sur la ligne de niveau, on a

$$f(a) = f(\gamma(t)) \quad (\text{III.28})$$

$$= f(a + t\mathbf{v} + t\epsilon(t)) \quad (\text{III.29})$$

2. Nous verrons ultérieurement qu'il en est autrement dans les variétés différentiables.

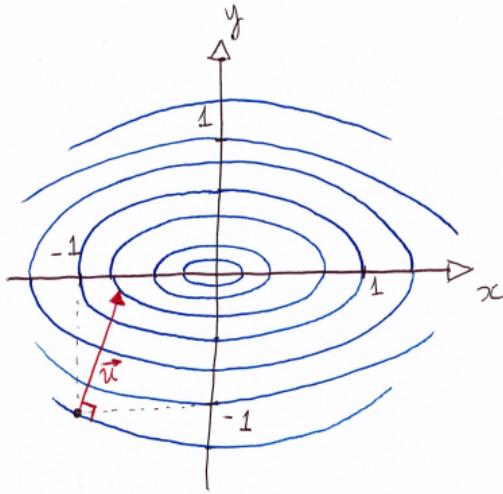


FIGURE III.4 – Lignes de niveau de la montagne d'équation $z = e^{-2x^2-y^2}$. On voit que le vecteur \mathbf{u} est bien orthogonal à la ligne de niveau.

$$= f(a) + t\langle \nabla f(a), t\mathbf{v} \rangle + t\langle \nabla f(a), \epsilon(t) \rangle + o(t) \quad (\text{III.30})$$

Comme $\|\epsilon(t)\| \xrightarrow[t \rightarrow 0]{} 0$, $t\langle \nabla f(a), \epsilon(t) \rangle = o(t)$ d'où

$$t\langle \nabla f(a), \mathbf{v} \rangle + o(t) = 0 \quad (\text{III.31})$$

Par unicité du développement limité, on en déduit $\langle \nabla f(a), \mathbf{v} \rangle = 0$, ce qui était à démontrer.

2. Soit une ligne passant par a et dirigée par un vecteur unitaire \mathbf{u} en a . Par le même raisonnement, on peut montrer que

$$f(\gamma(t)) = f(a) + t\langle \nabla f(a), \mathbf{u} \rangle + o(t) \quad (\text{III.32})$$

Si $t > 0$, l'accroissement local de f dépendra du terme $\langle \nabla f(a), \mathbf{u} \rangle$. Ou bien \mathbf{u} est colinéaire à $\nabla f(a)$, auquel cas on a nécessairement $\mathbf{u} = \pm \nabla f(a)/\|\nabla f(a)\|$, et pour maximiser l'accroissement il faut choisir $+$. Dans ce cas, on a

$$\langle \nabla f(a), \mathbf{u} \rangle = \|\nabla f(a)\| \quad (\text{III.33})$$

Ou bien \mathbf{u} n'est pas colinéaire à $\nabla f(a)$. Dans ce cas, on complète $\nabla f(a)/\|\nabla f(a)\|$ par $n - 1$ vecteurs pour faire une base orthonormée de \mathbb{R}^n . Dans ce cas, on a

$$\mathbf{u} = u_\nabla \nabla f(a)/\|\nabla f(a)\| + \sum_{i=1}^{n-1} u_i \mathbf{e}_i \quad (\text{III.34})$$

Comme $\|\mathbf{u}\| = 1$, nécessairement, $|u_\nabla| < 1$. Dans ce cas, on a

$$\langle \nabla f(a), \mathbf{u} \rangle = u_\nabla \|\nabla f(a)\| < \|\nabla f(a)\| \quad (\text{III.35})$$

Cela prouve que $\nabla f(a)$ est la direction qui maximise l'accroissement de f .

3. La direction que doit suivre le promeneur est bien entendu $\mathbf{u} = \nabla z(-1, -1)/\|\nabla z(-1, -1)\| = (2, 1)/\sqrt{5}$. Pour descendre le plus vite possible il faut aller vers $-\mathbf{u}$. Nous illustrons ce résultat dans la figure III.4.

III.3 Différentiabilité

Cette fois-ci on considère des fonctions à n variables à valeurs dans \mathbb{R}^p .

III.3.1 Différentielle

Dans cette section nous voulons généraliser et formaliser ce qui a été abordé de façon heuristique dans la première section. Rappelons la définition de la différentiabilité :

Définition III.3.1. Soit une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n$, soit $a \in \Omega$. On dit que f est différentiable en a s'il existe une application linéaire L_a , \mathcal{V}_a un voisinage de a , tels que pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $a + \mathbf{h} \in \mathcal{V}_a$,

$$f(a + \mathbf{h}) = f(a) + L_a(\mathbf{h}) + o(\|\mathbf{h}\|) \quad (\text{III.36})$$

Théorème III.3.2. (Définition). Si f est différentiable en a , alors avec les mêmes notations, L_a est unique. On appelle différentielle de f en a l'application linéaire ainsi définie : $\dot{f}(a) = L_a$.

Démonstration. Par exemple, supposer l'existence de L_1 et L_2 distinctes et calculer la limite $\lim_{t \rightarrow 0} \frac{f(a + t\mathbf{v}) - f(a)}{t}$ dans les deux cas et conclure. ■

Théorème III.3.3. Soit une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n$, soit $a \in \Omega$. Si f est différentiable en a , alors pour tout $\mathbf{h} \in \mathbb{R}^n / a + \mathbf{h} \in \mathcal{V}_a$ où \mathcal{V}_a est un voisinage de a tel que le développement limité à l'ordre 1 de f soit valable, on a

$$f(a + \mathbf{h}) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i + o(\|\mathbf{h}\|) \quad (\text{III.37})$$

où

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + t\mathbf{e}_i) - f(a)}{t} \quad (\text{III.38})$$

est la dérivée partielle de f en a selon la i -ième coordonnée.

Démonstration. C'est la même preuve que le gradient ! Il suffit de décomposer \mathbf{h} sur la base canonique. ■

Attention, ici $\partial_i f(a) \in \mathbb{R}^p$ et non \mathbb{R} , car f est à valeur dans \mathbb{R}^p . Dans ces conditions, la différentielle de f en a est une application linéaire de \mathbb{R}^n dans \mathbb{R}^p et vérifie

$$\dot{f}(a) : \mathbb{R}^n \rightarrow \mathbb{R}^p, \quad \mathbf{h} = \sum_i h_i \mathbf{e}_i \mapsto \sum_i \frac{\partial f}{\partial x_i}(a) h_i \quad (\text{III.39})$$

Représentation matricielle Si f est à valeur dans \mathbb{R}^p et différentiable en a on a

$$f(a + \mathbf{h}) = \begin{pmatrix} f_1(a_1 + h_1, \dots, a_n + h_n) \\ \vdots \\ f_p(a_1 + h_1, \dots, a_n + h_n) \end{pmatrix} = \begin{pmatrix} f_1(a_1, \dots, a_n) \\ \vdots \\ f_p(a_1, \dots, a_n) \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^n \partial_i f_1(a) h_i \\ \vdots \\ \sum_{i=1}^n \partial_i f_p(a) h_i \end{pmatrix} + o(\mathbf{h}) \quad (\text{III.40})$$

Le deuxième terme est une application linéaire qu'on peut représenter matriciellement. La matrice représentative de $\dot{f}(a)$ est appelée la matrice jacobienne de f en a relative à un système de coordonnées (ici la base canonique) et s'exprime ainsi

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_j}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & & & \vdots \\ \frac{\partial f_i}{\partial x_1}(a) & \dots & \frac{\partial f_i}{\partial x_j}(a) & \dots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & & & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_j}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \quad (\text{III.41})$$

En effet on vérifie que

$$f(a + \mathbf{h}) = \begin{pmatrix} f_1(a_1, \dots, a_n) \\ \vdots \\ f_p(a_1, \dots, a_n) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_j}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & & & \vdots \\ \frac{\partial f_i}{\partial x_1}(a) & \dots & \frac{\partial f_i}{\partial x_j}(a) & \dots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & & & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_j}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_j \\ \vdots \\ h_n \end{pmatrix} + o(\mathbf{h}) \quad (\text{III.42})$$

Ici le vecteur colonne \mathbf{h} a n lignes, et la matrice $J_f(a)$ a p lignes et n colonnes. Retenons que la matrice jacobienne a pour composantes à la i -ième ligne et à la j -ième colonne

$$J_f(a)_{[i,j]} = \frac{\partial f_i}{\partial x_j}(a) \quad (\text{III.43})$$

ou pas... en cas d'oubli il suffit de développer $f(a + \mathbf{h})$ et de représenter matriciellement l'application linéaire obtenue pour retrouver immédiatement l'expression de la jacobienne.

Exemple Soit f la fonction définie ainsi

$$\begin{aligned} f :]-\pi, \pi[\times \mathbb{R} &\rightarrow \mathbb{R}^3 \\ \begin{pmatrix} \theta \\ u \end{pmatrix} &\mapsto \begin{pmatrix} \cos \theta \cosh u \\ \sin \theta \cosh u \\ \sinh u \end{pmatrix} \end{aligned} \quad (\text{III.44})$$

où \cosh et \sinh sont les fonctions hyperboliques. On admet que f est différentiable en tout point de $]-\pi, \pi[\times \mathbb{R}$. On calcule tout simplement

$$J_f(\theta, u) = \begin{pmatrix} -\sin \theta \cosh u & \cos \theta \sinh u \\ \cos \theta \cosh u & \sin \theta \sinh u \\ 0 & \cosh u \end{pmatrix} \quad (\text{III.45})$$

En fait, comme on le verra plus tard, f est une immersion qui permet de représenter une surface plongée dans \mathbb{R}^3 . Si on impose $\theta = cst$, on voit que dans le plan vertical qui fait un angle θ avec l'axe Ox , on trace une branche d'hyperbole. Une fois cela fait, lorsque l'on fait varier θ , cela fait une surface de révolution autour de l'axe Oz . Il suffit donc de faire tourner une branche d'hyperbole d'équation cartésienne $x^2 - z^2 = 1$ autour de l'axe Oz pour dessiner cette figure. De plus, les deux colonnes de la matrice jacobienne, $\partial_\theta f$ et $\partial_u f$, sont des vecteurs tangents à cette surface : ils indiquent comment un point varie localement lorsqu'on fait varier localement les coordonnées θ et u . Toute cette heuristique sera développée formellement et rigoureusement dans le chapitre sur les surfaces régulières. En attendant on peut s'en faire une approche intuitive avec la figure III.5

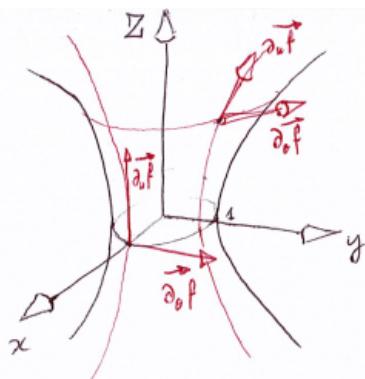


FIGURE III.5 – On représente la surface plongée dont le paramétrage est décrit par l'équation (III.44). Nous avons représenté les vecteurs colonnes du jacobien (eq. (III.45)), $\partial_\theta f(\theta, u)$ et $\partial_u f(\theta, u)$ en deux points de la surface. Nous voyons qu'ils sont tangents aux lignes $u = cst$ et $\theta = cst$ respectivement.

III.3.2 Propriétés

- Si f est différentiable en a , alors f est continue en a . C'est immédiat car toutes les applications linéaires sont continues en dimension finie. En dimension infinie, il faudra ajouter à la définition de la différentiabilité que l'application linéaire en question soit continue.
- Corollaire : si f n'est pas continue en a , elle n'est pas différentiable en a .
- f peut admettre des dérivées partielles en a voire même des dérivées le long de tout vecteur en a , sans être différentiable ! On a déjà vu que cela était possible sans même que f ne soit continue avec l'exemple $f(x, y) = xy^2/(x^2 + y^4)$ si $(x, y) \neq 0$ et 0 sinon.
- Si f est différentiable en a , alors

$$\frac{f(a + \mathbf{h}) - f(a) - \mathbf{f}(a)\mathbf{h}}{\|\mathbf{h}\|} \xrightarrow[\|\mathbf{h}\| \rightarrow 0]{} 0 \quad (\text{III.46})$$

III.3.3 Condition suffisante

Nous allons (enfin) donner une condition suffisante pour vérifier que f est différentiable en a .

Théorème III.3.4. Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$, $a \in \Omega$. S'il existe un voisinage ouvert de a tel que toutes les dérivées partielles de f existent et sont continues en a , alors f est différentiable en a .

La preuve de ce théorème est technique mais pas insurmontable. Nous préférons la reporter en annexe de ce chapitre pour éviter d'alourdir la première lecture de l'exposé. L'idée de la preuve consiste tout simplement à majorer le reste en cheminant d'abord de (a_1, a_2) à $(a_1 + h_1, a_2)$ puis $(a_1 + h_1, a_2 + h_2)$, en utilisant le théorème des accroissements finis pour exprimer le reste à chaque fois. L'idée est ensuite de majorer la norme de la différence avec la différentielle pour conclure que le développement limité peut s'écrire dans un voisinage de a . Nous recommandons aux lecteurs motivés d'essayer de faire et/ou de lire la preuve en deuxième lecture de ce cours s'ils découvrent le calcul différentiel en lisant ce texte, ou de le faire d'emblée s'ils ont déjà suivi des cours.

Définition III.3.5. On appelle fonction de classe \mathcal{C}^1 sur Ω l'ensemble des fonctions différentiables en tout point de Ω . On note $\mathcal{C}^1(\Omega)$ cet ensemble.

On a $\mathcal{C}^1(\Omega) \subset \mathcal{C}^0(\Omega)$.

Exemples

- $(x, y, z) \mapsto (xyz - 42x^{42}, z^{2020})$ est \mathcal{C}^1 sur \mathbb{R}^3 .
- $(x, y) \mapsto \sqrt{x^2 + y^2}$ est \mathcal{C}^0 sur \mathbb{R}^2 et \mathcal{C}^1 sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.
- $(x, y) \mapsto 1/\sqrt{x^2 + y^2}$ est \mathcal{C}^0 et \mathcal{C}^1 sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Une fois qu'on sait que la fonction est différentiable sur tout un ensemble on peut essayer d'en dégager des propriétés. En voici une.

Théorème III.3.6. Soit $f : \mathcal{U} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^p$. Si \mathcal{U} est un ouvert convexe alors les deux propositions suivantes sont équivalentes

III.3.6.1. $\forall (x, y) \in \mathcal{U}, \partial_x f(x, y) = 0$

III.3.6.2. Il existe une fonction d'une variable g, \mathcal{C}^1 , telle que $\forall (x, y) \in \mathcal{U}, f(x, y) = g(y)$

Démonstration. $2 \Rightarrow 1$ est évident, montrons $1 \Rightarrow 2$. Supposons les hypothèses vérifiées. Soient (x, y) et (x', y) dans \mathcal{U} . Pour tout t dans $[x, x']$, $(t, y) \in \mathcal{U}$ grâce à sa convexité. D'autre part on a

$$f(x', y) = f(x, y) + \int_x^{x'} \partial_t f(t, y) dt = f(x, y) \quad (\text{III.47})$$

ce qui montre que la fonction dont le graphe est $g(x) = f(x, y)$ est bien une fonction d'une variable qui satisfait ce qui est à démontrer. ■

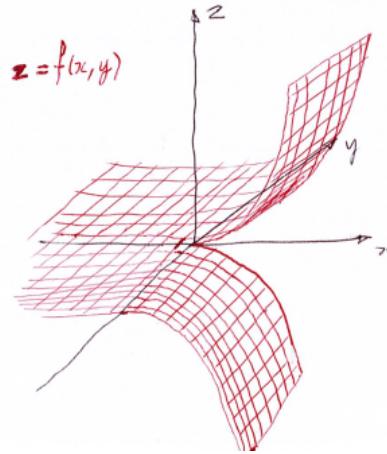


FIGURE III.6 – Représentation du graphe de la fonction définie à l'équation (III.48). On voit clairement que toutes les dérivées le long de la direction y sont nulles mais que ça n'empêche aucunement la fonction de dépendre de y à cause de la non convexité de l'ensemble de définition.

Ce genre de théorème est très utile en physique, d'ailleurs les physiciens ne se privent pas d'écrire des choses du genre " $\partial_x f(x, y) = 0 \Rightarrow f(y)$ " qui est faux d'un point de vue strictement mathématique (car les deux objets désignés par f ont une nature mathématique différente) mais correct du point de vue de la physique si f désigne une quantité et non une fonction mathématique.

La convexité est une hypothèse très importante. Voici un contre-exemple :

$$f : \mathbb{R}^2 \setminus \{(x, y) | y = 0, x \geq 0\} \rightarrow \mathbb{R}$$

$$(x, y) \mapsto \begin{cases} x^2 & \text{si } x \geq 0, y > 0 \\ -x^2 & \text{si } x \geq 0, y < 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (\text{III.48})$$

On peut vérifier que f est \mathcal{C}^1 sur son ensemble de définition et que partout, on a $\partial_y f = 0$. Pourtant, on a $f(1, 1) = 1$ et $f(1, -1) = -1$, on ne peut donc pas trouver une fonction ne dépendant que de x qui suffirait à décrire le comportement de f . Ici, c'est la non convexité de Ω qui nous empêche de passer d'un point à l'autre par une droite. Nous avons représenté ce contre-exemple sur la figure III.6. Nous montrerons comment ce théorème peut servir à résoudre des équations aux dérivées partielles avec l'exemple de l'équation d'onde (voir sec. III.4.5).

III.4 Composition et opération sur les différentielles

III.4.1 Opérations élémentaires

Proposition III.4.1. Soient $\alpha : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, f et $g : \Omega \rightarrow \mathbb{R}^p$, avec Ω ouvert. Soit $a \in \Omega$ tel que α, f, g soient différentiables en a . Alors $\alpha f + g$ est différentiable en a et on a

$$d(\alpha f + g)(a) = f(a)d\alpha(a) + \alpha(a)df(a) + dg(a) \quad (\text{III.49})$$

Matriciellement, on a

$$f(a)^T \nabla \alpha(a) + \alpha(a) J_f(a) + J_g(a) = \begin{pmatrix} f_1(a) \\ \vdots \\ f_p(a) \end{pmatrix} (\partial_1 \alpha(a), \dots, \partial_n \alpha(a)) + \alpha(a) \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix}$$

$$+ \begin{pmatrix} \partial_1 g_1(a) & \dots & \partial_n g_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 g_p(a) & \dots & \partial_n g_q(a) \end{pmatrix} \quad (\text{III.50})$$

Le premier terme de cette équation est une matrice de taille $p \times n$ dont le coefficient i, j vaut $f_i(a)\partial_j\alpha(a)$. Appliquée à un vecteur h , on a

$$\begin{pmatrix} f_1(a) \\ \vdots \\ f_p(a) \end{pmatrix} (\partial_1\alpha(a), \dots, \partial_n\alpha(a)) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n f_1(a)\partial_i\alpha(a)h_i \\ \vdots \\ \sum_{i=1}^n f_p(a)\partial_i\alpha(a)h_i \end{pmatrix} \quad (\text{III.51})$$

Démonstration. Comme α, f, g sont différentiables en a , on peut trouver un voisinage de a sur lequel les trois fonctions admettent un développement limité à l'ordre 1 (il suffit de prendre l'intersection des trois voisinages). Soit \mathcal{V}_a un tel voisinage, soit $h \in \mathbb{R}^n$ tel que $a + h \in \mathcal{V}_a$. On a

$$\begin{aligned} \alpha(a + h)f(ah) + g(a + h) &= (\alpha(a) + d\alpha(a)h + o(h))(f(a) + df(a)h + o(h)) + g(a) + dg(a)h + o(h) \\ &= \alpha(a)f(a) + g(a) + (d\alpha(a)h)f(a) + \alpha(a)df(a)h + dg(a)h \\ &\quad + f(a)o(h) + df(a)h o(h) + \alpha(a)o(h) + d\alpha(a)h o(h) + o(h) \end{aligned} \quad (\text{III.52})$$

$$(III.53)$$

Comme les normes de $f(a)$ et $\alpha(a)$ sont finies, $f(a)o(h) + \alpha(a)o(h) = o(h)$. Comme les applications linéaires en dimension finie ont des normes finies on a $\|df(a)h\| \leq |||df(a)||| \|h\|$ et $|d\alpha(a)h| \leq \|\nabla\alpha(a)\| \|h\|$, donc $d\alpha(a)h o(h) = o(h)$ et $df(a)h o(h) = o(h)$. De telle sorte que

$$\alpha(a + h)f(ah) + g(a + h) = \alpha(a)f(a) + g(a) + f(a)d\alpha(a)h + \alpha(a)df(a)h + dg(a)h + o(h) \quad (\text{III.54})$$

$h \mapsto f(a)(d\alpha(a)h) + \alpha(a)df(a)h + dg(a)h$ est une application linéaire, donc $\alpha f + g$ est différentiable en a et on a l'égalité demandée. ■

III.4.2 Composition des différentielles

Théorème III.4.2. Soient $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathcal{V} \subset \mathbb{R}^p$, $g : \mathcal{V} \rightarrow \mathbb{R}^q$ avec \mathcal{U}, \mathcal{V} , ouverts. Soit $a \in \mathcal{U}$. Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et on a

$$d(g \circ f)(a) = dg(f(a)) \circ df(a) \quad (\text{III.55})$$

Démonstration. g est différentiable en $f(a)$ donc il existe une boule ouverte de rayon ϵ dans laquelle pour tout vecteur ℓ tel que $f(a) + \ell$ est dans ce voisinage,

$$g(f(a) + \ell) = g(f(a)) + dg(f(a))\ell + o(\ell) \quad (\text{III.56})$$

D'autre part, comme f est différentiable en a on a l'existence d'un voisinage ouvert de a tel que tant que $a + h$ est dans ce voisinage,

$$f(a + h) = f(a) + df(a)h + o(h) \quad (\text{III.57})$$

Comme $df(a)$ est continue, on sait que $df(a)h + o(h) \xrightarrow{\|h\| \rightarrow 0} 0$ donc il existe une boule ouverte de rayon r telle que tant que $\|h\| < r$, $\|df(a)h + o(h)\| < \epsilon$. Il existe donc un voisinage de a tel que tant que $a + h$ est dans ce voisinage,

$$g(f(a + h)) = g(f(a) + df(a)h + o(h)) = g(f(a)) + dg(f(a))(df(a)h + o(h)) + o(df(a)h + o(h)) \quad (\text{III.58})$$

Comme $\forall h \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\|df(a)h\|/\|h\| \leq |||df(a)|||$, on a

$$o(\|df(a)h + o(h)\|) = o\left(\|h\| \frac{\|df(a)h\|}{\|h\|} + o(h)\right) = o(\|h\| |||df(a)||| + o(h)) = o(h) \quad (\text{III.59})$$

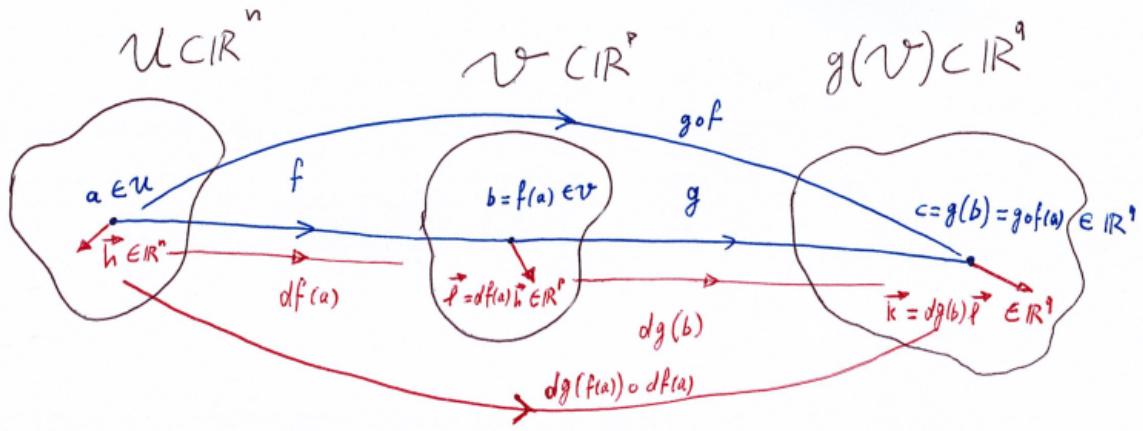


FIGURE III.7 – Représentation symbolique de la composition des différentielles. En bleu : composition de f et g . En rouge : composition de $df(a)$ et $dg(f(a))$

De plus,

$$dg(f(a))o(h) \leq |||dg(f(a))|||o(h) = o(h) \quad (\text{III.60})$$

D'où l'égalité demandée. ■

Les deux démonstrations des deux théorèmes précédents illustrent bien ce qu'annonçait l'introduction de ce cours. Le calcul différentiel consiste à majorer suffisamment des normes pour que l'algèbre linéaire puisse s'appliquer localement.

Matrices La représentation matricielle de $dg(f(a)) \circ df(a)$ est tout simplement le produit des jacobiniennes. On a donc

$$J_g(f(a))J_f(a)h = \begin{pmatrix} \partial_1 g_1(f(a)) & \dots & \partial_p g_1(f(a)) \\ \vdots & \ddots & \vdots \\ \partial_1 g_q(f(a)) & \dots & \partial_p g_q(f(a)) \end{pmatrix} \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \quad (\text{III.61})$$

La première matrice a q lignes et p colonnes, la deuxième matrice a p lignes et n colonnes, et la dernière a évidemment n lignes. Le produit de ces deux matrices est une matrice à q lignes et n colonnes et a pour composante i, j

$$(J_g(f(a))J_f(a))_{[i,j]} = \sum_{k=1}^p \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a) \quad (\text{III.62})$$

Appliqué au vecteur h , le résultat d'une telle opération est un vecteur de p colonnes donc la i -ième ligne vaut

$$(J_g(f(a))J_f(a)h)_i = \sum_{j=1}^n \sum_{k=1}^p \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a) h_j \quad (\text{III.63})$$

Ici nous avons posé $\partial g_i / \partial y_j = \partial_j g_i$ avec la lettre y et non x pour éviter la confusion entre les deux espaces vectoriels distincts sur lesquels portent la dérivation. La différentielle de f en a agit sur un vecteur de \mathbb{R}^n et le transporte linéairement dans \mathbb{R}^p , alors que la différentielle de g en $b = f(a)$ agit sur un vecteur de \mathbb{R}^p et le transporte linéairement dans \mathbb{R}^q . Voir la figure III.7.

III.4.3 Difféomorphisme et changement de variables

Il arrive souvent aux physiciens de noter abusivement $y = y(x)$ puis $f(x) = f(y)$ puis $df/dx = df/dy \times dy/dx$ pour noter un changement de variables. Par exemple, considérons la trajectoire d'une pierre en chute libre dans le champ de pesanteur au voisinage de la surface de la Terre lancée à 45 degrés vers le ciel. Il est possible de paramétriser l'altitude z de la pierre avec le temps, ou avec l'espace. Les physiciens noteront $z = z(t) = z(x)$. Pourtant, la fonction $t \mapsto z$ est différente de la fonction $x \mapsto z$ car la dépendance en x et en t est différente, l'équation $z(x) = z(t)$ est donc nécessairement fausse. Il serait plus rigoureux de noter \tilde{z} l'altitude exprimée comme fonction du temps et z l'altitude exprimée comme fonction de l'espace, et relier les deux ainsi : $\tilde{z}(t) = z(x(t))$ par exemple. Cependant, les physiciens s'intéressent plus aux quantités mesurables qu'aux mathématiques (ce qui est parfaitement leur droit) et à ce titre ils ont raison de nommer les mêmes quantités par les mêmes lettres. L'inconvénient de cette approche est qu'elle peut semer la confusion en mathématiques. Il est plus prudent d'avoir à l'esprit les subtilités mathématiques que nous allons voir lorsque l'on fait de la physique, même si on utilise les mêmes lettres pour les mêmes quantités mais qui font références à des fonctions différentes.

Le changement de variable consiste à utiliser un point de vue différent plus simple mais équivalent au précédent. Il faut donc considérer une transformation d'un jeu de variables défini dans un ouvert, à un autre jeu de variables défini dans ce même ouvert. La définition la plus adaptée au changement de variables est la suivante.

Règle de la chaîne

Définition III.4.3. Soient \mathcal{U}, \mathcal{V} , deux ouverts de \mathbb{R}^n . On dit que f est un \mathcal{C}^1 difféomorphisme, ou difféomorphisme de classe \mathcal{C}^1 , de \mathcal{U} dans \mathcal{V} , si f est une bijection de \mathcal{U} dans \mathcal{V} et que f et f^{-1} sont \mathcal{C}^1 respectivement sur \mathcal{U} et \mathcal{V} .

Proposition III.4.4. Si f est un \mathcal{C}^1 difféomorphisme de \mathcal{U} dans \mathcal{V} , alors

$$\forall a \in \mathcal{U}, df^{-1}(f(a)) = (df(a))^{-1} \quad (\text{III.64})$$

Démonstration. $f^{-1} \circ f(a) = a = Id(a)$ donc $d(f^{-1} \circ f)(a) = Id = df^{-1}(f(a)) \circ df(a)$ d'où le résultat. ■

L'idée du changement de variables est que si l'on dispose d'un difféomorphisme dont le graphe est $y = \phi(x)$, on peut utiliser indifféremment y ou x comme jeu de variables, y compris s'il y a besoin de dériver les fonctions qui en dépendent. Si l'on note une fonction f de \mathcal{U} dans \mathbb{R}^p , on peut très bien poser la fonction g définie de \mathcal{V} dans \mathbb{R}^p telle que $g(\phi(x)) = g(y) = f(x)$. Pour réconcilier les physiciens et les mathématiciens, on peut noter que la "quantité" z peut s'exprimer ou bien comme une fonction de x ou bien comme une fonction de y : $z = f(x) = g(y)$. Une question très importante en physique est de savoir comment varie la quantité z lorsque x varie, ou lorsque y varie. Lorsqu'on considère z comme une fonction de x , il suffit d'écrire

$$\frac{\partial z}{\partial x_i} = \frac{\partial f}{\partial x_i} \quad (\text{III.65})$$

et de même en substituant x_i par y_i et f par g . Comment obtenir l'une des expressions en fonction de l'autre ? Il suffit pour cela d'utiliser les résultats sur la composition des différentielles. On peut calculer très facilement que

$$\frac{\partial z}{\partial x_i} = \frac{\partial f}{\partial x_i} = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial g}{\partial y_k} \quad (\text{III.66})$$

ou, noté autrement

$$\frac{\partial z}{\partial x_i} = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial z}{\partial y_k} \quad (\text{III.67})$$

Matriciellement, on a multiplié la matrice jacobienne de g par la matrice jacobienne de ϕ . Comme cette équation est valable pour toute quantité z , on peut exprimer cette équation dans l'algèbre des opérateurs et écrire simplement

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial}{\partial y_k} \quad (\text{III.68})$$

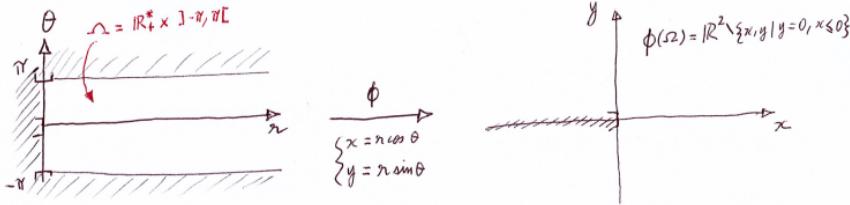


FIGURE III.8 – Représentation du changement de variables en coordonnées polaires.

Cette équation est ce qu'on appelle la "règle de la chaîne". On peut évidemment l'écrire dans l'autre sens

$$\frac{\partial}{\partial y_i} = \sum_{k=1}^n \frac{\partial x_k}{\partial y_i} \frac{\partial}{\partial x_k} \quad (\text{III.69})$$

et là il s'agit de multiplier la jacobienne de la fonction potentiellement considérée par la jacobienne de ϕ^{-1} .

Exemple simple

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^n$, \mathcal{C}^1 . On pose $u = x + y$ et $v = 2x - y$. On pose $z = f(x, y)$. Exprimer $\partial z / \partial x$, puis $\partial z / \partial y$ en fonction de $\partial z / \partial u$ et $\partial z / \partial v$.

Solution : d'abord, on remarque que $\phi : (x, y) \mapsto (x + y, 2x - y)$ est bien un \mathcal{C}^1 difféomorphisme : ϕ est différentiable partout et on peut calculer son inverse, $\phi^{-1}(u, v) = ((u + v)/3, (2u - v)/3)$ qui est aussi différentiable partout. Pour le reste il suffit d'utiliser la règle de la chaîne

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial}{\partial u} + \frac{\partial v}{\partial x} \frac{\partial}{\partial v} = \frac{\partial}{\partial u} + 2 \frac{\partial}{\partial v} \quad (\text{III.70})$$

de même, on obtient

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial u} - \frac{\partial}{\partial v} \quad (\text{III.71})$$

Coordonnées polaires

On pose $x = r \cos \theta$ et $y = r \sin \theta$ et $\phi : (r, \theta) \mapsto (x, y)$. Déterminer les ensembles de départ et d'arrivée tels que ϕ soit un \mathcal{C}^1 difféomorphisme. Écrire la bijection réciproque. Exprimer les dérivées partielles d'un jeu de variables en fonction de l'autre.

Solution. D'abord on constate que ϕ est différentiable sur \mathbb{R}^2 . Son jacobien vaut $\det J_\phi = r$ et s'annule en $r = 0$. Pour que ϕ soit inversible il faut donc enlever la droite $r = 0$ de \mathbb{R}^2 . Ensuite il faut que ϕ soit une bijection. Pour l'instant nous admettons le théorème d'inversion globale III.5.4 qui affirme que si une fonction définie sur un ouvert est différentiable et de jacobien non nul, et injective, sur un ouvert, alors c'est un \mathcal{C}^1 difféomorphisme. Il faut donc trouver l'ensemble de départ tel que ϕ soit injective. On constate que $\phi(r, \theta + 2\pi) = \phi(r, \theta)$. On décide donc de garder un intervalle de largeur 2π pour θ , par exemple $]-\pi, \pi[$. Il reste encore un problème, car on a $\phi(-r, \theta + \pi) = \phi(r, \theta)$. On décide donc de ne garder que les r strictement positifs. L'ensemble de départ sera donc le suivant : $\Omega = \mathbb{R}_+^* \times]-\pi, \pi[$. L'ensemble d'arrivée sera $\mathbb{R}^2 \setminus \{x, y | y = 0, x \leq 0\}$ qui est la demi-droite $(-xO)$. Voir une représentation figure III.8.

Écrivons la bijection réciproque. Tout d'abord on a évidemment $r = \sqrt{x^2 + y^2}$. On voit d'ailleurs ici la non inversibilité de $d\phi$ en $r = 0$ puisque r n'est pas différentiable en $x = y = 0$.

On serait tenté d'écrire $\theta = \text{atan}(y/x)$ mais cette formule présente un inconvénient : elle n'est pas valable quand $x = 0$. Une façon de procéder est de considérer le losange de la figure III.9. Nous y considérons le point $M(x, y)$. Le point B est obtenu en reportant la longueur r sur l'axe (Ox) . Le point A est obtenu de telle sorte que $\overrightarrow{OM} = \overrightarrow{BA}$. Dans ce cas, le quadrilatère $MOBA$ est un losange. De là, il suit que $(\overrightarrow{OB}, \overrightarrow{OA}) = (\overrightarrow{OB}, \overrightarrow{OM})/2 = \theta/2$. D'autre part, on voit sur la figure III.9 que

$$\tan(\overrightarrow{OB}, \overrightarrow{OA}) = \frac{\theta}{2} = \frac{y}{r + x} = \frac{y}{\sqrt{x^2 + y^2} + x} \quad (\text{III.72})$$

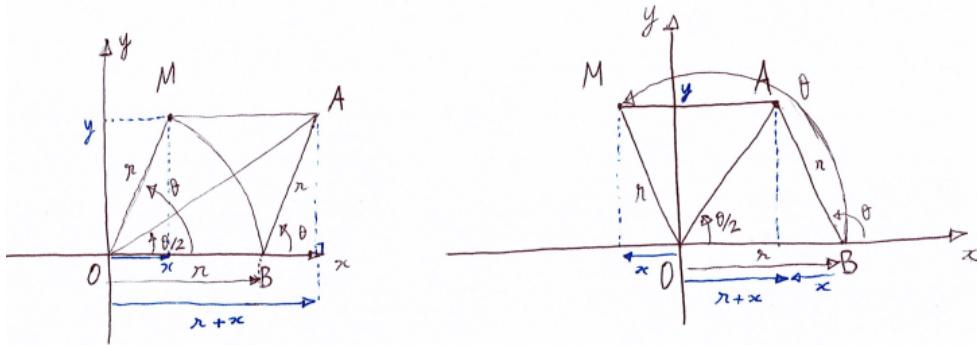
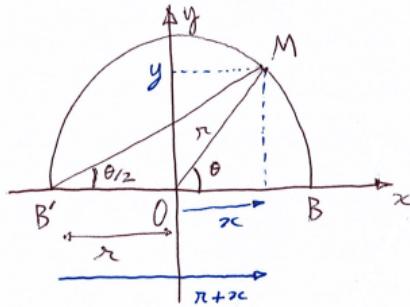

 FIGURE III.9 – À gauche on a représenté un cas où $x > 0$, à droite un cas où $x < 0$.


FIGURE III.10 – Cercle polaire.

d'où

$$\theta = 2 \operatorname{atan} \frac{y}{\sqrt{x^2 + y^2} + x} \quad (\text{III.73})$$

Cette équation est valable sur tout $\phi(\Omega)$. Une autre façon de procéder est la suivante. On constate sur la figure III.10 que l'angle θ est un angle au centre et que $(\overrightarrow{B'B}, \overrightarrow{B'M})$ est un angle inscrit qui coupe la même corde de cercle que θ , d'où $(\overrightarrow{B'B}, \overrightarrow{B'M}) = \theta/2$. Par un raisonnement géométrique simple on en déduit la même formule.

À présent calculons les dérivées partielles. On a

$$\partial_r = \frac{\partial x}{\partial r} \partial_x + \frac{\partial y}{\partial r} \partial_y = \cos \theta \partial_x + \sin \theta \partial_y = \frac{x}{\sqrt{x^2 + y^2}} \partial_x + \frac{y}{\sqrt{x^2 + y^2}} \partial_y \quad (\text{III.74})$$

ou de façon plus compacte et sans perdre en généralité

$$r \partial_r = x \partial_x + y \partial_y \quad (\text{III.75})$$

D'autre part

$$\partial_\theta = \frac{\partial x}{\partial \theta} \partial_x + \frac{\partial y}{\partial \theta} \partial_y = -r \sin \theta \partial_x + r \cos \theta \partial_y = -y \partial_x + x \partial_y \quad (\text{III.76})$$

De là on serait tenté de prendre l'expression de ϕ^{-1} et de différentier. Le calcul devient assez vite immonde. Le plus simple est d'inverser les équations (III.75) et (III.76). C'est un système linéaire à deux inconnues (∂_x et ∂_y) dont l'inversion s'apprend en classe secondaire, on trouve après quelques calculs élémentaires

$$\partial_x = \sin \theta \partial_r + \frac{\cos \theta}{r} \partial_\theta, \quad \partial_y = \cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \quad (\text{III.77})$$

III.4.4 Dérivées d'ordre supérieur

Si une fonction est \mathcal{C}^1 sur un ouvert et que ses dérivées partielles sont non seulement continues mais aussi elles-mêmes \mathcal{C}^1 , on peut se demander ce qu'il advient des dérivées partielles des dérivées

partielles, et pourquoi pas aller plus loin dans l'ordre des dérivées. Nous verrons ensuite comment les fonctions évoluent au voisinage d'un point aux ordres supérieurs.

Théorème de Schwarz

Théorème III.4.5. Soit $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ où \mathcal{U} est un ouvert, de classe \mathcal{C}^2 (elle est \mathcal{C}^1 et ses dérivées partielles sont toutes \mathcal{C}^1 sur \mathcal{U}). Alors

$$\forall a \in \mathcal{U}, \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (a) = \frac{\partial^2 f}{\partial x_i \partial x_j} (a) = \frac{\partial^2 f}{\partial x_j \partial x_i} (a) \quad (\text{III.78})$$

Démonstration. Soit $a = (a_1, a_2, \dots, a_n) \in \mathcal{U}$. Soit $\eta > 0$ tel que

$$\forall (x, y) \in \Omega = [a_i - \eta, a_i + \eta] \times [a_j - \eta, a_j + \eta], a + x e_i + y e_j \in \mathcal{U} \quad (\text{III.79})$$

Pour $(x, y) \in \Omega$, on pose

$$F(x, y) = \int_{x'=a_i}^x \left(\int_{y'=a_j}^y \partial_j (\partial_i f(a_1, \dots, x', \dots, y', \dots, a_n)) dy' \right) dx' \quad (\text{III.80})$$

où x' est à la i -ième position et y' à la j -ième. En intégrant deux fois on calcule immédiatement que

$$F(x, y) = f(\dots, x, \dots, y, \dots) - f(\dots, a_i, \dots, y, \dots) + f(a) - f(\dots, x, \dots, a_j, \dots) \quad (\text{III.81})$$

Mais d'après le théorème de Fubini, on a aussi

$$F(x, y) = \int_{y'=a_j}^y \left(\int_{x'=a_i}^x \partial_j (\partial_i f(\dots, x', \dots, y', \dots)) dx' \right) dy' \quad (\text{III.82})$$

D'après (III.81) on a

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right) (a_i, a_j) = \partial_i \partial_j f(a) \quad (\text{III.83})$$

mais d'après (III.82) on calcule aussi

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right) (a_i, a_j) = \partial_j \partial_i f(a) \quad (\text{III.84})$$

d'où le résultat. ■

Il y a une arnaque : dans cette preuve, nous avons admis le théorème de Fubini. Le théorème de Fubini est difficile à démontrer et s'inscrit dans la théorie de l'intégration. Cela montre que la notion d'intégration est plus puissante que la notion de dérivation.

Remarque : en terme d'opérateurs, cela veut dire que $\partial_i \partial_j = \partial_j \partial_i$.

Développement de Taylor

Théorème III.4.6. (Formule de Taylor-Lagrange)

Soit $f : [a, b] \rightarrow \mathbb{R}$, de classe \mathcal{C}^n , telle que $f^{(n+1)}$ existe sur $]a, b[$. Alors

$$\exists c \in]a, b[/ f(b) = f(a) + \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)^{n+1}} f^{(n+1)}(c) \quad (\text{III.85})$$

Démonstration. Considérons la fonction

$$\varphi : [a, b] \rightarrow \mathbb{R}; x \mapsto f(b) - f(x) - \sum_{k=1}^n \frac{(b-x)^k}{k!} f^{(k)}(x) - A \frac{(b-x)^{n+1}}{(n+1)!} \quad (\text{III.86})$$

On choisit A telle que $\varphi(a) = 0$. φ est continue sur $[a, b]$ et dérivable sur $]a, b[$ et un calcul par télescopage de la somme montre que

$$\forall x \in]a, b[, \varphi'(x) = -\frac{(b-x)^n}{n!} f^{(n+1)}(x) + A \frac{(b-x)^n}{n!} \quad (\text{III.87})$$

On a $\varphi(a) = \varphi(b) = 0$ donc d'après le théorème de Rolle, il existe c dans $]a, b[$ tel que $\varphi'(c) = 0$, ce qui implique $A = f^{(n+1)}(c)$, ce qui achève la preuve. ■

Corrigé III.4.7. Sous les mêmes hypothèses, on a

$$f(b) = f(a) + \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \mathcal{O}((b-a)^{n+1}) \quad (\text{III.88})$$

Nous avions besoin de ces préliminaires sur les développements limités à une variable pour passer à plusieurs variables comme nous allons le voir.

Théorème III.4.8. (Taylor-Lagrange à plusieurs variables)

Soit $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ où \mathcal{U} est ouvert, f est \mathcal{C}^n , $x \in \mathcal{U}$, $\mathbf{h} \in \mathbb{R}^p$ tel que $\forall t \in [-1, 1], x + t\mathbf{h} \in \mathcal{U}$. Alors il existe $\theta \in]0, 1[$ tel que

$$f(x + h) = f(x) + \sum_{k=1}^{n-1} \frac{1}{k!} \partial_{[k]} f(x) h^{[k]} + \frac{1}{n!} \partial_{[n]} f^{[n]}(x + \theta \mathbf{h}) \quad (\text{III.89})$$

où l'on a posé la notation suivante

$$\partial_{[k]} f^{[k]}(x) h^{[k]} = \sum_{i_1+i_2+\dots+i_p=k, i_j \geq 0} \frac{k!}{i_1! i_2! \dots i_p!} \frac{\partial^n f}{\partial x_1^{i_1} \dots \partial x_p^{i_p}}(x) h_1^{i_1} \dots h_p^{i_p} \quad (\text{III.90})$$

Démonstration. Il suffit d'appliquer la formule de Taylor-Lagrange à une variable sur la fonction $F : [0, 1] \rightarrow \mathbb{R}^q, t \mapsto f(x + t\mathbf{h})$. ■

Corrigé III.4.9. Sous les mêmes hypothèses on a

$$f(x + h) = \sum_{k=0}^{n-1} \frac{1}{k!} \partial_{[k]} f(x) h^{[k]} + \mathcal{O}(\|\mathbf{h}\|^n) \quad (\text{III.91})$$

III.4.5 Exercices

Équation d'onde Trouver l'ensemble des fonctions \mathcal{C}^2 sur \mathbb{R}^2 qui vérifient l'équation d'onde

$$\frac{\partial^2 y}{\partial x^2} - \frac{\partial^2 y}{\partial t^2} = 0 \quad (\text{III.92})$$

Indication : utiliser le changement de variables $u = x + t$, $v = x - t$.

Solution On remarque que le changement de variables s'inverse aisément $x = (u + v)/2$, $t = (v - u)/2$ et que tout se dérive bien dans les deux sens. On a donc bien un \mathcal{C}^1 difféomorphisme de \mathbb{R}^2 dans \mathbb{R}^2 , et même un \mathcal{C}^2 difféomorphisme. Calculons comment se transforment les dérivées partielles.

$$\partial_x = (\partial_x u) \partial_u + (\partial_x v) \partial_v = \partial_u + \partial_v \quad (\text{III.93})$$

$$\partial_x^2 = \partial_x(\partial_x) = (\partial_u + \partial_v)(\partial_u + \partial_v) = \partial_u^2 + \partial_v^2 + \partial_u \partial_v + \partial_v \partial_u = \partial_u^2 + \partial_v^2 + 2\partial_u \partial_v \quad (\text{III.94})$$

en vertu du théorème de Schwarz. De même, on calcule que

$$\partial_t^2 = \partial_u^2 + \partial_v^2 - 2\partial_u \partial_v \quad (\text{III.95})$$

De telle sorte que si l'on pose $y(x, t) = z(u, v)$, l'équation d'onde s'écrit

$$\partial_x^2 y - \partial_t^2 y = 4\partial_v \partial_u z = 0 \quad (\text{III.96})$$

La dernière égalité nous indique que la fonction $\partial_u z$ ne dépend que de u puisque sa dérivée partielle par rapport à v est nulle sur \mathbb{R}^2 . \mathbb{R}^2 est convexe, donc d'après le théorème III.3.6, il existe une fonction f d'une variable \mathcal{C}^1 telle que

$$\partial_u z = f(u) \quad (\text{III.97})$$

f est intégrable sur \mathbb{R} . Si l'on note F une primitive (qui est au moins \mathcal{C}^2), alors on peut récrire l'équation précédente ainsi

$$\partial_u(z - F(u)) = 0 \quad (\text{III.98})$$

donc il existe une fonction d'une variable G , \mathcal{C}^2 sur \mathbb{R} telle que $z - F(u) = G(v)$. Réexprimé dans les anciennes variables, on résoud l'équation d'onde ainsi :

$$\exists F, G \in \mathcal{C}^2(\mathbb{R}) / \forall (x, t) \in \mathbb{R}^2, y(x, t) = F(x + t) + G(x - t) \quad (\text{III.99})$$

Comme on l'enseigne en physique ondulatoire, le premier terme correspond à un signal se propageant dans la direction $-x$ et l'autre à un signal se propageant dans la direction $+x$. Ici nous avons pris des unités naturelles en posant que la célérité des ondes était égale à 1. Pour retrouver le terme physique il suffirait de poser $\tau = ct$, ce qui ne correspond à rien d'autre qu'à un changement d'unité de mesure du temps.

Laplacien

Définition III.4.10. Soit une fonction de classe \mathcal{C}^2 sur un ouvert \mathcal{U} de \mathbb{R}^2 . Alors on définit le laplacien de f ainsi :

$$\forall (x, y) \in \mathcal{U}, \Delta f(x, y) = \partial_x^2 f + \partial_y^2 f. \quad (\text{III.100})$$

Autrement dit, on a $\Delta = \partial_x^2 + \partial_y^2$.

1. On reprend le changement en coordonnées polaires évoqué plus haut. Montrer que

$$\Delta = \partial_r^2 + \frac{1}{r} \partial_r + \frac{1}{r^2} \partial_\theta^2 \quad (\text{III.101})$$

2. Trouver les fonctions harmoniques, c'est-à-dire les fonctions qui vérifient $\Delta f = 0$, qui ne dépendent que de $r = \sqrt{x^2 + y^2}$.
3. Généralisation : pour n variables, on pose $\Delta = \sum_{i=1}^n \partial_i^2$. Trouver les fonctions harmoniques ne dépendant que de $r = \sqrt{x_1^2 + \dots + x_n^2}$.

Solution.

1. On a $\partial_x = \cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta$. On calcule

$$\partial_x^2 = \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) \quad (\text{III.102})$$

$$= \cos \theta \partial_r \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) - \frac{\sin \theta}{r} \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) \quad (\text{III.103})$$

$$= \cos^2 \theta \partial_r^2 + 2 \frac{\cos \theta \sin \theta}{r^2} \partial_\theta \partial_r + \frac{\sin^2 \theta}{r} \partial_r^2 + \frac{\sin^2 \theta}{r^2} \partial_\theta^2 - 2 \frac{\cos \theta \sin \theta}{r} \partial_\theta \partial_r \quad (\text{III.104})$$

Avec la même technique on calcule

$$\partial_y^2 = \sin^2 \theta \partial_r^2 - 2 \frac{\sin \theta \cos \theta}{r^2} \partial_\theta \partial_r + \frac{\cos^2 \theta}{r} \partial_r^2 + \frac{\cos^2 \theta}{r^2} \partial_\theta^2 + 2 \frac{\sin \theta \cos \theta}{r} \partial_\theta \partial_r \quad (\text{III.105})$$

En utilisant le théorème de Schwarz et en faisant la somme, on trouve l'égalité annoncée.

2. Si f ne dépend que de r alors on a $\partial_\theta f = 0$, d'où

$$\partial_r^2 f + \frac{1}{r} \partial_r f = 0 \quad (\text{III.106})$$

Comme $\mathbb{R}_+^* \times]-\pi, \pi[$ est convexe, on peut considérer f comme une fonction d'une variable, de telle sorte que $f(x, y) = g(r)$. La fonction g est solution de l'équation différentielle suivante

$$g'' + \frac{1}{r} g' = 0 \quad (\text{III.107})$$

sur $]0, +\infty[$. On sait que la solution générale est $g'(r) = a \exp(-\ln r) = a/r$ où $a \in \mathbb{R}$, soit $g(r) = a \ln r + b$ où $a, b \in \mathbb{R}$.

3. Si f ne dépend que de r , alors elle ne dépend que de $R = r^2$. Cela revient à étudier F avec $f(r) = F(R)$. On a

$$\Delta F = \sum_i \partial_i^2 \left(F \left(\sum_j x_j^2 \right) \right) = \sum_i \partial_i (2x_i F'(R)) = 2(2RF''(R) + nF'(R)) \quad (\text{III.108})$$

Les fonctions harmoniques sont donc solutions de

$$F'' + \frac{n}{2R} F' = 0 \quad (\text{III.109})$$

La solution générale est de la forme $F'(R) = \alpha/R^{n/2}$ où $\alpha \in \mathbb{R}$. Examinons dimension par dimension ce que l'on trouve.

- $n = 1 : F(R) = a\sqrt{R} + b$ ou encore $f(r) = ar + b$ avec $a, b \in \mathbb{R}$.
- $n = 2 : F(R) = \tilde{a} \ln R + b$ ou encore $f(r) = a \ln r + b$ où $a, b \in \mathbb{R}$ (en fait on a $a = 2\tilde{a}$).
- On retrouve bien le cas précédent.
- $n \geq 3 : F(R) = a/R^{\frac{n}{2}-1} + b$ soit $f(r) = a/r^{n-2} + b$ avec $a, b \in \mathbb{R}$. Dans le cas particulier où $n = 3$, on retrouve $f(r) = a/r + b$ qui est le potentiel de Coulomb. En effet, pour un champ électrique à symétrie sphérique, pour tout point en un lieu différent de la charge, l'équation du champ est bien $\Delta\phi = 0$.

III.5 Théorèmes d'inversion

En première approche, on se contente d'énoncer les théorèmes, d'en donner l'heuristique et de les appliquer. Pour ne pas alourdir la première lecture, les démonstrations sont reportées en annexe. Ces théorèmes vont révéler la grande puissance du calcul différentiel, notamment tout ce qui concerne les problèmes inverses, mais permettra aussi de définir les variétés différentielles plongées qui permettent de définir les fonctions implicites, les courbes, les surfaces, les systèmes de coordonnées généralisées. Ici nul besoin d'insister sur la richesse et la multitude des applications physiques qu'on peut en tirer. Nous montrerons quelques exemples.

III.5.1 Inversion locale

C'est le théorème le plus important de cette section. De ce théorème va découler tout le reste.

Théorème III.5.1. (inversion locale).

Soit $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, avec \mathcal{U} ouvert, et f de classe \mathcal{C}^1 . Soit $a \in \mathcal{U}$. Si $df(a)$ est une application linéaire bijective, alors

- Il existe $\mathcal{V}_a \subset \mathcal{U}$ un voisinage ouvert de a ,
- Il existe $\mathcal{W}_b \subset \mathbb{R}^n$ un voisinage ouvert de $b = f(a)$,

tels que f soit un \mathcal{C}^1 difféomorphisme de \mathcal{V}_a dans $\mathcal{W}_b = f(\mathcal{V}_a)$. De plus, en notant $f^{-1} : \mathcal{W}_b \rightarrow \mathcal{V}_a$ la bijection réciproque de f restreinte à \mathcal{W}_b , on a

$$\forall (x, y) \in \mathcal{V}_a \times \mathcal{W}_b / y = f(x), \quad df^{-1}(y) = (df(x))^{-1}. \quad (\text{III.110})$$

En pratique, il faudra vérifier que le jacobien de f en a est non nul. Le théorème semble intuitivement évident : si f est différentiable en a et que sa différentielle est bijective, alors f est localement bijective puisqu'elle se comporte localement comme sa différentielle. Pourtant la preuve est très technique car les majorations de normes nécessaires sont assez subtiles.

Comme on l'a vu avec les difféomorphismes, on peut interpréter ce théorème comme servant de base aux changements de coordonnées. Pour ce faire on a besoin du corollaire suivant.

Corrigé III.5.2. Soit $f : \mathcal{U} \rightarrow \mathbb{R}^n$ où \mathcal{U} est un ouvert de \mathbb{R}^n , telle que $\forall x \in \mathcal{U}$, $df(x)$ est inversible. Alors $f(\mathcal{U})$ est ouvert.

Démonstration. Soit $x \in \mathcal{U}$. D'après le théorème d'inversion locale, il existe deux ouverts $\mathcal{V}_x, \mathcal{W}_y$, ouverts de \mathcal{U} et \mathbb{R}^n , voisinages de x et $y = f(x)$, tels que $f(\mathcal{V}_x) = \mathcal{W}_y$. On en déduit, puisque $\mathcal{U} = \bigcup_{x \in \mathcal{U}} \mathcal{V}_x$,

$$f(\mathcal{U}) = f \left(\bigcup_{x \in \mathcal{U}} \mathcal{V}_x \right) = \bigcup_{x \in \mathcal{U}} \mathcal{W}_y \quad (\text{III.111})$$

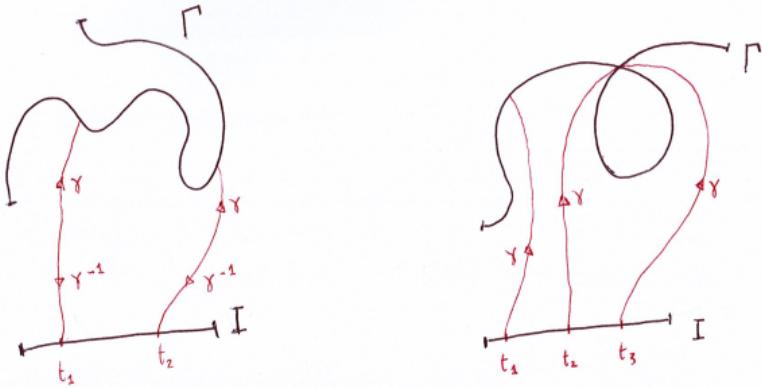


FIGURE III.11 – À gauche, un arc paramétré injectif et régulier, donc bijectif. À droite, un arc paramétré régulier, mais non injectif : on voit que $\gamma(t_2) = \gamma(t_3)$ avec $t_2 \neq t_3$.

qui est ouvert. ■

De là il suit le théorème suivant.

Théorème III.5.3. Soient f_1, \dots, f_n des fonctions réelles C^1 au voisinage de $a \in \mathbb{R}^n$. Les relations $u_i = f(x_1, \dots, x_n)$ définissent un changement de coordonnées sur un voisinage de a si et seulement si $df_i(a)$ sont des formes linéaires indépendants.

C'est une autre façon d'exprimer le théorème d'inversion locale.

III.5.2 Inversion globale

Corrigé III.5.4. (Théorème d'inversion globale).

Soit $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathcal{V} = f(\mathcal{U})$ où \mathcal{U} est ouvert et $f \in C^1$. Si f est injective alors on a l'équivalence entre les deux propositions suivantes.

III.5.4.1. $\forall x \in \mathcal{U}$, $df(x)$ est inversible.

III.5.4.2. $\mathcal{V} = f(\mathcal{U})$ est ouvert, f est globalement inversible et f^{-1} est C^1 .

Démonstration. 2⇒1. f et f^{-1} sont toutes deux C^1 et en tout point on a $f^{-1} \circ f = Id$, soit $\det(df^{-1}(f(x)))\det(df(x)) = 1$, ce qui prouve que $df(x)$ est inversible pour tout x .

1⇒2. Il est évident que f est surjective par construction, par hypothèse elle est donc bijective. Pour prouver que son inverse est C^1 il suffit d'appliquer le théorème d'inversion locale en chaque point pour obtenir l'expression de df^{-1} qui est donc bien définie. ■

Exemple important : coordonnées polaires Nous invitons les lecteurs à relire l'exercice III.4.3 à la lumière des nouveaux théorèmes que nous avions admis à ce moment-là. Remarquons que nous pourrions prendre tout intervalle ouvert de longueur 2π comme ensemble de définition de l'angle pour assurer l'injectivité de la transformation. Alors la demi-droite que l'on excluerait serait l'origine arbitrairement choisie du repère polaire.

Courbe paramétrée Soit $\gamma : I \subset \mathbb{R} \rightarrow \Gamma \subset \mathbb{R}^3$, $t \mapsto \gamma(t) = (x(t), y(t), z(t))$. Si I est ouvert, γ C^1 sur I (x, y, z le sont), et injective, alors γ est un C^1 difféomorphisme de I dans Γ . Remarquons que dire que γ est injective revient à dire que la courbe Γ ne se recoupe pas elle-même (voir figure III.11).

Cet exemple devrait vous rappeler quelque chose

$$\gamma(t) = \left(-\frac{1}{2}gt^2 + v_{0z}t \right) \mathbf{e}_z + v_{0x}t \mathbf{e}_x, \quad t \in \left[0, \frac{2v_{0z}}{g} \right], \quad v_{0x}, v_{0z}, g > 0 \quad (\text{III.112})$$

Il est aisément de montrer que cette courbe paramétrée réalise une bijection entre $[0, 2v_{0z}/g]$ et la courbe d'équation cartésienne (voir figure III.12)

$$z = -\frac{g}{2v_{0x}^2}x^2 + \frac{v_{0z}}{v_{0x}}x, \quad y = 0, \quad x \in [0, 2v_{0x}v_{0z}/g]. \quad (\text{III.113})$$

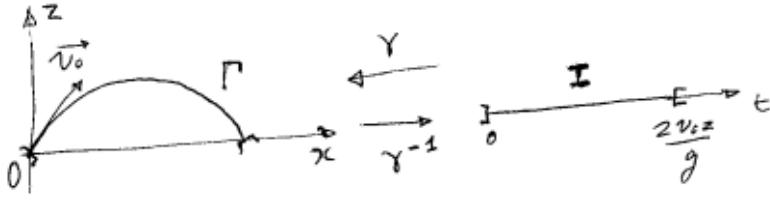


FIGURE III.12 – Graphe de la trajectoire d'une chute libre.

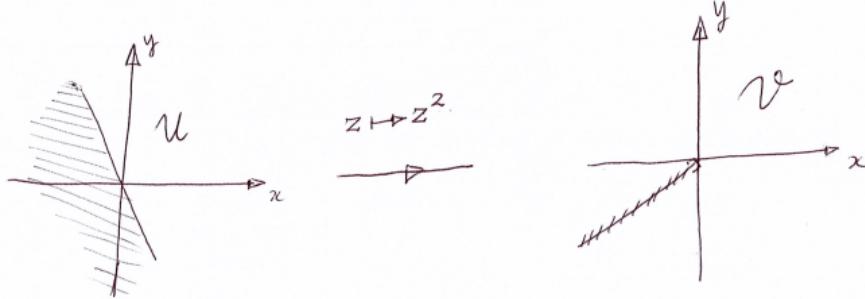


FIGURE III.13 – Exemple d'ensemble de départ et d'ensemble d'arrivée pour que la fonction carré soit un \mathcal{C}^1 difféomorphisme.

Un exemple que nous traiterons abondamment ultérieurement constitue les surfaces régulières.

Exercice Soit $\mathcal{U} = \mathbb{R}^2 \setminus \{(0,0)\}$. On pose $f : \mathcal{U} \rightarrow \mathbb{R}^2, (x, y) \mapsto (x^2 - y^2, 2xy)$.

1. Montrer que f est localement inversible sur tout \mathcal{U} mais pas globalement inversible.
2. Déterminer deux ouverts \mathcal{V} et \mathcal{W} tels que f soit un \mathcal{C}^1 difféomorphisme de \mathcal{V} dans \mathcal{W} .

Solution.

1. f est différentiable sur \mathcal{U} et on calcule aisément $J_f = 4(x^2 + y^2) \neq 0$ sur tout \mathcal{U} . D'après le théorème d'inversion locale, en tout point de \mathcal{U} , il existe un voisinage et un voisinage de l'image de ce point tels que f soit un \mathcal{C}^1 difféomorphisme reliant ces deux ensembles. Autrement dit, f est localement inversible sur tout \mathcal{U} . Cependant, on constate qu'en tout point de \mathcal{U} , on a $f(x, y) = f(-x, -y)$. f n'est donc pas injective, donc n'est pas globalement inversible.
2. En posant $z = x + iy$, on constate que dans \mathbb{C} , f est représentée par la fonction carré. Le résultat de la question précédente montre qu'il est possible de trouver en tout point une détermination unique locale de la racine carrée d'un nombre complexe. Pour trouver le plus "gros" \mathcal{C}^1 difféomorphisme global, il faut trouver un ensemble qui ne contient pas à la fois z et $-z$ afin de rendre la fonction carré injective. Une solution est de prendre tout demi-plan dont la coupure contient l'origine pour \mathcal{U} , et l'image de \mathcal{U} donnera \mathbb{R}^2 privé de la demi-droite dont l'origine est $(0, 0)$ et dont le vecteur directeur sera orthogonal à la coupure du demi-plan antécédent et dirigé vers le complémentaire du demi-plan. Voir figure III.13.

III.5.3 Fonctions implicites

Approche heuristique

Imaginons qu'on cherche à résoudre une équation du type $f(x, y) = 0$ au voisinage de (a, b) sachant que $f(a, b) = 0$. Autrement dit, on cherche à déterminer la ligne de niveau de f au voisinage de (a, b) et passant par (a, b) , représentée comme le graphe d'une fonction de x . Sans souci de rigueur pour le moment, on peut écrire

$$f(x, y) \approx f(a, b) + \partial_x f(a, b)(x - a) + \partial_y f(a, b)(y - b) \quad (\text{III.114})$$

et en déduire que, si $\partial_y f(a, b) \neq 0$,

$$y = \phi(x) \approx b + \frac{\partial_x f(a, b)}{\partial_y f(a, b)}(x - a) \quad (\text{III.115})$$

et on peut même calculer $\phi'(x) = \partial_x f(a, b)/\partial_y f(a, b)$. Le théorème d'inversion locale nous assurera l'existence et l'unicité de ϕ .

Théorème des fonctions implicites

Avant d'énoncer le théorème nous aurons besoin d'une notation.

Définition III.5.5. Soient E_1, \dots, E_n des espaces vectoriels normés et $f : \Omega \subset E_1 \times \dots \times E_n \rightarrow F$, où Ω est ouvert. Soit $a = (a_1, \dots, a_n) \in \Omega$. Pour $i \in \llbracket 1, n \rrbracket$, la fonction $f_i : x \mapsto f(a_1, \dots, x, \dots, a_n)$ où x est à la i -ième position, est définie sur un voisinage de a_i dans E_i . Si f_i est différentiable en a_i , alors on dit que f admet une différentielle partielle d'indice i , et on note

$$\partial_i f(a) = df_i(a_i) \quad (\text{III.116})$$

Évidemment, si f est différentiable en a , on a

$$df(a)\mathbf{h} = \sum_{i=1}^n \partial_i f(a) \mathbf{h}_i \quad (\text{III.117})$$

Si $E_i = \mathbb{R}$, $\partial_i f(a) = \partial f(a)/\partial x_i$.

Théorème III.5.6. (fonctions implicites)

Soient E, F deux espaces vectoriels normés, $\mathcal{U} \subset E$, $\mathcal{V} \subset F$; ouverts, et $f : \mathcal{U} \times \mathcal{V} \rightarrow F$ de classe \mathcal{C}^1 . Soit $(a, b) \in \mathcal{U} \times \mathcal{V}$. Si $\partial_2 f(a, b)$ est une application linéaire bijective, alors

- il existe $\mathcal{U}_a \subset \mathcal{U}$ un voisinage ouvert de a ;
- il existe $\mathcal{V}_b \subset \mathcal{V}$ un voisinage ouvert de b ;
- il existe $\mathcal{W}_c \subset F$ un voisinage ouvert de $c = f(a, b)$;
- il existe une fonction φ , unique, de classe \mathcal{C}^1 , $\varphi : \mathcal{U}_a \times \mathcal{W}_c \rightarrow \mathcal{V}_b$

telles que

$$\forall z \in \mathcal{W}_c, \forall x \in \mathcal{U}_a, f(x, \varphi(x, z)) = z \quad (\text{III.118})$$

L'idée de la preuve est d'appliquer le théorème d'inversion locale à la fonction $(x, y) \mapsto (x, f(x, y))$. Nous reportons cette preuve en annexe.

Souvent, lorsqu'on cherche des fonctions implicites, on fixe la valeur de z et on ne cherche qu'une ligne de niveau de f . Mais la version générale de ce théorème permet aussi de caractériser comment les lignes de niveau varient quand l'image de f varie.

Interprétation graphique

Essayons d'interpréter graphiquement ce théorème dans le cas où $E = F = \mathbb{R}$. Nous avons représenté un cas fictif sur la figure III.14. Pour que le théorème puisse s'appliquer, il faut que f varie le long de e_y . Tant que les lignes de niveau ne sont pas verticales, alors le long de l'axe y , puisqu'on traverse des lignes de niveau, la valeur de $f(x, y)$ varie, donc $\partial_y f \neq 0$. Pour avoir un graphe local de y comme fonction de x , il faut donc que ses lignes de niveau ne soient pas tangentes à e_y , autrement dit il ne faut pas qu'elles soient verticales si on les représente sur un repère cartésien classique. Si tel est le cas, il est toujours possible d'appliquer ce théorème à la variable x et non y , et d'exprimer localement x comme le graphe d'une fonction de y .

Différentielle de la fonction implicite

Supposons les conditions du théorème des fonctions implicites réalisées. Localement, on a $f(x, \varphi(x, z)) = z$ pour tous x et z dans le bon voisinage. Soient $\mathbf{h}_1, \mathbf{h}_3 \in E \times F$. On a

$$d(f(x, \varphi(x, z)))(\mathbf{h}_1, \mathbf{h}_3) = \partial_1 f(x, \varphi(x, z))\mathbf{h}_1 + \partial_2 f(x, \varphi(x, z)) \circ d\varphi(x, z)(\mathbf{h}_1, \mathbf{h}_3) \quad (\text{III.119})$$

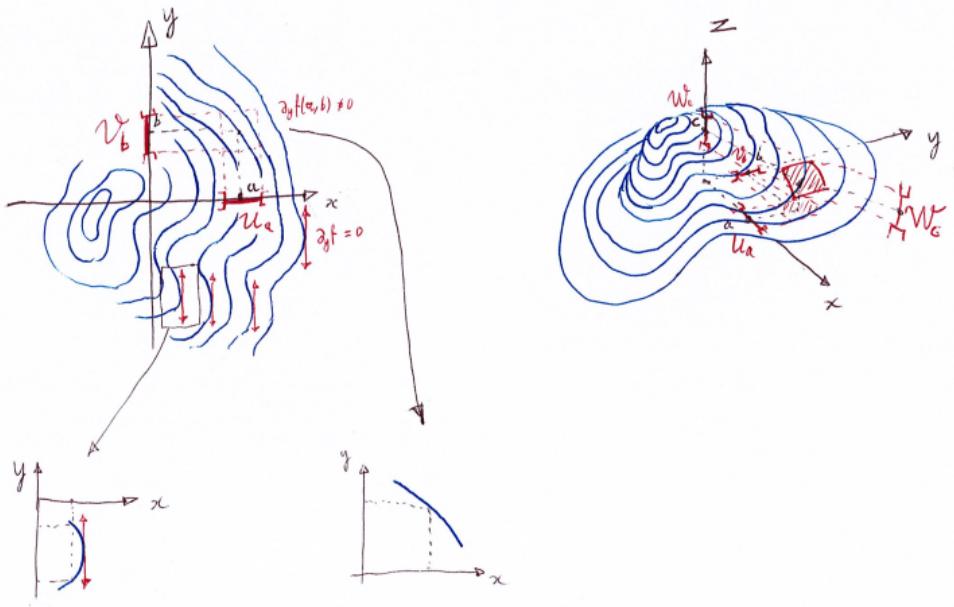


FIGURE III.14 – Nous représentons en haut à gauche les lignes de niveau d'une fonction de deux variables à valeurs dans \mathbb{R} . Nous avons représenté les voisinages correspondant aux notations du théorème en un point où $\partial_2 f(a, b) \neq 0$. Nous avons indiqué quelques cas où $\partial_2 f(a, b) \neq 0$ donc avec une tangente verticale. En dessous, nous avons représenté deux zooms : l'un où $\partial_2 f = 0$, auquel cas la ligne n'est localement pas le graphe d'une fonction (un antécédent peut donner plusieurs images), et l'autre où $\partial_2 f \neq 0$, dans ce cas la ligne est localement le graphe d'une fonction de x . À droite nous avons tenté de faire une représentation en perspective pour montrer à quoi correspond $\mathcal{W}_c = f(\mathcal{U}_a \times \mathcal{V}_b)$.

mais aussi

$$d(f(x, \varphi(x, z)))(\mathbf{h}_1, \mathbf{h}_3) = \mathbf{h}_3 \quad (\text{III.120})$$

Par hypothèse, $\partial_2 f(x, \varphi(x, z))$ est inversible, on peut donc calculer $d\varphi(x, z)$:

$$d\varphi(x, z)(\mathbf{h}_1, \mathbf{h}_3) = (\partial_2 f(x, \varphi(x, z)))^{-1} (\mathbf{h}_3 - \partial_1 f(x, \varphi(x, z))\mathbf{h}_1) \quad (\text{III.121})$$

soit

$$d\varphi(x, z) = (\partial_2 f(x, \varphi(x, z)))^{-1} \circ (Id_F - \partial_1 f(x, \varphi(x, z))) \quad (\text{III.122})$$

Remarquons que si l'on considère que $z = c$ n'est pas une variable, la formule devient simplement

$$d\varphi_c(x) = -(\partial_2 f(x, \varphi_c(x)))^{-1} \circ \partial_1 f(x, \varphi_c(x)) \quad (\text{III.123})$$

en ayant posé $\varphi_c(x) = \varphi(x, c)$.

En règle générale, il est absolument inutile d'apprendre par coeur cette formule. En effet, on peut toujours retomber sur ses pieds en différentiant l'équation $f(x, y) = 0$ comme nous allons le voir dans les exemples qui suivent.

Exercices sur les fonctions implicites

Échauffement : le cercle On souhaite résoudre l'équation $x^2 + y^2 = 1$ en y . On pose $f(x, y) = x^2 + y^2 - 1$ et on cherche les lignes de niveau de f lorsqu'elle prend la valeur 0. On calcule $\partial_y f(x, y) = 2y$, pour pouvoir appliquer le théorème des fonctions implicites il faut donc exclure les points où $y = 0$. Cette équation définit deux fonctions "implicites" (qu'on peut ici déterminer explicitement, en général ce n'est pas le cas) :

$$\varphi_+ :]-1, 1[\rightarrow]0, 1[, x \mapsto \sqrt{1 - x^2} \quad (\text{III.124})$$

$$\varphi_- :]-1, 1[\rightarrow]0, 1[, x \mapsto -\sqrt{1-x^2} \quad (\text{III.125})$$

qui définissent les deux demi-cercles supérieurs et inférieurs. On ne peut pas définir de fonction de x au voisinage des points $(-1, 0)$ et $(1, 0)$, mais on peut le faire dans l'autre sens. Cela donnerait les deux fonctions suivantes

$$\psi_d :]-1, 1[\rightarrow]0, 1[, y \mapsto \sqrt{1-y^2} \quad (\text{III.126})$$

$$\psi_g :]-1, 1[\rightarrow]0, 1[, y \mapsto -\sqrt{1-y^2} \quad (\text{III.127})$$

qui correspondent aux demi-cercles à droite et à gauche de l'axe ($0y$). Calculons les différentielles (ici, simplement les dérivées) des fonctions φ_{\pm} . On pourrait dériver brutalement les fonctions explicitement obtenues ou utiliser la formule obtenue plus haut, mais soyons plus malins. En différentiant l'équation $f(x, \varphi_{\pm}(x)) = 0$ on obtient $xdx + \varphi_{\pm}(x)d\varphi_{\pm} = 0$ soit

$$\varphi'_{\pm}(x) = -\frac{x}{\varphi_{\pm}(x)} = \mp \frac{x}{\sqrt{1-x^2}} \quad (\text{III.128})$$

Ce petit exemple montre la méthode pour obtenir la dérivée de la fonction implicite.

Théorie des perturbations Le théorème des fonctions implicites est à la base de la théorie des perturbations. En effet, la plupart des "vraies" équations de la physique sont insolubles, mais on peut tenter de les résoudre approximativement au voisinage d'un point connu. On la résoud d'abord "à l'ordre 0" en assimilant le voisinage du point au point connu, puis "à l'ordre 1" en utilisant les dérivées premières et en y injectant le résultat à l'ordre 0, et éventuellement aux ordres supérieurs pour repousser le reste à des ordres plus élevés. Une question mathématique intéressante est de savoir si la série converge et si le reste devient nul lorsque l'ordre tend vers l'infini. Ce genre de question occupe les mathématiques des systèmes dynamiques non intégrables qui s'écrivent sous la forme d'équations différentielles insolubles du type $y' = f(y) + \varepsilon g(y)$ dans le cas où l'on connaît la solution de $y' = f(y)$ et que le paramètre ε est petit. La recherche mathématique actuelle s'occupe de déterminer des conditions pour que la théorie des perturbation converge vers une solution exacte (dans la plupart des cas, les séries ne convergent pas). Ce type de problème est notamment étudié en M2 DSG. Pour ce cours, nous allons donner une illustration plus simple qui permet de toucher le problème du doigt.

Soit l'équation (E) : $x^7 + 0,99x - 2,03 = 0$.

1. Montrer que cette équation a une solution unique dans \mathbb{R} .
2. Trouver sans calculatrice une solution à 10^{-1} près, à 10^{-3} près, puis à 10^{-5} près. On pensera à poser $f : \mathbb{R}^3 \rightarrow \mathbb{R}, (x, p, q) \mapsto x^7 + px - q$ et à rechercher les bonnes fonctions implicites.
3. Calculer les dérivées partielles à l'ordre 1 et à l'ordre 2 de la fonction implicite φ qui vérifie $f(\varphi(p, q), p, q) = 0$ au voisinage de $(p, q) = (1, 2)$.

Solution.

1. Soit $g : x \mapsto x^7 + 0,99x - 2,03$. g est continue et dérivable sur \mathbb{R} et $\forall x \in \mathbb{R}, g'(x) = 7x^6 + 0,99 > 0$. De plus on a $\lim_{\pm\infty} g = \pm\infty$ donc g réalise une bijection de \mathbb{R} dans \mathbb{R} donc $g(x) = 0$ admet une solution unique.
2. On pose la fonction f suggérée. Pour $(p, q) = (1, 2)$, une solution évidente (et unique) est $x = 1$. C'est la solution "à l'ordre 0" de (E) , précise à 10^{-1} près.

Pour appliquer le théorème des fonctions implicites, il faut que $\partial_x f(p, q) = 7x^6 + p \neq 0$. Or cela n'est possible que si $p \leq 0$ ce qui n'est pas le cas ici. Donc il existe un voisinage ouvert de $(1, 2)$ et une fonction unique $\varphi(p, q)$ définie sur ce voisinage qui vérifie $f(\varphi(p, q), p, q) = 0$. D'après ce qui précède, on a $\varphi(1, 2) = 1$ et on va essayer de déterminer approximativement $\varphi(0, 99; 2, 03)$ grâce à un développement limité.

À présent posons $(p, q) = (1 + \delta p, 2 + \delta q)$, et notons symboliquement

$$\tilde{x} = x + dx + d^2x + \mathcal{O}(dx^3) = \varphi(1, 2) + d\varphi(1, 2)(\delta p, \delta q) + (1/2)d^2\varphi(1, 2)(\delta p, \delta q) + \mathcal{O}((\delta p, \delta q))^3 \quad (\text{III.129})$$

afin de simplifier les notations. Différention l'équation à résoudre.

$$df(x, p, q) = 7x^6 dx + x dp + pdx - dq = 0 \quad (\text{III.130})$$

En substituant le résultat à l'ordre 0 en x , p , et q , les termes d'ordre 0 se sont implicitement éliminés pour $x = 1$, $p = 1$, $q = 2$ et il ne reste que les termes d'ordre 1, ce qui permet de déduire dx :

$$dx = -\frac{x}{7x^6 + p} dp + \frac{1}{7x^6 + p} dq \quad (\text{III.131})$$

En remplaçant x par 1, p par 1, q par 2, dp par $\delta p = -0.001$, dq par $\delta q = 0.02$, on obtient $\delta x = 0.005$ ce qui donne $\tilde{x} \approx x + \delta x = 1.005$, résultat précis à 10^{-3} .

Pour augmenter la précision il faut aller chercher le résultat à l'ordre 2. Si l'on fait varier l'équation (E) à l'ordre 2 on trouve ceci

$$\begin{aligned} (x + dx + d^2x)^7 + xp + xdp + pdx - q - dq &= x^7 + 7x^6(dx + d^2x) + 21x^5(dx)^2 \\ &\quad + xp + xdp + pdx - q - dq + \mathcal{O}(\|(dp, dq)\|^3) \end{aligned} \quad (\text{III.132})$$

mais on sait que $x^7 + px - q = 0$ et que $7x^6dx + xdp + pdx - dq = 0$, il ne reste donc que les termes d'ordre 2, ce qui donne

$$d^2x = -\frac{3dx^2}{x} - \frac{dxdp}{7x^6} \quad (\text{III.133})$$

et en remplaçant par les bonnes valeurs (notamment $dx = 0.005$) on trouve $d^2x = -6 \cdot 10^{-5}$ ce qui donne un résultat précis à 10^{-5} : $x \approx 1 + 0.05 - 0.00006 = 1.00494$.

3. Pour les dérivées d'ordre 1 il suffit d'identifier les termes de la différentielle de l'équation (III.131). Cela nous donne donc

$$\partial_p \varphi(p, q) = -\frac{x}{7x^6 + p} = -\frac{\varphi(p, q)}{7\varphi(p, q)^6 + p} \quad (\text{III.134})$$

$$\partial_q \varphi(p, q) = \frac{1}{7x^6 + p} = \frac{1}{7\varphi(p, q)^6 + p} \quad (\text{III.135})$$

Pour les dérivées d'ordre 2 il faut développer l'équation (III.133) en y injectant l'expression de dx . Après calculs on trouve ceci

$$d^2x = \frac{x}{7x^6 + p} \left(-\frac{3}{7x^6 + p} + \frac{1}{7x^6} \right) dp^2 - \frac{3}{x(7x^6 + p)} dq^2 + \frac{1}{7x^6 + p} \left(\frac{6}{7x^6 + p} - \frac{1}{7x^6} \right) dpdq \quad (\text{III.136})$$

De là il suffit d'identifier avec

$$d^2x = \frac{1}{2} d^2\varphi(p, q) = \frac{1}{2} (\partial_p^2 \varphi(p, q) dp^2 + \partial_q^2 \varphi(p, q) dq^2 + 2\partial_p \partial_q \varphi(p, q) dpdq) \quad (\text{III.137})$$

que l'on déduit du corollaire III.4.9. Soit

$$\partial_p^2 \varphi(p, q) = 2 \frac{x}{7x^6 + p} \left(-\frac{3}{7x^6 + p} + \frac{1}{7x^6} \right) \quad (\text{III.138})$$

$$\partial_q^2 \varphi(p, q) = -\frac{6}{x(7x^6 + p)} \quad (\text{III.139})$$

$$\partial_p \partial_q \varphi(p, q) = \frac{1}{7x^6 + p} \left(\frac{6}{7x^6 + p} - \frac{1}{7x^6} \right) \quad (\text{III.140})$$

où l'on substitue $\varphi(p, q)$ à x (mentalement seulement pour éviter d'alourdir les notations).

Thermodynamique En thermodynamique des gaz, les trois variables p, v, t sont souvent reliées par une équation d'état de la forme $f(p, v, t) = 0$. Notamment, l'équation des gaz parfait correspond à $f(p, v, t) = pv - t$ en unités naturelles. Si l'on suppose que $p, v, t > 0$ et que toutes les dérivées partielles de f sont non nulles en p_0, v_0, t_0 , alors on peut appliquer le théorème des fonctions implicites trois fois : il existe trois uniques fonctions de deux variables π, ν, τ telles qu'on ait, au voisinage de (p_0, v_0, t_0) :

$$f(\pi(v, t), v, t) = 0 \quad (\text{III.141})$$

$$f(p, \nu(v, t), t) = 0 \quad (\text{III.142})$$

$$f(p, v, \tau(p, v)) = 0 \quad (\text{III.143})$$

Lorsque les physiciens choisissent de considérer l'une de ces trois fonctions, cela signifie qu'ils choisissent de considérer que toutes les autres fonctions ne dépendront que des deux autres variables.

Les physiciens considèrent souvent des fonctions d'état : $u : \mathbb{R}_+^{*3} \rightarrow \mathbb{R}$, $(p, v, t) \mapsto u(p, v, t)$. Cependant, comme les variables (p, v, t) sont reliées par l'équation d'état, les physiciens notent souvent par exemple $u(p, v, t) = u(p, v)$, ou encore $u(p, v, t) = u(p, t)$ ou $u(p, v, t) = u(v, t)$ pour signifier qu'ils choisissent un jeu de variables, alors qu'ils considèrent en fait respectivement les fonctions $\alpha(p, v) = u(p, v, \tau(p, v))$, $\beta(p, t) = u(p, \nu(p, t), t)$, et $\gamma(v, t) = u(\pi(v, t), v, t)$. L'avantage qu'ils en tirent est que les mêmes quantités physiques gardent les mêmes lettres. L'inconvénient est qu'ils se mélangent les pinceaux lorsqu'ils calculent les dérivées partielles. Pour éviter de s'embrouiller, ils ont trouvé une astuce et notent les dérivées partielles ainsi

$$\partial_p \alpha = \left. \frac{\partial u}{\partial p} \right)_v \quad (\text{III.144})$$

$$\partial_p \beta = \left. \frac{\partial u}{\partial p} \right)_t \quad (\text{III.145})$$

et ainsi de suite.

1. Calculer les dérivées partielles de π , ν , τ , dans le cas du gaz parfait et dans le cas général.
2. On considère le gaz parfait. Montrer que

$$\left. \frac{\partial u}{\partial p} \right)_t - \left. \frac{\partial u}{\partial p} \right)_v = -v \left. \frac{\partial u}{\partial t} \right)_p = -\frac{v}{p} \left. \frac{\partial u}{\partial v} \right)_p \quad (\text{III.146})$$

3. Cas général : montrer que

$$\left. \frac{\partial u}{\partial p} \right)_t - \left. \frac{\partial u}{\partial p} \right)_v = \left. \frac{\partial_p f}{\partial_t f} \right. \left. \frac{\partial u}{\partial t} \right)_p = -\left. \frac{\partial_p f}{\partial_v f} \right. \left. \frac{\partial u}{\partial v} \right)_p \quad (\text{III.147})$$

4. Vérifier que

$$\left. \frac{\partial p}{\partial v} \right)_t \left. \frac{\partial v}{\partial t} \right)_p \left. \frac{\partial t}{\partial p} \right)_v = -1 \quad (\text{III.148})$$

Solution.

1. Dans le cas du gaz parfait on a $\pi(v, t) = t/v$ d'où l'on tire aisément $\partial_v \pi = -t/v^2$ et $\partial_t \pi = 1/v$. De même, on trouve $\nu(p, t) = t/p$ d'où $\partial_p \nu = -t/p^2$, $\partial_t \nu = 1/p$, et de même $\tau(p, v) = pv$ d'où $\partial_p \tau = v$ et $\partial_v \tau = p$.

Dans le cas général il suffit d'écrire $df = \partial_p f dp + \partial_v f dv + \partial_t f dt = 0$ et pour obtenir par exemple les dérivées partielles de π , on aurait

$$d(f(\pi(v, t), v, t)) = \partial_p f d\pi(v, t) + \partial_v f dv + \partial_t f dt = 0 \quad (\text{III.149})$$

ce qui nous donne

$$d\pi = -\left. \frac{\partial_v f}{\partial_p f} \right. dv - \left. \frac{\partial_t f}{\partial_p f} \right. dt \quad (\text{III.150})$$

d'où

$$\partial_v \pi = -\left. \frac{\partial_v f}{\partial_p f} \right. \quad \partial_t \pi = -\left. \frac{\partial_t f}{\partial_p f} \right. \quad (\text{III.151})$$

De même

$$\partial_p \nu = -\left. \frac{\partial_p f}{\partial_v f} \right. \quad \partial_t \nu = -\left. \frac{\partial_t f}{\partial_v f} \right. \quad \partial_p \tau = -\left. \frac{\partial_p f}{\partial_t f} \right. \quad \partial_v \tau = -\left. \frac{\partial_v f}{\partial_t f} \right. \quad (\text{III.152})$$

2. Dans le cas du gaz parfait, on a $\pi(v, t) = v/t$, $\nu(p, t) = t/p$, $\tau(p, v) = pv$. On a

$$\left. \frac{\partial u}{\partial p} \right)_t = \partial_p \beta = \partial_p(u(p, t/p, t)) = \partial_1 u - \frac{t}{p^2} \partial_2 u = \partial_1 u - \frac{v}{p} \partial_2 u \quad (\text{III.153})$$

$$\left. \frac{\partial u}{\partial p} \right)_v = \partial_p \alpha = \partial_p(u(p, v, pv)) = \partial_1 u + v \partial_3 u \quad (\text{III.154})$$

$$\left. \frac{\partial u}{\partial t} \right)_p = \partial_t \beta = \partial_t(u(p, t/p, t)) = \frac{1}{p \partial_2 u} + \partial_3 u \quad (\text{III.155})$$

$$\left. \frac{\partial u}{\partial v} \right)_p = \partial_v \alpha = \partial_v(u(p, v, pv)) = \partial_2 u + p \partial_3 u \quad (\text{III.156})$$

En combinant les bonnes équations on en déduit aisément le résultat demandé.

3. Avec la même méthode on a

$$\left. \frac{\partial u}{\partial p} \right)_t = \partial_p \beta = \partial_p(u(p, \nu(p, t), t)) = \partial_1 u + \partial_p \nu \partial_2 u = \partial_1 u - \frac{\partial_p f}{\partial_v f} \partial_2 u \quad (\text{III.157})$$

$$\left. \frac{\partial u}{\partial p} \right)_v = \partial_p \alpha = \partial_p(u(p, v, \tau(p, v))) = \partial_1 u + \partial_p \tau \partial_3 u = \partial_1 u - \frac{\partial_p f}{\partial_t f} \partial_3 u \quad (\text{III.158})$$

$$\left. \frac{\partial u}{\partial t} \right)_p = \partial_t \beta = \partial_t(u(p, \nu(p, t), t)) = -\frac{\partial_t f}{\partial_v f} \partial_2 u + \partial_3 u \quad (\text{III.159})$$

$$\left. \frac{\partial u}{\partial v} \right)_p = \partial_v \alpha = \partial_v(u(p, v, \tau(p, v))) = \partial_2 u + \partial_v \tau \partial_3 u = \partial_2 u - \frac{\partial_v f}{\partial_t f} \partial_3 u \quad (\text{III.160})$$

En combinant les bonnes équations on en déduit aisément le résultat demandé.

4. D'après 1. on a

$$\left. \frac{\partial p}{\partial v} \right)_t \left. \frac{\partial v}{\partial t} \right)_p \left. \frac{\partial t}{\partial p} \right)_v = \partial_v \pi \partial_t \nu \partial_p \tau = \left(-\frac{\partial_v f}{\partial_p f} \right) \left(-\frac{\partial_t f}{\partial_v f} \right) \left(-\frac{\partial_p f}{\partial_t f} \right) = -1 \quad (\text{III.161})$$

III.6 Annexe : preuves techniques

III.6.1 Différentiabilité

III.6.2 Théorèmes d'inversion

Théorème d'inversion locale

Théorème des fonctions implicites

CHAPITRE IV

Courbes

Sommaire

IV.1 Courbes en dimension finie	61
IV.2 Courbes planes	64
IV.3 Courbes gauches	71
IV.4 Exercices	76

Qu'il s'agisse d'une ligne de niveau, d'une trajectoire, du bord d'une surface la notion de courbe est très présente en physique. C'est pourquoi, l'objectif de cette partie est de définir et d'exposer proprement les principales caractéristiques et propriétés des courbes de \mathbb{R}^n . Nous nous intéresserons plus particulièrement au cas plan ($n = 2$ courbes planes) et tridimensionnel ($n = 3$ courbes gauches). On verra comment chacune d'entre elles est localement paramétrée par une abscisse curviligne. Nous explorerons également les notions de point singulier, de courbures, de torsion. Nous verrons également les conditions à fixer sur une courbe pour qu'elle soit une isométrie globale d'une autre courbe.

Nous souhaitons remercier GUEDJ 2015 dont les documents de cours nous ont beaucoup aidé lors de la rédaction de ce chapitre.

IV.1 Courbes en dimension finie

Dans ce qui suit, I, J désigneront deux intervalles ouverts de \mathbb{R} et $\varphi : I \mapsto \mathbb{R}^n$ une application. Afin d'éviter d'éventuel problème symptomatiques, dans ce qui suivra nous travaillerons souvent avec des applications de \mathcal{C}^∞ : les application *lisses*.

Définition IV.1.1 (Application lisse est immersion). *Soit \mathcal{U} un ouvert de \mathbb{R}^n . Une application différentiable $\psi : \mathcal{U} \mapsto \mathbb{R}^k$*

- *est une immersion si et seulement si en tout point de $x \in \mathcal{U}$ $d\psi_x$ est injective.*
- *est lisse si et seulement si $\psi \in \mathcal{C}^\infty$.*

IV.1.1 Généralité et Définition

Définition IV.1.2 (Courbes paramétrée et géométrique). *On appelle courbe paramétrée de \mathbb{R}^n la donnée d'une application lisse $\varphi : I \mapsto \mathbb{R}^n$. L'image $\Gamma = \varphi(I)$ s'appelle courbe géométrique associée.*

On dit qu'un point $\varphi(t_0)$ est régulier si $\frac{d\varphi}{dt} \neq 0$, c'est-à-dire si φ est une immersion au voisinage de t_0 . Un point $\varphi(t_0)$ de Γ tel que $\frac{d\varphi}{dt} = \mathbf{0}$ est appelé un point singulier de Γ .

Il est important de distinguer propriété de paramétrisation des courbes paramétrée et propriété géométrique des courbe géométrique. Par exemple, à une courbe paramétrée correspond une courbe géométrique mais la réciproque n'est pas vraie, on peut définir un *changement admissible de paramétrisation* :

Définition IV.1.3 (Changement admissible de paramétrisation). *Une application lisse $\alpha : J \mapsto I$ est changement admissible de paramétrisation de la courbe géométrique $\Gamma = \varphi(I)$ si est seulement si $t \mapsto \frac{d\alpha}{dt}$ ne s'annule pas sur J .*

Ainsi lors d'un changement de paramètre admissible α , $\frac{d\alpha}{dt}(t_0)$ ne change pas de signe, si $\forall t_0 \in J \frac{d\alpha}{dt}(t_0) > 0$ on dit que α préserve l'orientation, sinon on dit que α change l'orientation.

Exemple IV.1.4.

1. Considérons l'application $\varphi : \mathbb{R} \mapsto \mathbb{R}^2$ définie par $\varphi(t) = (t, t^2)$, la courbe géométrique est la parabole d'équation $\Gamma = \varphi(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$, dont chaque point est régulier, on dit que la paramétrisation est régulière. Posons $\alpha : t \in \mathbb{R} \longrightarrow t^3 \in \mathbb{R}$, ce changement de paramétrisation n'est pas admissible car $\alpha'(0) = 0$. La courbe géométrique $\tilde{\Gamma} = \varphi \circ \alpha(\mathbb{R})$ est un objet géométrique différent de $\Gamma = \varphi(\mathbb{R})$ même si $\tilde{\Gamma} = \Gamma$, en effet $(0, 0) \in \Gamma = \tilde{\Gamma}$ est un point régulier pour le paramétrage φ et singulier pour $\varphi \circ \alpha$.
2. La courbe paramétrée $\varphi : \mathbb{R} \mapsto \mathbb{R}^2$ définie par $\varphi(t) = (t^2, t^3)$ n'est pas régulière au point $(0, 0)$ (ce point est un point de rebroussement de première espèce, on caractérisera plus tard ce genre de point).
3. La courbe paramétrée $\varphi : \mathbb{R} \mapsto \mathbb{R}^2$ définie par $\varphi(t) = (\cos t, \sin t)$ définit un paramétrisation du cercle unité $\mathbb{U} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, parcouru en sens trigonométrique. $\alpha : t \in \mathbb{R} \longrightarrow -t \in \mathbb{R}$ est un changement de paramétrisation admissible et $\varphi \circ \alpha$ parcourt le cercle unité dans le sens des aiguilles d'une montre.

Définition IV.1.5 (Droite). Soit $M, N \in \mathbb{R}^n$ ($M \neq N$) et $u \in \mathbb{R}^n \setminus 0$. On appelle droite passant par M et de direction u (qu'on note $\mathcal{D}_{M,u}$) la courbe géométrique associé au paramétrage $\varphi : t \in \mathbb{R} \mapsto M + tu$.

On appelle droite passant par M et N (qu'on notera (MN)) la droite passant par M et de direction $MN = N - M$.

Théorème (et définition) IV.1.6 (Tangente). Soit une courbe paramétrée $\varphi : I \mapsto \mathbb{R}^n$ et $\Gamma = \varphi(I)$ la courbe géométrique associé. Pour tout point régulier $P = \varphi(t_0) \in \Gamma$ de Γ ($\varphi'(t_0) \neq 0$), alors $\forall t, t' \in I$ tel que $\varphi(t) \neq \varphi(t')$ la droite $\mathcal{D}_{t,t'} = (\varphi(t)\varphi(t'))$ converge vers une droite appelée *tangente* à Γ au point P . On vérifie aisément qu'elle a pour équation paramétrique

$$t \in \mathbb{R} \mapsto (t - t_0) \frac{d^n \varphi}{dt^n}(t_0) + \varphi(t_0)$$

où n est le premier ordre non nuls des dérivée de φ en t_0

Définition IV.1.7 (Courbe fermée). Soit $a, b \in I$ avec $a < b$ et $\varphi : I \longrightarrow \mathbb{R}^n$ tel que $\varphi(a) = \varphi(b)$, on dit alors que $\Gamma = \varphi([a, b])$ est une courbe fermée.

Définition IV.1.8 (Courbe fermée simple). Une courbe fermée $\varphi([a, b])$ est dites simple si et seulement si elle n'a pas d'autre intersection, c'est-à-dire $\forall t \neq s \in]a, b[\implies \varphi(s) \neq \varphi(t)$.

Remarque IV.1.9. Il est important de remarquer que les courbes géométrique des courbes fermé sont défini sur des intervalles dont une borne au moins est fermée.

IV.1.2 Abscisse Curviligne

On a vu dans IV.1.7 qu'une courbe fermée est l'image d'un segment fermé *particulier* contenu dans I d'une courbe paramétrée φ . Ce qui nous amène à la définition de l'arc paramétré.

Théorème (et définition) IV.1.10 (Arc paramétré et longueur d'un arc). Soit $\varphi : I \mapsto \mathbb{R}^n$ une courbe paramétrée et soient $a, b \in I$ donc $[a, b] \subset I$, alors l'image du segment fermé $A = \varphi(J) \subset \Gamma = \varphi(I)$ est un *arc fermée* de la courbe géométrique Γ . La *longueur* de A est :

$$\ell(A) \hat{=} \int_a^b \|\varphi'(t)\| dt,$$

elle est indépendante de la paramétrisation admissible choisie.

Preuve. Déjà l'application $t \mapsto \|\varphi'(t)\|$ est continue (φ est lisse), sur $[a, b]$ (qui n'est pas infini), ce qui justifie l'existence de $\ell(A)$. Considérons $\alpha : J \longrightarrow I$ une paramétrisation admissible qu'on considérera sans perte de généralité conserver l'orientation (sinon on travaille avec $\beta(t) = \alpha(-t)$),

on note $\tilde{a} = \varphi^{-1}(a)$ et $\tilde{b} = \varphi^{-1}(b)$. Notons $\psi = \varphi \circ \alpha$, on a $\psi([\tilde{a}, \tilde{b}]) = \varphi([a, b])$. Alors on a par la formule du changement de variable :

$$\begin{aligned}\ell(\phi([a, b])) &= \int_a^b \|\varphi'(t)\| dt = \int_{\tilde{a}}^{\tilde{b}} \|\varphi'(\alpha(t))\| \alpha'(t) dt = \int_{\tilde{a}}^{\tilde{b}} \|\varphi'(\alpha(t)) \alpha'(t)\| dt \\ &= \int_{\tilde{a}}^{\tilde{b}} \left\| \frac{d\varphi \circ \alpha}{dt}(t) \right\| dt = \int_{\tilde{a}}^{\tilde{b}} \|\psi'(t)\| dt = \ell(\varphi \circ \alpha([\tilde{a}, \tilde{b}]))\end{aligned}$$

■

Théorème (et définition) IV.1.11 (Abscisse curviligne). Soit $\varphi : I \mapsto \mathbb{R}^n$ une courbe paramétrée régulière. On pose $L = \sup(\{\ell(\varphi([a, b])) \mid a < b \in I\}) \in \mathbb{R}_+ \cup +\infty$, alors $\forall t_0 \in I$ l'application,

$$\begin{array}{rcl}s & : & I \rightarrow [-L, L] \\ t & \mapsto & \int_{t_0}^t \|\varphi'(u)\| du\end{array}$$

est lisse et strictement croissante de dérivée $\forall t \in I \ s'(t) = \|\varphi'(t)\| \neq 0$ (paramétrisation admissible), elle est bijective et sa bijection est lisse, on l'appelle *paramétrisation par l'abscisse curviligne*. On note $J = s(I)$, alors la paramétrisation $\psi = \varphi \circ s^{-1}$ vérifie :

$$\|\psi'(t)\| = \|\varphi(s^{-1}(t)) \frac{ds^{-1}}{dt}(t)\| = \|\varphi(s^{-1}(t)) \frac{1}{s'(s^{-1}(t))}\| = \|\varphi(s^{-1}(t)) \frac{1}{\|\varphi'(s^{-1}(t))\|}\| = 1$$

on dis qu'elle est à *paramétrisation unitaire*.

Proposition IV.1.12. Soient $\varphi : I \longrightarrow \mathbb{R}^n$ et $\psi = \varphi \circ \alpha : \alpha(I) \longrightarrow \mathbb{R}^n$ deux paramétrisation unitaire de la même courbe géométrique Γ , alors $\exists t_0$ tel que $\alpha(t) = \pm t + t_0$.

Exercice 10. — Distance entre deux droites.

On considère deux droites $\mathcal{D}_{M_1, \mathbf{u}_1}$ et $\mathcal{D}_{M_2, \mathbf{u}_2}$ (non parallèles), montrer que la distance entre ces droites est

$$d(\mathcal{D}_{M_1, \mathbf{u}_1}, \mathcal{D}_{M_2, \mathbf{u}_2}) = \frac{|\det(\mathbf{M}_1 \mathbf{M}_2, \mathbf{u}_1, \mathbf{u}_2)|}{\|\mathbf{u}_1 \wedge \mathbf{u}_2\|}$$

et que les points de plus courtes approche sont :

$$A_1 = M_1 + \frac{(\mathbf{u}_1 \wedge \mathbf{u}_2) \cdot (\mathbf{M}_1 \mathbf{M}_2 \wedge \mathbf{u}_2)}{\|\mathbf{u}_1 \wedge \mathbf{u}_2\|^2} \mathbf{u}_1 \quad \text{et} \quad A_2 = M_2 + \frac{(\mathbf{u}_2 \wedge \mathbf{u}_1) \cdot (\mathbf{M}_2 \mathbf{M}_1 \wedge \mathbf{u}_1)}{\|\mathbf{u}_1 \wedge \mathbf{u}_2\|^2} \mathbf{u}_2$$

Afin de décomposer les propriétés des courbes géométriques nous pouvons les étudier localement. Toute courbe paramétrée autour d'un point régulier est "plongé" à l'ordre deux dans son abscisse curviligne dans un plan qu'on appellera *osculateur*, à l'ordre 3 la courbe est "plongé" dans un sous-espace vectoriel de dimension 3. Les propriétés géométriques local à l'ordre 2 ou 3 dans l'abscisses curviligne d'une courbe de \mathbb{R}^d ($d > 3$), peuvent donc être déduite des propriétés géométriques des courbes planes ou gauche (dimension 3).

IV.1.3 Etude locale et Branche infinie

Le but ici est tout d'abord d'étudier localement la courbe paramétrée φ . Considérons un point régulier de cette courbe supposons qu'on ait

$$\varphi(t_0 + \epsilon) = \varphi(t_0) + P_n(\epsilon) \frac{d\varphi}{dt}(t_0) + \frac{\epsilon^n}{n!} \frac{d^n \varphi}{ds^n} + \epsilon^{n+1} r(\epsilon)$$

Où n est le premier ordre supérieur ou égal à 2 non proportionnel à $\frac{d\varphi}{dt}(t_0)$ et P_n un polynôme de degré $n - 1$. Pour n pair la courbe se situe localement du même côté de la tangente. Pour n impair la courbe "traverse sa tangente".

Exemple IV.1.13. Considérer $g(x) = 1 - \exp(1/x^2)$ prolongée par continuité en 0. On montrer que g est infiniment dérivable en 0 de dérivée nulle. Donc l'indice n de la formule précédente n'existe pas.

Exercice 11. — Paramétrisation du cercle.
Trouver les paramétrisations unitaires du cercle de rayon R .

Exercice 12. — Hélice.

L'hélice circulaire de \mathbb{R}^3 de rayon a et de pas b est la courbe paramétrée $\varphi : t \rightarrow (a \cos t, b \sin t, bt/2\pi)$, trouver ses paramétrisations unitaires.

Exercice 13. — Calculs d'abscisse curviligne.
Soit $\varphi : t \in]0, +\infty[\rightarrow (t^3, t^3) \in \mathbb{R}^2$. Montrer que la longueur d'arc comptée à partir du point $(0, 0)$ est :

$$\ell(t) = \frac{(4 + 9t^2)^{3/2} - 8}{27}.$$

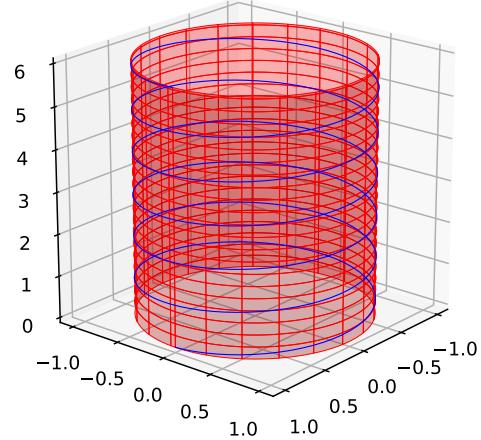


FIGURE IV.1 – Représentation d'une hélice (bleu) de rayon $a = 1$ et de pas $b = 1$ à l'intérieur d'un cylindre (rouge).

On considère maintenant un point singulier de cette courbe paramétrée φ . On suppose alors que,

$$\varphi(t_0 + \epsilon) = \varphi(t_0) + P_q(\epsilon) \frac{d^p \varphi}{ds^p} + \frac{\epsilon^q}{q!} \frac{d^q \varphi}{ds^q} + \epsilon^{q+1} r(\epsilon)$$

Où p est le premier ordre non nul et $q > p$ le premier ordre supérieur à p pour lequel $\frac{\epsilon^q}{q!} \frac{d^q \varphi}{ds^q}$ n'est pas colinéaire à $\frac{d^p \varphi}{ds^p}$

Définition IV.1.14 (Classification point singulier).

- Si p est impair et q pair on parle de point de concavité.
- Si p est impair et q impair on parle de point d'inflexion.
- Si p est pair et q impair on parle de point de rebroussement de première espèce.
- Si p est pair et q pair on parle de point de rebroussement de seconde espèce.

Exemple IV.1.15. Etudiez la nature du point $(0, 0)$ de la courbe géométrique $\Gamma = \{(x, y) \in \mathbb{R}^2 \mid x^a = y^b\}$ pour différentes valeurs de a et b .

Définition IV.1.16 (Branche infinie, Direction et asymptote). Soit $\varphi : I \mapsto \mathbb{R}^n$ une courbe paramétrée qui diverge vers l'infini en t_0 (t_0 peut être égal à $\pm\infty$). Alors on dit que φ présente une branche infinie en t_0 .

Cette branche infinie présente une direction asymptotique si $\forall P \in \mathbb{R}^d$ les droites $(P\varphi(t))$ tendent vers une droite limite L lorsque $t \rightarrow t_0$. De plus si il existe une droite D tel que la distance entre D et $\varphi(t)$ tend vers t_0 on dit que $\Gamma = \varphi(I)$ admet D pour asymptote.

IV.2 Courbes planes

IV.2.1 Courbure

Soit $s \in I \mapsto \varphi(s) \in \mathbb{R}^2$ une courbe régulière paramétrée *plane* par son abscisse curviligne, alors $s \mapsto \varphi'(s)$ est une courbe paramétrée sur le cercle unité, celle-ci possède un vecteur accélération $s \mapsto \frac{d^2 \varphi}{ds^2}$. Au signe près la courbure est la norme de l'accélération. Puisque $\frac{d\varphi}{ds}(s) \perp \frac{d^2 \varphi}{ds^2}(s)$, si la courbure n'est pas nulle $(\frac{d\varphi}{ds}(s), \frac{d^2 \varphi}{ds^2})$ forme une base. On conviendra que la courbure est strictement positive en s si $(\frac{d\varphi}{ds}(s), \frac{d^2 \varphi}{ds^2})$ forme une base directe, strictement négative sinon.

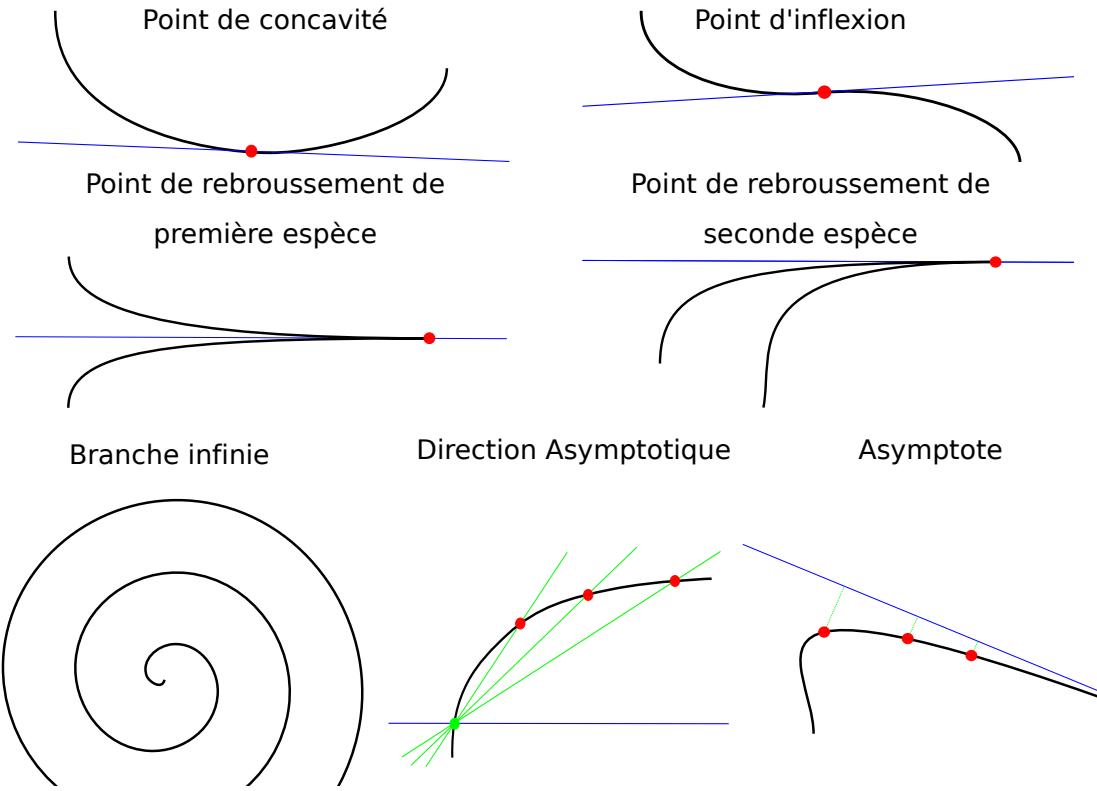


FIGURE IV.2 – Représentation des différents type de point sur une courbe géométrique

Définition IV.2.1 (Courbure). *La courbure d'une courbe paramétrique régulière paramétrée par son abscisse curviligne est*

$$\kappa(s) \hat{=} \epsilon(s) \left\| \frac{d^2\varphi}{ds^2}(s) \right\|,$$

où le signe $\epsilon(s) \in \{-1, +1\}$ est positif si $(\frac{d\varphi}{ds}(s), \frac{d^2\varphi}{ds^2}(s))$ forme une base directe (déterminant positif dans la base canonique), négatif sinon.

Proposition IV.2.2. La courbure d'une courbe paramétrique régulière paramétrée par son abscisse curviligne est :

$$\kappa(s) \hat{=} \det \left(\frac{d\varphi}{ds}(s), \frac{d^2\varphi}{ds^2}(s) \right)$$

Remarque IV.2.3. Le long d'une courbe régulière, la courbure ne dépend pas de la paramétrisation de la courbe seulement de son *orientation*. En effet il n'existe que deux orientation pour la paramétrisations unitaires, dans la seconde paramétrisation $\varphi'(s) \leftrightarrow -\varphi'(s)$ et donc la courbure change de signe pour l'autre orientation de la courbe régulière. Le signe de la courbure dépend également du choix de la convention concernant l'orientation du plan.

Intuitivement, imaginez que $s \rightarrow \varphi(s)$ soit la position d'une voiture. Alors la courbure est positive sur les portions de trajectoire pour lesquelles le conducteur tourne vers la gauche et négative quand il tourne vers la droite, on comprend donc que si le conducteur fais demi-tour et parcours la même trajectoire dans l'autre sens (changement d'orientation de la paramétrisation unitaire), alors la courbure à changer de signe.

On appelle *point d'inflexion* un point de courbure nulle.

Comme l'abscisse curviligne est souvent difficile à calculer il est utiles de pouvoir calculer la courbure dans une paramétrisation quelconque

Proposition IV.2.4. Soit $t \mapsto \psi(t) \in \mathbb{R}^2$ une paramétrisation quelconque d'une courbe géométrique orienté Γ , alors la courbure $\kappa(t)$ au point $\psi(t)$ est donnée par

$$\kappa(t) \hat{=} \frac{\det\left(\frac{d\varphi}{dt}(t), \frac{d^2\varphi}{dt^2}(t)\right)}{\left\|\frac{d\varphi}{dt}(t)\right\|^3}$$

Preuve. La courbe étant régulière, $\forall t_0 \in I$ il existe un unique paramétrage unitaire orienté correctement $s : t \mapsto s(t)$ s'annulant en t_0 . Etant admissible le paramétrage unitaire possède une fonction inverse lisse $t : s \mapsto t(s)$, la courbe paramétrée $\varphi = \psi \circ t$ est un paramétrage unitaire orienté correctement de la courbe géométrique orientée Γ . Alors on a,

$$\mathbf{T}(s) = \frac{d\varphi}{ds} = \psi' \circ t(s) \frac{dt}{ds} = \frac{\psi' \circ t(s)}{\|\psi' \circ t(s)\|}$$

Ou l'on a utilisé le fait que les paramétrisations t et s étaient inverse l'une de l'autre $t \circ s(x) = x$ et que par définition $\frac{ds}{dt} = \|\psi'(t)\|$ et donc $\frac{dt}{ds} = \frac{1}{\|\psi' \circ t(s)\|}$. On pose alors, $\mathbf{N}(s)$ le vecteur normal unitaire directe à $\mathbf{T}(s)$. On a immédiatement,

$$\frac{d\mathbf{T}}{ds}(s) = \frac{d^2\varphi}{ds^2} = \kappa(s) \mathbf{N}(s)$$

en effet $\forall s \|\mathbf{T}(s)\| = 1$, donc $\frac{d\mathbf{T}}{ds}(s) \perp \mathbf{T}(s)$ et donc $\frac{d\mathbf{T}}{ds}(s) \parallel N$ la définition de la courbure fait le reste puisque pour tout s , $(\mathbf{T}(s), \mathbf{N}(s))$ forme une base directe par construction. Pour exprimer la courbure κ dérivons la dernière expression de $\mathbf{T}(s)$ par rapport à s , on obtient,

$$\begin{aligned} \frac{d\mathbf{T}}{ds}(s) &= \frac{d}{ds} \left(\frac{\psi' \circ t(s)}{\|\psi' \circ t(s)\|} \right) = \frac{\psi'' \circ t(s)}{\|\psi' \circ t(s)\|} \frac{dt}{ds}(s) + \frac{d}{ds} \left(\frac{1}{\|\psi' \circ t(s)\|} \right) \psi' \circ t(s) \\ &= \frac{\psi'' \circ t(s)}{\|\psi' \circ t(s)\|^2} + \frac{d}{ds} \left(\frac{1}{\|\psi' \circ t(s)\|} \right) \psi' \circ t(s) \end{aligned}$$

En utilisant $\kappa(s) \hat{=} \det\left(\frac{d\varphi}{ds}(s), \frac{d^2\varphi}{ds^2}(s)\right)$, et la linéarité du déterminant on obtient immédiatement la relation demandée. ■

Théorème (et définition) IV.2.5 (Base de Frenet). Soit Γ une courbe géométrique régulière orientée et φ un paramétrage unitaire orienté de la même façon que Γ , on appelle *base othonormale directe plane de Frenet* la base constituée de $\mathbf{T}(s) = \frac{d\varphi}{ds}$ et de $\mathbf{N}(s)$ le vecteur normal unitaire tel que (\mathbf{T}, \mathbf{N}) soit direct. Pour tout $\varphi(s)$ de courbure non nul, on pose $\kappa(s)$ la courbure orientée alors,

$$\begin{cases} \mathbf{T}(s) = \frac{d\varphi}{ds} \\ \mathbf{N}(s) = \frac{1}{\kappa(s)} \frac{d^2\varphi}{ds^2} \end{cases}$$

La base othonormale directe plane de Frenet vérifie en tout point,

$$\begin{cases} \frac{d\mathbf{T}}{ds}(s) = \kappa(s) \mathbf{N}(s) \\ \frac{d\mathbf{N}}{ds}(s) = -\kappa(s) \mathbf{T}(s) \end{cases}$$

On exprime la base de Frenet à partir d'un paramétrage quelconque $t \mapsto \gamma(t)$.

$$\begin{cases} \mathbf{T}(s) = \frac{1}{\|\dot{\gamma}(t)\|} \dot{\gamma}(t) \\ \mathbf{N}(s) = \frac{1}{\|\dot{\gamma}(t)\|} \frac{\dot{\gamma}(t) \wedge (\ddot{\gamma}(t) \wedge \dot{\gamma}(t))}{\det(\dot{\gamma}(t), \ddot{\gamma}(t))} \end{cases}$$

La base de Frenet en un point définit un repère direct du plan

On laisse la démonstration à titre d'exercice.

Proposition IV.2.6. Toute courbe géométrique régulière est localement le graphe d'une fonction $y = f(x)$ dans son repère de Frenet.

Preuve. Pour tout point P , on pose φ la paramétrisation unitaire centré sur $P = \varphi(0)$, on note $\mathbf{T}_0, \mathbf{N}_0$ la base de Frenet en P on pose alors,

$$\begin{cases} x(s) = (\varphi(s) - P) \cdot \mathbf{T}_0 \\ y(s) = (\varphi(s) - P) \cdot \mathbf{N}_0 \end{cases}$$

Par construction x et y sont lisses et donc C^1 . De plus $\frac{dx}{ds}(0) = 1 > 0$ et donc le théorème d'inversion locale implique qu'il existe $\epsilon > 0$ et $a, b \in \mathbb{R}_{+\star}$ tel que x soit un C^1 -difféomorphisme de $]-\epsilon, \epsilon[$ sur $]a, b[$, on note S défini sur $]a, b[$ et à valeurs dans $]-\epsilon, \epsilon[$ la bijection de x , S vérifie, $\frac{dS}{dx} = \frac{1}{T(S(x)) \cdot T(0)}$ qui est positive sur $]a, b[$.

La courbe géométrique définie par $x \mapsto P + (x, y(S(x)))$ sur $]a, b[$ est donc le graphe d'une fonction contenant P dans le repère de Frenet en P qui est localement inclus dans Γ

Exercice 14. — Courbures....

Montrer que les courbures, du cercles $t \mapsto (R \cos t, R \sin t)$, d'une ellipse $\mapsto (a \cos t, b \sin t)$.

$$\kappa(t) = \frac{1}{R} \text{ Cercle} \quad \kappa(t) = \frac{ab}{\sqrt{a^2 \cos^2 t + b^2 \sin^2 t}} \text{ Ellipse}$$

Proposition IV.2.7. Les seules courbes paramétrées planes dont la courbure est constante sont les portions de cercles ou les portions de droites.

Preuve. Soit Γ une courbe géométrique régulière et φ un paramétrage unitaire. Si la courbure est nulle, on a $\varphi''(s) = 0$, en intégrant il vient immédiatement que Γ est une portion de droite. Sinon l'idée de la démonstration est de trouver les centres du cercle. Notons $\mathbf{T}(s) = \varphi'(s)$ et $\mathbf{N}(s) = \mathbf{T}'(s)/\kappa$. Or $\mathbf{N}(s)$ est unitaire et donc $\mathbf{N}'(s) \perp \mathbf{N}(s)$ et donc $\mathbf{N}'(s) = \lambda(s)\mathbf{T}(s)$. En dérivant $\mathbf{T}(s) \cdot \mathbf{N}(s) = 0$ on obtient,

$$\mathbf{N}'(s) = -\kappa \mathbf{T}(s)$$

Posons alors $M(s) = \varphi(s) + \frac{1}{\kappa} \mathbf{N}(s)$, on vérifie par construction que $M'(s) = 0$ et donc $\forall s$, on a $M(s) = M_0$ et $\|M_0 - \varphi(s)\| = \frac{1}{|\kappa|}$ la courbe est donc une portion de cercle de centre M_0 et de rayon $1/|\kappa|$. ■

On a, lors de la précédentes preuve défini un point $M(s) = \varphi(s) + \frac{1}{\kappa} \mathbf{N}(s)$ s'avérant pour une courbure constante être le centre du cercle de rayon $1/\kappa$ dans lequel la courbe étais incluse. Quand la courbe a une courbure variable, le cercle de centre $M(s) = \varphi(s) + \frac{1}{\kappa} \mathbf{N}(s)$ et de rayon $\kappa(s)$ définira le *cercle osculateur* de la courbe géométrique.

Théorème (et définition) IV.2.8. Soit Γ une courbe géométrique, et $M \in \Gamma$ un point régulier de courbure non nulle, alors il existe un unique cercle ayant un contact d'ordre 3 avec Γ . Alors le centre de ce cercle est $C = M + \frac{1}{\kappa} \mathbf{N}(0)$ et son rayon est $\frac{1}{|\kappa|}$.

Preuve. Posons $\varphi : t \mapsto \varphi(t) = (x(t), y(t))$ est une paramétrisation unitaire vérifiant $\varphi(0) = M$.

On cherche $C \in \mathbb{R}^2$ et $R > 0$ tel que si l'on pose $\tilde{\mathbf{C}}(t) = C + R \frac{\mathbf{C}\varphi(t)}{\|\mathbf{C}\varphi(t)\|}$ la distance $\|\tilde{\mathbf{C}}(t)\varphi(t)\| = t^3 O(1)$. On a $\tilde{\mathbf{C}}(t)\varphi(t) = (1 - \frac{R}{\|\mathbf{C}\varphi(t)\|})\mathbf{C}\varphi(t)$ et par conséquent $\|\mathbf{C}\varphi(t)\|^2 - R^2 = \pm 2Rt^3 \epsilon(t) = t^3 O(1)$ équivaut à $\|\tilde{\mathbf{C}}(t)\varphi(t)\| = t^3 O(1)$.

$$\begin{cases} \tilde{\mathbf{C}}(t) = C + R \frac{\mathbf{C}\varphi(t)}{\|\mathbf{C}\varphi(t)\|} \\ \|\tilde{\mathbf{C}}(t)\varphi(t)\| = t^3 O(1) \end{cases} \iff \begin{cases} \tilde{\mathbf{C}}(t)\varphi(t) = (1 - \frac{R}{\|\mathbf{C}\varphi(t)\|})\mathbf{C}\varphi(t) \\ \|\mathbf{C}\varphi(t)\|^2 - R^2 = 2Rt^3 \epsilon(t) = t^3 O(1) \end{cases}$$

Or φ est lisse et régulier de courbure non nuls on à donc :

$$\begin{aligned} \varphi(t) &= M + t \frac{d\varphi}{dt}(0) + \frac{t^2}{2} \frac{d^2\varphi}{dt^2}(0) + t^3 O(t) \\ \implies \|\mathbf{C}\varphi(t)\|^2 &= \|CM\|^2 + t \mathbf{C}M \cdot \frac{d\varphi}{dt}(0) + t^2 \left(\mathbf{C}M \cdot \frac{d^2\varphi}{dt^2}(0) + \|\frac{d\varphi}{dt}(0)\|^2 \right) + t^3 O(t) \end{aligned}$$

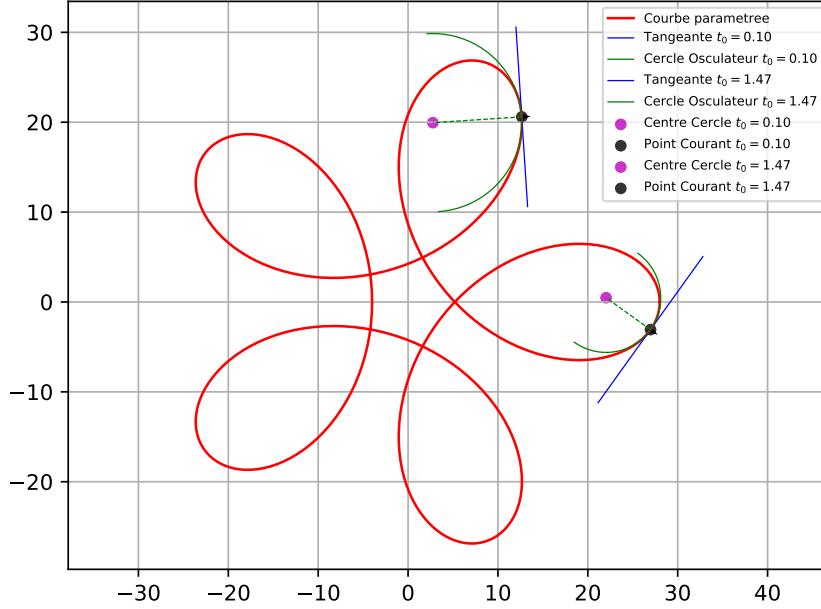


FIGURE IV.3 – Représentation de la courbe paramétrée $t \mapsto a(q-1)u(t) + ap u((1-q)t)$ pour $a = 4 p = 3 q = 5$. Deux cercles osculateur sont représenté en vert et les tangente en bleu.

Et donc le contacte d'ordre 3 de $\tilde{C}(t) = C + R \frac{C\varphi(t)}{\|C\varphi(t)\|}$, est équivalent à,

$$\Leftrightarrow \begin{cases} \tilde{C}(t)\varphi(t) = (1 - \frac{R}{\|C\varphi(t)\|})C\varphi(t) \\ \|CM\|^2 + tCM \cdot \frac{d\varphi}{dt}(0) + t^2 \left(CM \cdot \frac{d^2\varphi}{dt^2}(0) + \left\| \frac{d\varphi}{dt}(0) \right\|^2 \right) - R^2 = t^3 O(1) \end{cases}$$

Et encore,

$$\Leftrightarrow \begin{cases} \tilde{C}(t)\varphi(t) = (1 - \frac{R}{\|C\varphi(t)\|})C\varphi(t) \\ \|CM\|^2 = R^2 \\ CM \cdot \frac{d\varphi}{dt}(0) = 0 \\ CM \cdot \frac{d^2\varphi}{dt^2}(0) + \left\| \frac{d\varphi}{dt}(0) \right\|^2 = 0 \end{cases}$$

Puisque φ est une paramétrisation unitaire on a,

$$\Leftrightarrow \begin{cases} \tilde{C}(t)\varphi(t) = (1 - \frac{R}{\|C\varphi(t)\|})C\varphi(t) \\ \|CM\| = R \\ CM \cdot T(0) = 0 \\ \kappa CM \cdot N(0) + 1 = 0 \end{cases}$$

Ce qui aboutit à

$$\Leftrightarrow \begin{cases} \tilde{C}(t)\varphi(t) = (1 - \frac{R}{\|C\varphi(t)\|})C\varphi(t) \\ C = M + \frac{1}{\kappa}N(0) \end{cases}$$

Ce qui donne donc le résultat escompté (On a travailler avec des équivalence et donc la réciproques et l'unicité sont justes). ■

IV.2.2 Convexité

On rappelle qu'une partie (notamment du plan) est convexe si et seulement si pour tout couple de point le segment passant par ces points est inclus dans la partie. On va étendre cette définition aux courbes du plan.

Définition IV.2.9 (Convexité Locale). *On dit qu'une courbe géométrique $\Gamma = \varphi(I)$ est localement convexe si et seulement si elle borde un ensemble convexe au voisinage de chacun de ses points, c'est-à-dire $\forall t \in I, \exists t_1, t_2 \in I, \exists A$ partie du convexe du plan tel que l'arc ouvert $\varphi([t_1, t_2]) = Fr(A) = A - \overset{\circ}{A}$*

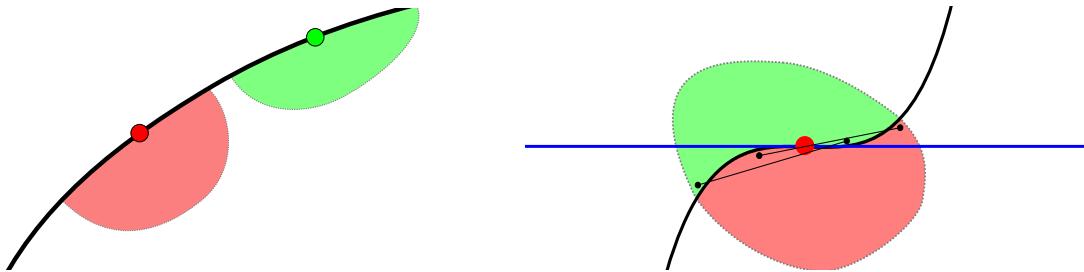


FIGURE IV.4 – Représentation des différents types de possibilité concernant la convexité locale. À gauche, une courbe localement convexe, est représentée autour d'un point en rouge l'ensemble convexe bordant localement le point considéré. À droite, une courbe avec un point d'inflexion pour lequel il n'existe aucun ensemble convexe qui le borde, on comprend à l'aide des segments en noir et du fait que la courbe traverse sa tangente en ce point (en bleu), que quelle que soit la manière dont on cherche un ensemble qui borde avec la condition de l'arc ouvert, cet ensemble ne sera pas convexe.

Proposition IV.2.10. Soit une courbe géométrique régulière ne comportant pas de segment de droite, graphe d'une fonction (c'est-à-dire $x \mapsto (x, f(x))$ avec f lisse sur $]a, b[$) et localement convexe alors $\frac{d^2f}{dx^2}$ garde un signe constant.

Preuve. Sens directe : considérons Γ une courbe géométrique graphe d'une fonction qui est localement convexe. On note $\varphi : x \in]a, b[\mapsto (x, f(x))$ la courbe paramétrée associée.

La courbe étant convexe on a $\forall x_0 \in I, \exists x_1 < x_0 < x_2 \in I, \exists A$ partie convexe du plan $\varphi([x_1, x_2]) = Fr(A)$.

Une partie convexe auxquels on soustrait un demi-plan reste convexe, ainsi $\tilde{A} = \{(x, y) | x_1 < x < x_2\} \cap A$ est convexe. On vérifie que $Fr(\tilde{A}) = Fr(A) = \varphi([x_1, x_2])$.

On note $C_+ = \{(x, y) \in]x_1, x_2[\times \mathbb{R} | y > f(x)\}$ et $C_- = \{(x, y) \in]x_1, x_2[\times \mathbb{R} | y < f(x)\}$ les ouverts constitués des points d'abscisses x comprises entre x_1 et x_2 et d'ordonnée strictement supérieures ou strictement inférieures à $f(x)$.

Si \tilde{A} est convexe alors $\overset{\circ}{A}$ l'est également. On vérifie que $Fr(\tilde{A}) = Fr(A) = \varphi([x_1, x_2])$. $\forall M \in \tilde{A}$, $y(M) \neq f(x(M))$ sinon $M \in \varphi([x_1, x_2]) = Fr(\tilde{A}) = \tilde{A} - \overset{\circ}{A}$ donc $\tilde{A} = (C_+) \cap \overset{\circ}{A} \cup (C_-) \cap \overset{\circ}{A} = A_+ \cap A_-$ une partition en deux ouverts d'un convexe donc connexe par arc. Ce genre de partition (Cf : Partition en ouvert d'un connexe par arc) ne peut contenir deux ouvert non vide. On supposera par symétrie que $A_- = \emptyset$. Et donc $\tilde{A} = A_+$. On remarque que $C_+ \cup Fr(A)$ est donc un ensemble convexe bordant Γ .

On a donc par convexité de $C_+ \cup Fr(A)$, $\forall t_1 \leq t_2 \in]x_1, x_2[\quad \forall a \in]0, 1[$ on a $f(at_1 + (1-a)t_2) \leq af(t_1) + (1-a)f(t_2)$ et donc $\forall t_1 \leq t_2 \leq t_3 \in]x_1, x_2[$ quelques manipulations algébriques permettent d'arriver rapidement à :

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

Ce qui en faisant tendre x_2 vers x_1 , par caractère C^∞ de f implique que f' est croissante et donc $\frac{d^2f}{dx^2} \geq 0$ sur $]x_1, x_2[$. Si on avait eu $A_+ = \emptyset$ alors on aurait montré de la même manière que $\frac{d^2f}{dx^2} \leq 0$ sur $]x_1, x_2[$. Dans tous les cas on obtient la propriété $P_1 : \forall x_o \in I, \exists x_1 < x_o < x_2 \in I$ tel que $\frac{d^2f}{dx^2}$ ne change pas de signe sur $]x_1, x_2[$.

Donc $\forall \epsilon > 0 \forall x \in [a + \epsilon, b\epsilon]$, il existe $a_x < x < b_x \in I$ tel que $\frac{d^2f}{dx^2}$ ne change pas de signe sur $]a_x, b_x[$ et donc $\bigcup_{x \in [a + \epsilon, b - \epsilon]}]a_x, b_x[$ un recouvrement d'ouvert du segment $[a + \epsilon, b - \epsilon]$, puisque les segments fermés sont compacts on peut donc en extraire un sous-recouvrement fini (Bolzano-Weierstrass). Notons $]a_i, b_i|_{i=1 \dots n}$ un sous-recouvrement minimal¹ sur laquelle P_1 est vrai.

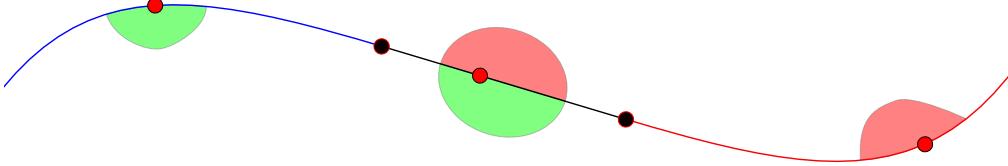


FIGURE IV.5 – Cas où la courbe géométrique comporte un morceau de droite, cette courbe est localement convexe, néanmoins sa courbure change de signe. Concernant les points sur le morceau de droite les ensembles convexes peuvent être choisis d'un côté ou de l'autre de la courbe.

Si $n > 1$.

Montrons tout d'abord qu'il y a forcément une intersection non vide entre les ouverts. En effet, si $\forall j \in [|2 \dots n|],]a_1, b_1[\cap a_j, b_j[= \emptyset$.

Alors $a_1 \in]a, a + \epsilon \cup b + \epsilon, b[$ en effet dans le cas contraire $a_1 \in [a + \epsilon, b - \epsilon]$ et donc il existe $j \in [|2 \dots n|]$ tel que $a_1 \in]a_j, b_j[$ et donc $]a_1, b_1[\cap a_j, b_j[\neq \emptyset$ contradiction donc $a_1 \in]a, a + \epsilon \cup b + \epsilon, b[$. De même $b_1 \in]a, a + \epsilon \cup b + \epsilon, b[$.

- Soit $]a_1, b_1[$ recouvre $[a + \epsilon, b\epsilon]$ et donc cela entraîne en contradiction avec :
- soit le caractère minimal du sous-recouvrement
- soit le fait que $n > 1$.
- Soit $]a_1, b_1[\cap [a + \epsilon, b\epsilon] = \emptyset$ et cela entraîne en contradiction avec le caractère minimal du recouvrement, puisque on peut retirer $]a_1, b_1[$ de la liste sans changer le caractère de recouvrement de $[a + \epsilon, b\epsilon]$.

Donc on à forcément l'existence de $j \in [|2 \dots n|],]a_1, b_1[\cap a_j, b_j[\neq \emptyset$, on supposera quitte à réordonner les segments ouverts qu'il s'agit de $j = 2$. On note $]a_1, b_1[\cap a_2, b_2[=]\tilde{a}, \tilde{b}[$ et donc $\frac{d^2f}{dx^2}$ ne change pas de signe sur $]\tilde{a}, \tilde{b}[$ et donc $\frac{d^2f}{dx^2}$ est du même signe sur $]a_1, b_1[\cup a_2, b_2[=]\bar{a}, \bar{b}[$. On a donc créé un sous-recouvrement fini minimal de cardinal $n - 1$, puisque on peut remplacer $]a_1, b_1[$ et $]a_2, b_2[$ dans la liste des ouverts par l'ouvert $]\bar{a}, \bar{b}[$. En répétant ce processus on arrive en un nombre fini d'étape à $n = 1$

Si $n = 1$ alors puisque $]a_1, b_1[$ est un recouvrement et donc la propriété P_1 est vrai sur $[a + \epsilon, b\epsilon]$ qui est contenue dans $]a_1, b_1[$.

On est donc capable d'assurer la propriété P_1 sur $[a + \epsilon, b - \epsilon]$ quelque soit $\epsilon > 0$ et donc on peut l'assurer sur $]a, b[$ ■

Proposition IV.2.11. Si courbe géométrique $\Gamma = \psi(I)$ régulière ne comportant pas de segment de droite est localement convexe alors sa courbure garde un signe constant.

Preuve. La courbe étant régulière $\forall t \in I, \psi(t) = P, P \in \Gamma$ on peut la reparamétriser par son paramétrage unitaire φ centré en P ($\varphi(0) = P$) et définir ainsi la base de Frenet associée à ce paramétrage. La proposition IV.2.6 implique que pour tout point $P \in \Gamma$, l'arc est localement descriptible par le graphe d'une fonction, $\exists a, b \in \mathbb{R}_{+*}$ et $\exists f :]-a, b[\rightarrow \mathbb{R}$ tel que l'arc $y = f(x)$ du repère de Frenet en P contienne P ($f(0) = 0$) et soit inclus dans Γ .

Posons $\theta(s)$ l'angle entre $\mathbf{T}(0)$ et $\mathbf{T}(s)$ (c'est le même entre $\mathbf{N}(0)$ et $\mathbf{N}(s)$). On a $\frac{d\mathbf{T}}{ds} = \kappa(s)\mathbf{N}(s)$, de plus $\sin(\theta(s)) = N(0) \cdot \mathbf{T}(s)$, en dérivant cette relation on obtient $\frac{d\theta}{ds} = \kappa(s)$. La courbe géométrique associée à $x \mapsto (x, f(x))$ est régulière, ne comporte pas de segment de droite, est le graphe d'une fonction et localement convexe et donc la proposition IV.2.10 implique que $\frac{d^2f}{dx^2}$ garde un signe constant. Or $\frac{df}{dx} = \tan \theta(s(x))$ et donc $\frac{d^2f}{dx^2} = (1 + \tan^2 \theta(s(x))) \frac{d\theta}{ds} \frac{dx}{ds}$ puisque sur $]-a, b[$, $\frac{dx}{ds}$ garde le même signe cela implique que sur $]-a, b[\frac{d\theta}{ds} = \kappa(s(x))$ ne change pas de signe. En

1. que pour tout retrait d'un ouvert, la famille ne recouvre plus $[a + \epsilon, b - \epsilon]$

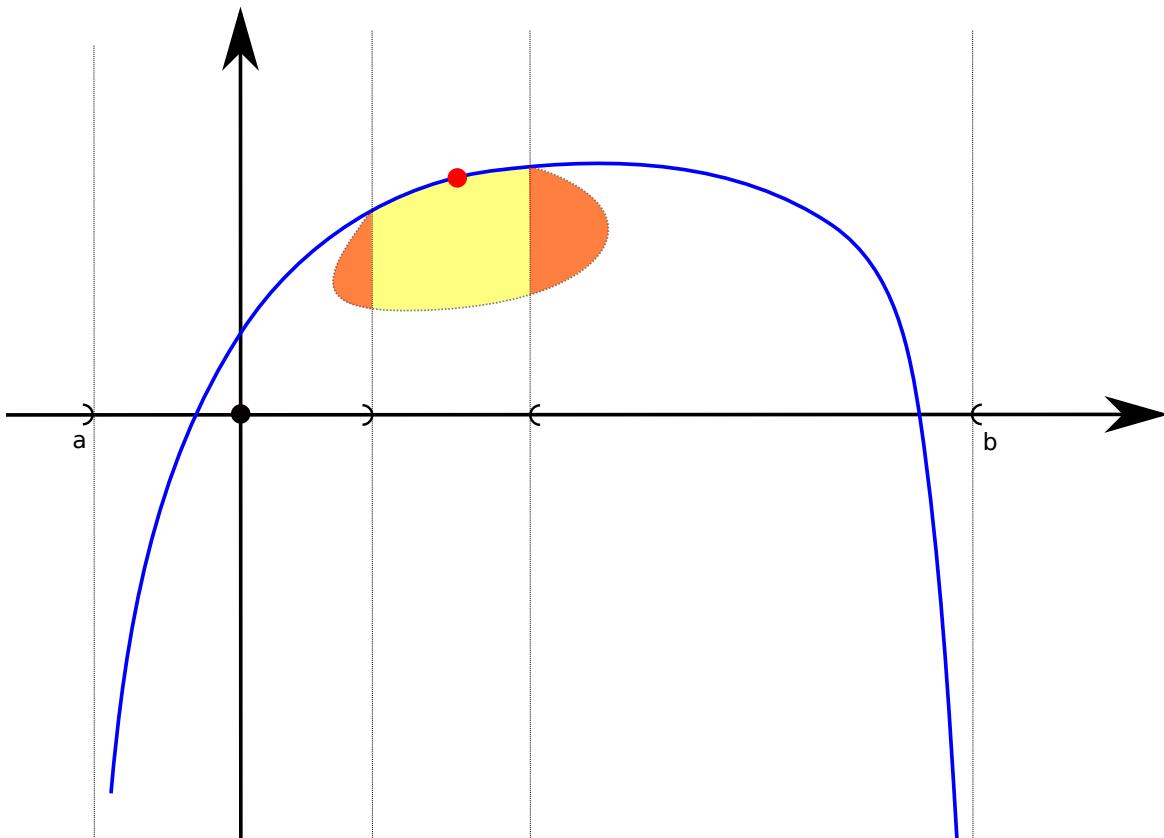


FIGURE IV.6 – Construction de \tilde{A} dans la première partie de la démonstration de la proposition IV.2.10.

notant $A_P = \varphi(s(-a))$ et $B_P = \varphi(s(b))$ on construit ainsi $\forall P \in \Gamma$ un arc ouvert de Γ contenant P sur lequel la courbure ne change pas de signe.

La fin de la démonstration est basé sur le même principe que la démonstration de IV.2.10. Un arc fermé sur lequel on construit un recouvrement d'ouvert sur laquelle la propriété de constance du signe de la courbure est valable. En utilisant la compacité de l'arc fermé on extrait une sous famille finie minimale que l'ont réduit à une famille à un élément en regroupant les arc ouverts entre eux. ■

IV.3 Courbes gauches

Les courbes gauches sont les courbes de \mathbb{R}^3 , ce nom provient du fait que ces courbes quittent à priori leurs *plan osculateur*. Ainsi contrairement à ce qui se passe pour les courbes planes, on ne peut orienter N uniquement avec la donnée de T . On est donc amené à donner à ce vecteur une définition plus *intrinsèque* à l'arc lui-même. Le même genre de considération impliquera qu'on ne pourra donner réellement d'orientation (et donc de signe) à la courbure. On pourra par contre définir un "scalaire orienté", il s'agira de la *torsion*. Elle mesurera la façon dont la courbe gauche quitte son plan osculateur.

IV.3.1 Base de Frenet, Courbure, Plan osculateur, Cercle osculateur et des courbes gauches

Commençons par généraliser dans \mathbb{R}^3 les outils construit pour les courbes planes. On considère une courbe géométrique régulière Γ de paramétrisation unitaire φ .

Définition IV.3.1 (Courbure et Base de Frenet). *On appelle vecteur tangent unitaire orienté avec*

la paramétrisation unitaire φ , le vecteur

$$\mathbf{T}(s) = \frac{d\varphi}{ds}(s)$$

On définit la courbure de notre courbe gauche par

$$\kappa(s) = \left\| \frac{d\mathbf{T}}{ds}(s) \right\|$$

C'est donc un scalaire appartenant à \mathbb{R}_+ . Le vecteur normal et binormal ne sont définis que pour les points de courbure non nuls, on a alors

$$\mathbf{N}(s) = \frac{1}{\kappa(s)} \frac{d\mathbf{T}}{ds}(s) \text{ Normal} \quad \mathbf{B}(s) = \mathbf{T}(s) \wedge \mathbf{N}(s) \text{ Binormal}$$

La base $\mathbf{T}, \mathbf{N}, \mathbf{B}$ est donc une base orthonormale directe appelée Base de Frenet.

Proposition IV.3.2. On exprime la base de Frenet à partir d'un paramétrage quelconque $t \mapsto \gamma(t)$. Alors si le point est régulier alors on a :

$$\begin{cases} \kappa(s) = \frac{\|\dot{\gamma}(t) \wedge \ddot{\gamma}(t)\|}{\|\dot{\gamma}\|^3} \\ \mathbf{T}(s) = \frac{1}{\|\dot{\gamma}(t)\|} \dot{\gamma}(t) \\ \mathbf{N}(s) = \frac{1}{\|\dot{\gamma}(t)\|} \frac{\dot{\gamma}(t) \wedge (\ddot{\gamma}(t) \wedge \dot{\gamma}(t))}{\|\dot{\gamma}(t) \wedge \ddot{\gamma}(t)\|} \end{cases}$$

Il s'agit juste de dérivé correctement.

La notions de tangente n'a pas à être redéfinie, car on l'a définie en toute généralité dans la définition IV.1.6. La tangente d'un point régulier est une droite qui a un ordre de contact d'ordre 2. Le plan osculateur est un plan qui aura un ordre de contacte d'ordre 3.

Théorème (et définition) IV.3.3 (Plan Osculateur). On considère un point régulier P d'abscisse curviligne 0 (reparamétrage) de la courbe géométrique Γ et de courbure non nulle. Alors il existe un unique plan ayant un point de contact d'ordre 3 avec la courbe géométrique en P ce plan est $P + \text{vect}(\mathbf{T}(0), \mathbf{N}(0))$, on l'appelle *Plan Osculateur*.

Preuve. La condition de contact en P est tout simplement que le plan s'écrit $P + \text{vect}(\mathbf{t}, \mathbf{n})$ (ou \mathbf{t}, \mathbf{n} est une base orthonormée de ce plan.). Le but de cette démonstration est de montrer que si le point de contact est d'ordre 3 alors $\text{vect}(\mathbf{t}, \mathbf{n}) = \text{vect}(\mathbf{T}, \mathbf{N})$. La distance entre le plan $P + \text{vect}(\mathbf{t}, \mathbf{n})$ et $\varphi(s)$ est la norme de $\det(P\varphi(s), \mathbf{t}, \mathbf{n})$. Le contacte est d'ordre 3 si il existe $s \mapsto e(s)$ borné dans un voisinage de 0 tel que $\det(P\varphi(s), \mathbf{n}, \mathbf{d}) = s^3 e(s)$. On développe l'arc autour de $s = 0$. On a $P\varphi(s) = \mathbf{T}_0 s + \kappa(0) \frac{s^2}{2} \mathbf{N}(0) + s^3 \mathbf{v}(s)$ (où \mathbf{v} est borné au voisinage de 0). Réinjecté dans l'expression du déterminant, on obtient quasi immédiatement (puisque $\kappa \neq 0$) $\det(\mathbf{T}, \mathbf{t}, \mathbf{n}) = \det(\mathbf{N}, \mathbf{t}, \mathbf{n}) = 0$ et donc $\text{vect}(\mathbf{t}, \mathbf{n}) = \text{vect}(\mathbf{T}, \mathbf{N})$. ■

Si le plan à un point de contact d'ordre 3 avec la courbe géométrique en P cela veut dire qu'il existe des courbes de ce plan ayant un contact d'ordre 3. Il existe en particulier un cercle de ce plan qui à point de contact d'ordre 3.

Théorème (et définition) IV.3.4 (Cercle Osculateur). On considère un point régulier P d'abscisse curviligne 0 (reparamétrage) de la courbe géométrique Γ et de courbure non nulle. Alors il existe un unique cercle ayant un point de contacte d'ordre 3 avec la courbe géométrique en P . Ce cercle est le cercle du plan osculateur de centre $P + \frac{1}{\kappa(0)} \mathbf{N}(0)$ et de rayon $\frac{1}{\kappa(0)}$.

La preuve un peu technique est laissée à titre d'exercice.

IV.3.2 Torsion

La façon dont le plan osculateur varie est déterminé par la variation le long de la courbe des vecteurs \mathbf{T} et \mathbf{N} . Or la variation de \mathbf{T} n'est fonction que de la courbure et de \mathbf{N} . La variation de \mathbf{N} possède une composante le long de \mathbf{T} (qui est connue $\frac{d\mathbf{N}}{ds} \cdot \mathbf{T} = -\kappa$) et une composante le long du vecteur binormal \mathbf{B} , cette composante s'appelle la *torsion*.

Définition IV.3.5 (Torsion). *La torsion de la courbe Γ au point $\varphi(s)$ est*

$$\tau(s) = \mathbf{B}(s) \cdot \frac{d\mathbf{N}}{ds}(s)$$

On déduit aisément les variation du vecteur binormal que

$$\mathbf{N}(s) \cdot \frac{d\mathbf{B}}{ds}(s) = -\tau(s) \quad \mathbf{T}(s) \cdot \frac{d\mathbf{B}}{ds}(s) = 0$$

On en déduit les formules

Proposition IV.3.6. La base de Frenet a pour variation

$$\begin{cases} \frac{d\mathbf{T}}{ds} = [\tau\mathbf{T} + \kappa\mathbf{B}] \wedge \mathbf{T} \\ \frac{d\mathbf{N}}{ds} = [\tau\mathbf{T} + \kappa\mathbf{B}] \wedge \mathbf{N} \\ \frac{d\mathbf{B}}{ds} = [\tau\mathbf{T} + \kappa\mathbf{B}] \wedge \mathbf{B} \end{cases}$$

Remarque IV.3.7. La formule précédente montre que la base de Frenet "précessse" autour d'un vecteur du plan $\varphi(s) + \text{vect}(\mathbf{T}, \mathbf{B})$, la courbure caractérise la variation de la tangente quand la torsion caractérise la variation du plan osculateur. En effet le plan osculateur étant normal à \mathbf{B} , sa variation n'est caractérisé que par la variation de \mathbf{B} qui est elle-même entièrement caractérisée par la torsion.

Proposition IV.3.8. Une courbe gauche dont la courbure ne s'annule pas, est incluse dans un plan si et seulement si sa torsion est identiquement nulle.

Preuve. Dans le sens direct, si une courbe gauche est incluse dans un plan alors si sa courbure ne s'annule pas, \mathbf{T} et \mathbf{N} sont parrallèles à la direction de ce plan. Et donc le vecteur binormal est un vecteur normal unitaire continu et donc constant ce qui implique que la torsion est nulle.

Réciiproquement si la torsion est nulle et la courbure non nulle, alors la binormale est constante. On pose alors $f(s) = \mathbf{B} \cdot \varphi$. On montre aisément que cette fonction est constante ce qui est équivalent à dire que φ est plongé dans un plan normal à \mathbf{B} . ■

Proposition IV.3.9. Soit $\Gamma = \gamma(I)$ une courbe géométrique donnée par un paramétrisation quelconque. La torsion est donnée par

$$\tau(t) = \frac{\dot{\gamma}(t) \cdot (\ddot{\gamma}(t) \wedge \dddot{\gamma}(t))}{\|\dot{\gamma}(t) \wedge \ddot{\gamma}(t)\|^2}$$

Preuve. On a $\tau(s) = -\mathbf{N} \cdot \frac{d\mathbf{B}}{ds} = -\mathbf{N} \cdot (\mathbf{T} \wedge \frac{d\mathbf{N}}{ds}) = \det(\mathbf{T}, \mathbf{N}, \frac{d\mathbf{N}}{ds})$.

Notons φ un paramétrage unitaire de la même orientation, alors

$$\begin{cases} \mathbf{T} = \frac{d\varphi}{ds} \\ \mathbf{N} = \frac{1}{\kappa} \frac{d^2\varphi}{ds^2} \\ \frac{d\mathbf{N}}{ds} = \frac{1}{\kappa} \frac{d^3\varphi}{ds^3} + \frac{d}{ds} \frac{1}{\kappa} \frac{d^2\varphi}{ds^2} \end{cases}$$

Et donc $\tau(s) = \det(\frac{d\varphi}{ds}, \frac{d^2\varphi}{ds^2}, \frac{d^3\varphi}{ds^3})/\kappa(s)^2$

On a de façon évidente $\frac{d\varphi}{ds} = \frac{\dot{\gamma}}{\|\dot{\gamma}\|}$ et donc en dérivant successivement et en notant que $\frac{d}{ds} = \frac{1}{\|\dot{\gamma}\|} \frac{d}{dt}$ on obtient

$$\begin{cases} \frac{d\varphi}{ds} = \frac{\dot{\gamma}}{\|\dot{\gamma}\|} \\ \frac{d^2\varphi}{ds^2} = \frac{1}{\|\dot{\gamma}\|^2} \ddot{\gamma} + \frac{1}{\|\dot{\gamma}\|} \frac{d}{dt} \left(\frac{1}{\|\dot{\gamma}\|} \right) \dot{\gamma} \\ \frac{d^3\varphi}{ds^3} = \frac{1}{\|\dot{\gamma}\|^3} \dddot{\gamma} + \frac{1}{\|\dot{\gamma}\|} \frac{d}{dt} \left(\frac{1}{\|\dot{\gamma}\|^2} \right) \ddot{\gamma} + \frac{1}{\|\dot{\gamma}\|} \frac{d}{dt} \left(\frac{1}{\|\dot{\gamma}\|} \frac{d}{dt} \left(\frac{1}{\|\dot{\gamma}\|} \right) \dot{\gamma} \right) \end{cases}$$

Le dernier terme étant une combinaison linéaire de $\dot{\gamma}$ et $\ddot{\gamma}$, on a

$$\begin{aligned} \tau(s) &= \det\left(\frac{d\varphi}{ds}, \frac{d^2\varphi}{ds^2}, \frac{d^3\varphi}{ds^3}\right)/\kappa(s)^2 \\ &= \frac{\det(\dot{\gamma}, \ddot{\gamma}, \dddot{\gamma})}{\|\dot{\gamma}\|^6 \kappa(s)^2} \end{aligned}$$

Ce qui en utilisant l'expression de la courbure dans une paramétrisation quelconque conduit au résultat escompté.

Proposition IV.3.10. Au voisinages d'un point on a :

$$\varphi(s_0 + \epsilon) = \varphi(s_0) + \epsilon \left(1 - \frac{\kappa^2 \epsilon^2}{6} \right) \mathbf{T}_0 + \frac{\epsilon^2}{2} \left(\kappa + \frac{\epsilon}{3} \frac{d\kappa}{ds} \right) \mathbf{N}_0 + \frac{\epsilon^3 \kappa \tau}{6} \mathbf{B}_0 + \epsilon^4 v(\epsilon)$$

Ainsi la torsion dans le voisinage de $\epsilon = 0$ exprime la position relative de la courbe vis-à-vis du plan osculateur.

IV.3.3 Hélice

Nous avons rencontré tracé en FigIV.1 la courbe gauche de l'hélice circulaire, cette courbe est à la torsion ce que le cercle est à la courbure. En effet, on peut montrer (mais c'est technique) que pour une courbe gauche régulière en un point de courbure et de torsion non nuls qu'il n'existe qu'une hélice circulaire ayant un point de contacte d'ordre 4 avec la courbe géométrique. On montre aisément que la torsion et la courbure d'une hélice circulaire sont constantes. Nous verrons que la réciproque est également vrai. Mais commençons par définir les hélices en toutes généralités.

Définition IV.3.11 (Hélice). *Une hélice dans \mathbb{R}^3 est une courbe géométrique dont la tangente fait un angle constant avec une direction fixe.*

Remarque IV.3.12. Si \mathbf{T} fais un angle avec une direction constante cela signifie qu'il est contenue dans un cône dont l'ouverture est égale à l'angle en question. Si l'angle en question est égal à $\pi/2$, alors T est plongé dans un plan et cela implique que la courbe est plane et que sa torsion est nulle. Si l'ouverture est nulle en revanche la courbe gauche est simplement une droite.

On peut en particulier pour chaque surface de révolution définir une famille d'hélices plongés dans cette surface. On a déjà vu l'exemple de l'hélice circulaire plongée dans un cylindre de révolution, on a tracée ci-dessous l'hélice plongée dans le cône de révolution, la sphère, et le paraboloïde de révolution. En fait on peut pour chaque surface de révolution tracer une famille d'hélices plongées dans cette surface.

Exercice 15. —

Calculer la courbure et la torsion de l'hélice circulaire de rayon R et de pas p

Exercice 16. —

On considère $t \mapsto (3t - t^3, 3t^2, 3t + t^3)$, montrer qu'il s'agit d'une hélice.

Théorème IV.3.13 (de Lancret). **Une courbe géométrique gauche régulière est une hélice si et seulement si le quotient de sa courbure avec sa torsion est constant.**

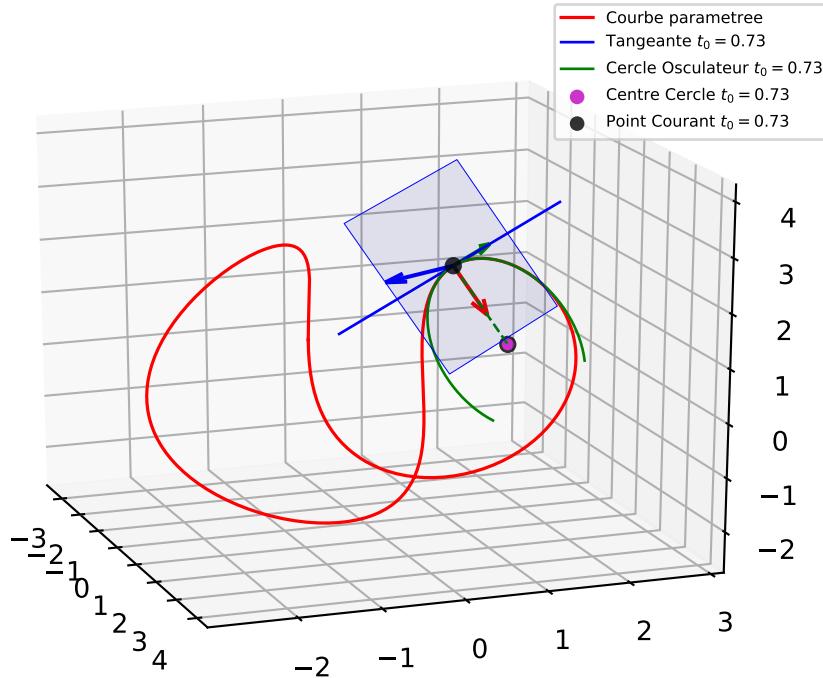


FIGURE IV.7 – Représentation du plan osculateur de la tangente et du cercle osculateur d'une courbe représentant la couture d'une balle de tennis $t \mapsto (a \cos(t) + b \cos(3t), a \sin(t) + b \sin(3t), c \sin(2t))$ (ou $c = 2\sqrt{ab}$ et $b = a/3$)

Preuve. On travaille avec la paramétrisation unitaire $s \mapsto \varphi(s)$, on a $\mathbf{T}(s) = \frac{d\varphi}{ds}$. Supposons qu'il existe \mathbf{V}_0 un vecteur fixe unitaire tel que l'angle entre \mathbf{T} et \mathbf{V}_0 est constant. On a $\mathbf{T} \cdot \mathbf{V}_0 = \cos \theta_0$ est constant. On puisque les courbes paramétrées sont lisses on peut donc dérivé deux fois cette équation. Et donc $\kappa(s)\mathbf{N} \cdot \mathbf{V} = 0$ et $\kappa(-\kappa\mathbf{T} + \tau\mathbf{B})$. Si pour tout s , $\kappa = 0$ alors on travaille avec une droite et donc une hélice sinon, on en déduit que $\mathbf{N} \perp \mathbf{V}_0$ et donc que \mathbf{V}_0 appartient au plan engendré par \mathbf{T}, \mathbf{B} et donc $\kappa(s) \cos \theta_0 = \pm \tau(s) \sin \theta$ soit encore $\frac{\kappa(s)}{\tau(s)} = \pm \tan \theta_0$.

Réiproquement si $\frac{\kappa(s)}{\tau(s)}$ est constant, $\exists \theta \in]-\pi, \pi[$ tel que $\frac{\kappa(s)}{\tau(s)} = \tan \theta$, on pose $\mathbf{V}(s) = \cos \theta \mathbf{T}(s) + \sin \theta \mathbf{B}(s)$ et donc $\frac{d\mathbf{V}}{ds} = (\kappa \cos \theta - \tau \sin \theta) \mathbf{N}(s) = 0$, d'où $\mathbf{V}(s) = \mathbf{V}_0$ est un vecteur unitaire qui fais un angle constant θ avec $\mathbf{T}(s)$. ■

Proposition IV.3.14. Les seules courbes dont la courbure et la torsion sont constantes sont les hélices circulaires.

Preuve. On remarque que dans le cas où torsion et courbure sont constante. $\tau\mathbf{T} + \kappa\mathbf{B}$ est également un vecteur constant. On pose $\tau\mathbf{T} + \kappa\mathbf{B} = \sqrt{\tau^2 + \kappa^2}\mathbf{u}_z = \Omega\mathbf{u}_z$ (ou \mathbf{u}_z est unitaire). On pose $\mathbf{u}_x, \mathbf{u}_y$ deux vecteurs unitaires tels que $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ est une base orthonormale directe. Alors on a $\frac{d\mathbf{T}}{ds} = \Omega\mathbf{u}_z \wedge \mathbf{T}(s)$, on a donc une équation de précession à vecteur constant. En posant $\mathbf{v}(\theta) = \cos(\theta)\mathbf{u}_x + \sin(\theta)\mathbf{u}_y$ et $\cos \psi = \mathbf{T}(0) \cdot \mathbf{u}_z$ (avec $\psi \in [0, \pi]$) sait que ce genre d'équation a pour solution $\mathbf{T}(s) = \cos \psi \mathbf{u}_z + \sin \psi \mathbf{v}(\Omega s + \phi)$. On a donc dans une paramétrisation φ unitaire orientée directement $\frac{d\varphi}{ds} = \mathbf{T}(s)$ et donc $\varphi(s) = \varphi(0) + \cos \psi s \mathbf{u}_z + \frac{\sin \psi}{\sqrt{\tau^2 + \kappa^2}} (\mathbf{v}(\Omega s + \phi - \frac{\pi}{2}) - \mathbf{v}(\phi - \frac{\pi}{2}))$ qui est une hélice circulaire de rayon $\sin \psi / \sqrt{\tau^2 + \kappa^2}$ d'axe $\varphi(0) - \mathbf{v}(\phi - \frac{\pi}{2}) + \mathbb{R}\mathbf{u}_z$ ■

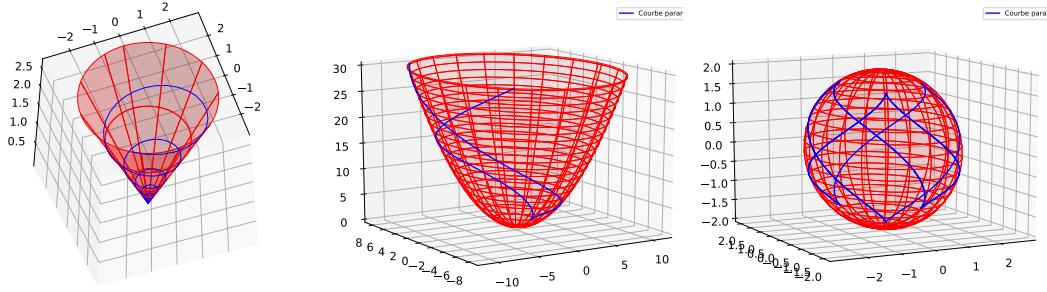


FIGURE IV.8 – Représentation de 3 hélice plongé dans des surfaces de révolutions, à gauche l'hélice plonger dans le cône (sa projection dans le plan correspond à une spirale logarithmique) $t \mapsto (a \exp(kt) \cos(t), a \exp(kt) \sin(t), a \cot \alpha \exp(kt))$, au centre l'hélice plonger dans le paraboloïde de révolution $t \mapsto (a \tan \alpha (\cos(t) + t \sin(t)), a \tan \alpha (\sin(t) - t \cos(t)), \frac{a}{2} \tan^2 \alpha (1 + t^2))$ ou on a pris $k = 1/10$, à droite enfin l'hélice plonger dans la sphère $t \mapsto (a(k \cos(t) \cos(kt) + \sin(t) \sin(kt)), a(k \sin(t) \cos(kt) - \cos(t) \sin(kt)), a\sqrt{1 - k^2} \cos(kt))$ ou on a pris $k = 3/5$

IV.4 Exercices

Exercice 17. —

Les coniques sont les lieux (x, y) du plan vérifiant $P(x, y)$ où P est un polynôme de degré deux en x et y . On note Γ_P la conique de polynôme P . Montrer que $\forall P$, Γ_P est une courbe géométrique associée à un paramétrage rationnel ($\exists r_1, r_2 \in \mathbb{R}(x)$ tel que $t \mapsto (r_1(t), r_2(t))$ est un paramétrage de la courbe géométrique Γ_P).

Exercice 18. —

Soit la courbe paramétrée $t \mapsto (t^2, t^3)$. Calculer l'abscisse curviligne en fonction de t .

Exercice 19. —

Peut-on calculer la longueur d'une ellipse ?

Exercice 20. —

La développée d'une courbe paramétrée plane est le lieu des centres des cercles osculateurs. Montrer que la tangente à la développée est portée par la normale à la courbe.

Exercice 21. —

Soit une courbe paramétrée plane φ telle que $\|\varphi(s) - \varphi(t)\|$ est une fonction de $|t - s|$. Montrer que φ est une portion de cercle ou de droite

Exercice 22. —

Comment évoluent la courbure et la torsion d'un courbe gauche sous l'action d'une homothétie ?

Exercice 23. —

On considère une courbe gauche $\Gamma = \varphi(I)$ et on note κ et τ la torsion, on note $R = \frac{1}{\kappa}$ et $\delta = \frac{1}{\tau}$, montrer que si Γ est plongé dans une sphère alors $R^2 + \frac{dR}{dt}^2 \delta^2$ est constant. Montrer la réciproque (On montrera que $\varphi(s) + R(s)N(s) - R'(s)\delta(s)B(s)$ est constant).

CHAPITRE V

Surfaces régulières

Sommaire

V.1	Surfaces sans bord	77
V.2	Première et seconde formes fondamentales	90
V.3	Theorema Egregium de Gauss	98
V.4	Classification des surfaces compactes à bord	108

L'étude des surfaces régulières est intéressante non seulement pour ces objets en eux-mêmes, mais aussi parce qu'elle est une excellente propédeutique à l'introduction des variétés différentielles que nous aborderons au chapitre suivant.

Une surface sans bord peut être définie de plusieurs manières : comme une paramétrisation, comme le lieu solution d'une équation implicite, comme un graphe d'une fonction à deux variables. Dans le cas des surfaces régulières, les théorèmes d'inversion locale et des fonctions implicites impliqueront l'équivalences de ces définitions. On définira les notions de plan et de fibré tangents, d'orientation, de première et de seconde forme fondamentale ainsi que la notion de courbure de Gauss. On prouvera le *theorema egregium* de Gauss et l'expression des équations de *Mainardi-Coddazzi*. On généralisera ces notions aux surfaces avec un bord, et on effectuera en particulier un classement des surfaces compactes à bords.

Nous souhaitons remercier GUEDJ 2015 dont les documents de cours nous ont beaucoup aidé lors de la rédaction de ce chapitre.

V.1 Surfaces sans bord

V.1.1 Définition

Les surfaces sont des objets de dimension 2. Intuitivement pour les surfaces plongées dans \mathbb{R}^3 , on voit qu'il va donc nous falloir une application suffisamment régulière qui va d'une partie de \mathbb{R}^2 dans \mathbb{R}^3 , cette application devra être injective afin d'éviter de perdre une dimension ou d'avoir des surfaces qui s'auto-intersectent. Afin de définir un plan tangent, il va falloir que les vecteurs des dérivées partielles ne soient pas alignés.

De façon semblable au développement des courbes paramétrées, on commence par définir la notion de nappe paramétrée.

Définition V.1.1 (Nappe paramétrée). *On appelle nappe paramétrée une application lisse injective $\varphi : \mathcal{U} \mapsto \mathbb{R}^3$ sur un ouvert \mathcal{U} de \mathbb{R}^2 . Cette nappe paramétrée est dite régulière si et seulement si la différentielle de φ est de rang 2 en chaque point.*

Si l'on note u et v les coordonnées sur \mathcal{U} , on rappelle que la différentielle d'une telle application peut s'exprimer à l'aide des formes coordonnées par :

$$d\varphi_{u,v}(\cdot) = \partial_u \varphi(u, v) du(\cdot) + \partial_v \varphi(u, v) dv(\cdot)$$

Comme pour les courbes, les paramétrages d'une surface ne sont pas uniques. On est ainsi conduit à définir une relation d'équivalence entre nappes paramétrées.

Deux nappes paramétrées (φ, \mathcal{U}) et (ψ, \mathcal{V}) sont équivalentes $(\varphi, \mathcal{U}) \sim (\psi, \mathcal{V})$ si et seulement si il existe un difféomorphisme lisse $f : \mathcal{U} \rightarrow \mathcal{V}$ tel que $\psi = \varphi \circ f$. On a donc $d\psi_P = d\varphi_{f(P)} \circ df_P$. Le caractère de régularité se transporte donc d'une nappe paramétrée à un autre nappe paramétrée équivalente. En effet puisque f est un difféomorphisme, le rang de $d\psi_P$ est égal à celui de $d\varphi_{f(P)}$. De plus les images de deux nappes paramétrées équivalentes sont les mêmes : $\psi(\mathcal{V}) = \varphi(\mathcal{U})$.

Les classes d'équivalences des nappes géométriques permettront de définir les *nappes géométriques*.

Définition V.1.2 (Nappes géométriques et surfaces). *On appelle nappe géométrique (régulière) S l'image d'une classe d'équivalences de nappes paramétrées (régulières).*

Une surface (régulière) S sans bords est une partie de \mathbb{R}^3 qui est, au voisinage ouvert relatif de chacun de ses points, une nappe géométrique (régulière). C'est-à-dire, $S \subset \mathbb{R}^3$ est une surface (régulière) sans bords si et seulement si $\forall P \in S, \exists O$ ouvert de \mathbb{R}^3 contenant $P, \exists \mathcal{U}$ ouvert de \mathbb{R}^2 et $\exists \varphi : \mathcal{U} \rightarrow \mathbb{R}^3$ homéomorphisme (difféomorphisme si surface régulière) entre \mathcal{U} avec $\varphi(\mathcal{U}) = O \cap S$. On appelle paramétrage local ou nappe paramétrée locale ce genre d'application.

On rappelle que le Théorème d'inversion locale et globale implique qu'une nappe paramétrée régulière est un difféomorphisme lisse.

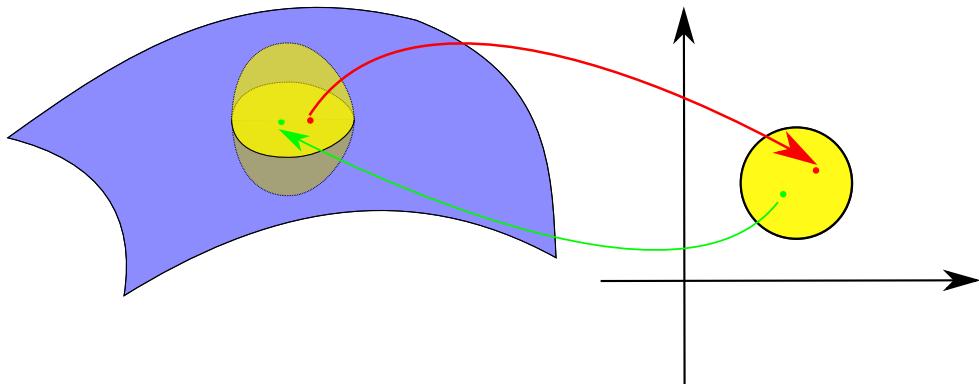


FIGURE V.1 – Représentation d'une surface (bleu clair) et de l'homéomorphisme avec un ouvert de \mathbb{R}^2 . Le point P de la définition est en vert, mais le difféomorphisme est valable dans l'intersection de l'ouvert O (jaune) avec S et donc également pour le point courant rouge.

Exercice 24. — Cône.

On considère le cône de révolution paramétré par $\varphi : (\theta, z) \in [0, 2\pi] \times \mathbb{R} \mapsto (z \cos \theta, z \sin \theta, z)$. Alors la matrice de la différentielle est égale à :

$$d\varphi(\theta, z) = \begin{pmatrix} -z \sin \theta & \cos \theta \\ -z \cos \theta & \sin \theta \\ 0 & 1 \end{pmatrix}$$

Qui est de rang deux si et seulement si $z \neq 0$, c'est-à-dire que le cône de révolution considéré est une surface régulière partout sauf en son sommet.

Exercice 25. — Sphère.

On considère la sphère unité $S^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$. Pour montrer que c'est une surface régulière on considère un point $P \in S^2$ on a alors $\|OP\| = 1$, on considère alors $\mathbf{u}_x, \mathbf{u}_y$ on pose $\mathbf{OM} = \mathbf{u}_z$ et on introduit une base orthonormée de \mathbf{u}_z^\perp de telle manière que $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ soit orthonormal. On définit alors sur $\mathcal{D}_1 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < 1\}$, le disque

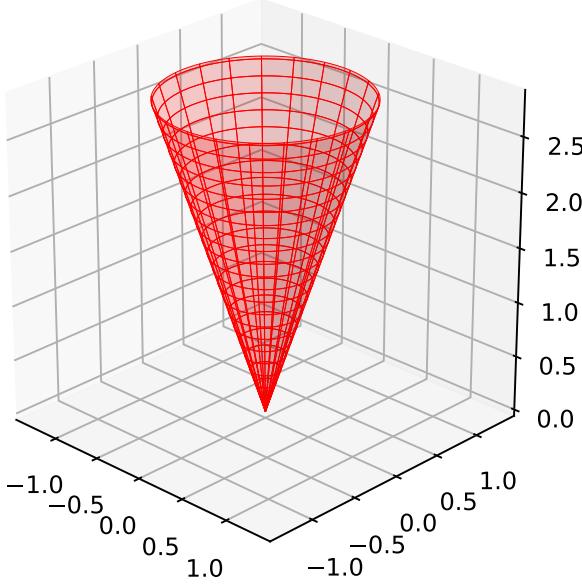


FIGURE V.2 – Représentation d'un Cône.

unité ouvert, l'application $\varphi_P : (x, y) \mapsto x\mathbf{u}_x + y\mathbf{u}_y + \sqrt{1-x^2-y^2}\mathbf{u}_z$, son image est la demi-sphère ouverte $S_{1/2, \mathbf{OP}}^2 = \{M \in \mathbb{R}^3 \mid \|\mathbf{OM}\| = 1 \text{ et } \mathbf{OM} \cdot \mathbf{OP} > 0\}$ qui est un ouvert relatif de S^2 (image réciproque de $]0, +\infty[$ sur l'application continue définie sur S^2 par $M \mapsto \mathbf{OM} \cdot \mathbf{OP}$). Cette application est injective (car $\varphi(x, y) = \varphi(x', y')$ implique $(x, y, \sqrt{1-x^2-y^2}) = (x', y', \sqrt{1-x'^2-y'^2})$ soit encore $(x, y) = (x', y')$). φ_P est également lisse et de rang 2 donc $S_{1/2, \mathbf{OP}}^2 = \{M \in \mathbb{R}^3 \mid \|\mathbf{OM}\| = 1 \text{ et } \mathbf{OM} \cdot \mathbf{OP} > 0\}$ est une nappe géométrique au voisinage de P . Et donc P est un point régulier et S est donc une surface régulière.

Proposition V.1.3. Une partie S de \mathbb{R}^3 est une surface régulière si et seulement si elle est difféomorphe au disque unité au voisinage de chacun de ses points.

Preuve. Considérons S de \mathbb{R}^3 une surface régulière. Alors pour tout point $P \in S$, $\exists O$ un ouvert relatif de S tel que $S \cap O$ soit une nappe géométrique, c'est-à-dire l'image d'une classe d'équivalences de nappes géométriques régulières. Il existe donc \mathcal{U} un ouvert de \mathbb{R}^2 et une application lisse injective (dont la différentielle est de rang 2) $\psi : \mathcal{U} \rightarrow \mathbb{R}^3$ tel que $\text{im } \psi = O \cup S$. Cela implique que $\psi : \mathcal{U} \rightarrow O \cap S$ est un difféomorphisme lisses. Or $P \in S \cap O$ et donc $\exists x_P \in \mathcal{U}$ tel que $\psi(x_P) = P$. Or \mathcal{U} est ouvert, il existe donc $r > 0$ tel que $B(x_P, r) \subset \mathcal{U}$. On pose alors $\forall x \in \mathcal{D}_1$, $\varphi : x \mapsto \psi(x_P + rx)$ cette application est parfaitement définie : c'est un difféomorphisme entre le disque unité et $O' = \psi(B(x_P, r))$ qui est un voisinage ouvert relatif de P (car ψ est un homéomorphisme).

Réciproquement, si S est difféomorphe au disque unité au voisinage de chacun de ses points, alors on considère un point P quelconque et les classes d'équivalences du difféomorphisme d'avec le disque unité. L'image de cette classe d'équivalences est un ouvert relatif de S contenant le point P , cette image est donc une nappe géométrique régulière. ■

V.1.2 Graphes de fonctions

Dans l'exemple 25 de la sphère, nous avons introduit la nappe paramétrée de la demi-sphère. Cette application, telle qu'elle a été construite, fait partie d'un ensemble de nappes paramétrées

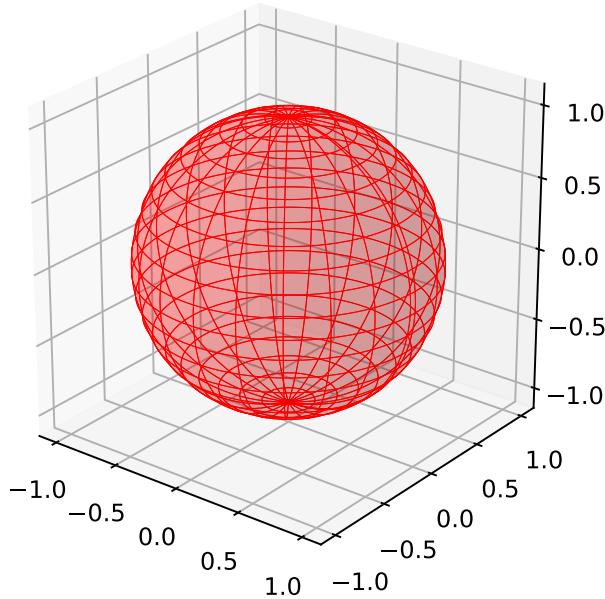


FIGURE V.3 – Représentation d'une sphère.

qu'on appelle graphe de fonction.

Définition V.1.4 (Surface graphes de fonction). *On dit d'une nappe géométrique S qu'elle est graphe d'une fonction si et seulement si l'une des nappes paramétrées φ dont elle est l'image est un graphe, c'est-à-dire s'il existe une base $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ de \mathbb{R}^3 , un point $O \in \mathbb{R}^3$, un ouvert $\mathcal{U} \subset \mathbb{R}^2$ et une fonction $f : \mathcal{U} \rightarrow \mathbb{R}$ tels que l'on ait $\forall (x_1, x_2) \in \mathcal{U}, \varphi(x_1, x_2) = x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + f(x_1, x_2) \mathbf{u}_3$. Un graphe est toujours une nappe géométrique régulière parce que sa forme implique l'injectivité et car on a $\partial_u \varphi = (1, 0, \partial_u f)$ et $\partial_v \varphi = (0, 1, \partial_v f)$ qui sont toujours 2 vecteurs libres.*

Avec une telle définition on voit qu'en tant que nappe géométrique, la sphère ou le cylindre de révolution ne sont pas des graphes de fonction. Cela tient au fait que chaque droite qui *traverse* cette surface, excepté ses tangentes, la traverse au moins deux fois et donc, la valeur de f devrait être double pour décrire l'entièreté de la nappe géométrique. Une autre option consiste à considérer la sphère ou le cylindre non pas en tant que nappe, mais comme une surface.

On a vu dans l'exemple précédent que la sphère pouvait être *localement* décrite comme un graphe. On va voir en particulier que c'est le cas de toutes les surfaces régulières.

Proposition V.1.5. Soit \mathcal{S} une surface régulière. Alors pour tout point $P \in \mathcal{S}$, il existe $O_{\mathcal{S}} = O \cap \mathcal{S}$ (O ouvert de \mathbb{R}^3) ouvert relatif de \mathcal{S} tel que $O_{\mathcal{S}}$ soit graphe d'une fonction.

Preuve. Si \mathcal{S} est une surface régulière, alors $\forall P \in \mathcal{S}, \exists O$ ouvert de \mathbb{R}^3 contenant $P, \exists \mathcal{U}$ ouvert de \mathbb{R}^2 et $\exists \varphi : \mathcal{U} \rightarrow \mathbb{R}^3$ difféomorphisme entre \mathcal{U} avec $\varphi(\mathcal{U}) = O \cap \mathcal{S}$. On note $(u_P, v_P) \in \mathcal{U}$ tel que $\varphi(u_P, v_P) = P$. Puisque la différentielle de φ est de rang 2, $\partial_u \varphi(u_P, v_P)$ et $\partial_v \varphi(u_P, v_P)$ sont libres. On note $F = P + \text{vect}(\partial_u \varphi(u_P, v_P), \partial_v \varphi(u_P, v_P))$, le sous-espace affine de dimension deux engendré par ces vecteurs. On pose $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ tel que $\mathbf{u}_x, \mathbf{u}_y$ soit une base orthonormée de F et

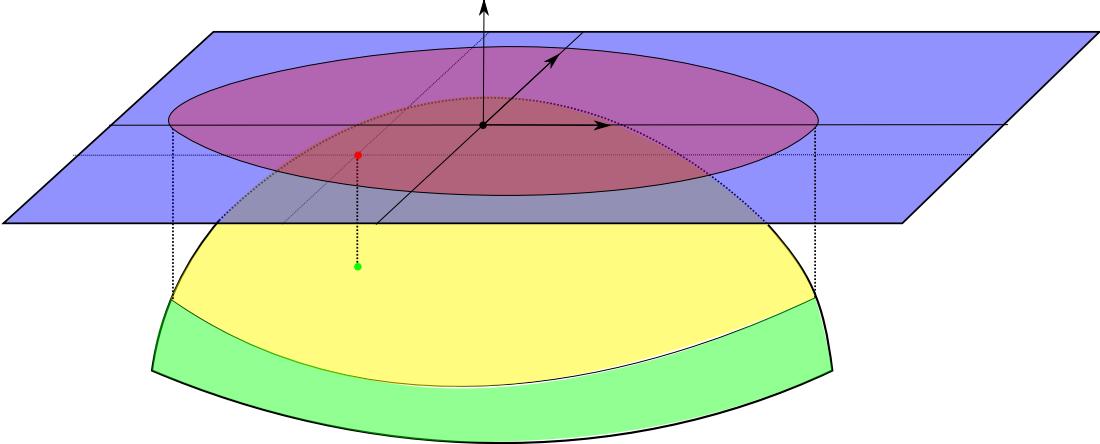


FIGURE V.4 – Représentation d'un graphe local. L'ouvert relatif $\tilde{O} \cap \mathcal{S}$ est en jaune, les vecteurs $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ sont en noir, F est en bleu, P est le point représenté en noir et \mathcal{S} en vert.

$\mathbf{u}_z = \mathbf{u}_x \wedge \mathbf{u}_y$. On définit sur \mathcal{U} les applications :

$$\begin{array}{rccc} & : & \mathcal{U} & \rightarrow & \mathbb{R} \\ x & (u, v) & \mapsto & \mathbf{u}_x \cdot \mathbf{P}\varphi(u, v) \\ y & (u, v) & \mapsto & \mathbf{u}_y \cdot \mathbf{P}\varphi(u, v) \\ z & (u, v) & \mapsto & \mathbf{u}_z \cdot \mathbf{P}\varphi(u, v) \end{array}$$

Alors l'application regroupant les deux premières,

$$\begin{array}{rccc} \psi & : & \mathcal{U} & \rightarrow & F \\ & & (u, v) & \mapsto & P + x(u, v)\mathbf{u}_x + y(u, v)\mathbf{u}_y \end{array}$$

est lisse par construction. De plus

$$\partial_u \psi = (\mathbf{u}_x \cdot \partial_u \varphi) \mathbf{u}_x + (\mathbf{u}_y \cdot \partial_u \varphi) \mathbf{u}_y$$

C'est la projection orthogonale dans F du vecteur $\partial_u \varphi$. Or $\partial_u \varphi(u_P, v_P) \in F$, donc $\partial_u \psi(u_P, v_P) = \partial_u \varphi(u_P, v_P)$. Symétriquement on a $\partial_v \psi(u_P, v_P) = \partial_v \varphi(u_P, v_P)$ et donc la différentielle $d\psi_{u_P, v_P}$ est de rang 2 au point $(u_P, v_P) \in \mathcal{U}$.

Puisque la différentielle $d\psi_{u_P, v_P} : \mathbb{R}^2 \rightarrow F$ est bijective, le théorème d'inversion locale assure l'existence d'un ouvert $\tilde{\mathcal{U}} \subset \mathcal{U}$ contenant (u_P, v_P) et d'un ouvert $\tilde{\mathcal{V}}$ de F tels que la restriction de $\tilde{\psi} = \psi|_{\tilde{\mathcal{U}}}$ est un difféomorphisme lisse, soit

$$\forall M \in \tilde{\mathcal{V}} \exists! x, y \in \mathbb{R}^2 / M = P + x\mathbf{u}_x + y\mathbf{u}_y \in \tilde{\mathcal{V}}$$

On appelle $\tilde{\xi} : \tilde{\mathcal{V}} \rightarrow \tilde{\mathcal{U}}$ l'application réciproque de $\tilde{\psi}$, alors $\tilde{\xi}(P + x\mathbf{u}_x + y\mathbf{u}_y) = (u(x, y), v(x, y))$ où $u(x, y)$ et $v(x, y)$ sont les applications composantes de $\tilde{\xi}$ qui sont lisses puisque $\tilde{\xi}$ l'est. On pose alors,

$$\begin{array}{rccc} f & : & \tilde{\mathcal{V}} & \rightarrow & \mathbb{R} \\ & & P + x\mathbf{u}_x + y\mathbf{u}_y & \mapsto & z(u(x, y), v(x, y)) \end{array}$$

qui est lisse par composition d'applications lisses.

Et donc, le graphe

$$\begin{aligned}\Theta : \tilde{\mathcal{V}} &\rightarrow \mathbb{R}^3 \\ P + x\mathbf{u}_x + y\mathbf{u}_y &\mapsto P + x\mathbf{u}_x + y\mathbf{u}_y + f(x, y)\mathbf{u}_z\end{aligned}$$

a pour image une partie de $O \cap \mathcal{S}$. En effet, pour $\forall u, v \in \tilde{\mathcal{V}}$ et $\forall P + x\mathbf{u}_x + y\mathbf{u}_y \in \tilde{\mathcal{V}}$,

$$\begin{aligned}\tilde{\varphi}(u, v) &= P + P\tilde{\varphi}(u, v) = \tilde{\psi}(u, v) + \tilde{z}(u, v)\mathbf{u}_z \\ \tilde{\varphi} \circ \tilde{\xi}(P + x\mathbf{u}_x + y\mathbf{u}_y) &= \tilde{\psi} \circ \tilde{\xi}(P + x\mathbf{u}_x + y\mathbf{u}_y) + z(\tilde{\xi}(P + x\mathbf{u}_x + y\mathbf{u}_y))\mathbf{u}_z \\ &= P + x\mathbf{u}_x + y\mathbf{u}_y + z(u(x, y), v(x, y))\mathbf{u}_z \\ &= P + x\mathbf{u}_x + y\mathbf{u}_y + f(x, y)\mathbf{u}_z = \Theta(P + x\mathbf{u}_x + y\mathbf{u}_y)\end{aligned}$$

Ainsi $(\tilde{\varphi}, \tilde{\mathcal{U}}) \sim (\Theta, \tilde{\mathcal{V}})$. L'image $\varphi(\tilde{\mathcal{U}})$ est un ouvert relatif de $S \cap O$. En effet, $\tilde{\mathcal{U}}$ est un ouvert inclus dans \mathcal{U} et puisque φ est un difféomorphisme de \mathcal{U} vers $S \cap O$, il transporte les ouverts relatifs de l'espace de départ sur des ouverts relatifs de l'espace d'arrivée. Il existe donc O' tel que $\varphi(\tilde{\mathcal{U}}) = O' \cap O \cap S = \tilde{\mathcal{O}} \cap S$ (on a noté $O' \cap O = \tilde{\mathcal{O}}$). De plus $(u_P, v_P) \in \tilde{\mathcal{U}}$ donc $P \in \tilde{\mathcal{O}} \cap S$.

Résumons : $\forall P \in S, \exists \tilde{\mathcal{U}}$ ouvert de $\mathbb{R}^2, \exists \tilde{\mathcal{O}}$ ouvert de \mathbb{R}^3 et un graphe Θ tel que $\Theta(\tilde{\mathcal{U}}) = \tilde{\mathcal{O}} \cap S$. ■

Exemple V.1.6. On considère le paraboloïde hyperbolique \mathcal{H} . C'est une nappe géométrique qui peut être définie par la nappe paramétrée $\varphi : (x, y) \mapsto x\mathbf{e}_x + y\mathbf{e}_y + xy\mathbf{e}_z$. C'est donc le graphe de $f : (x, y) \mapsto xy$ dans les vecteurs de la base canonique. Elle est également définie à l'aide de l'équation algébrique

$$\mathcal{H} = \{(x, y, z) \in \mathbb{R}^3 \mid z = xy\}$$

On remarque que la droite $\mathcal{D}_x : t \mapsto (x, 1, x) + t(0, 1, x)$ est incluse dans cette surface. Une de ces droites passe par chaque point. On parle de *surface réglée*, on en fera une définition plus précise plus loin.

V.1.3 Submersion

Nous avons vu dans l'exemple précédent que le paraboloïde hyperbolique était une surface régulière pouvant être définie à l'aide d'une équation implicite algébrique. Ce genre de définition est ce qu'on appelle une définition par *submersion*. Des conditions sur cette équation impliquent la régularité de la surface ainsi définie.

Proposition V.1.7 (Submersion). Soit $f : \mathcal{V} \rightarrow \mathbb{R}$ (où \mathcal{V} ouvert de \mathbb{R}^3 et f lisse) et définissons l'ensemble défini par l'équation implicite :

$$\mathcal{S} = f^{-1}(\{0\}) = \{(x, y, z) \in \mathcal{V} \mid f(x, y, z) = 0\}$$

Alors $\forall P \in \mathcal{S}$ tel que df_P est *surjective*, $\exists O$ ouvert de \mathbb{R}^3 tel que $\mathcal{S} \cap O$ soit une *surface régulière*.

En particulier si le gradient de f ne s'annule pas sur \mathcal{S} alors \mathcal{S} est une surface régulière.

Preuve. Si df_P est surjective alors $\nabla f(P) \neq 0$ (on dis que f est une *submersion* au voisinage de P), on peut donc, quitte à permuter les coordonnées, supposer que $\partial_z f(P) \neq 0$, puisque $f(P) = 0$ et f est lisse on est dans les hypothèses du théorème des fonctions implicites, il existe donc un ouvert \mathcal{U} de \mathbb{R}^2 , un voisinage ouvert O de P et une fonction lisse $h : \mathcal{U} \rightarrow \mathbb{R}$ tel que

$$(x, y, z) \in \mathcal{S} \cap O \iff z = h(x, y)$$

$\mathcal{S} \cap O$ est donc un graphe de la fonction h , et donc une nappe géométrique liée à la nappe paramétrée $(x, y) \in \mathcal{U} \mapsto (x, y, h(x, y)) \in \mathcal{S} \cap O$. Cette nappe est régulière puisque c'est un graphe. $\mathcal{S} \cap O$ est donc une surface régulière. ■

Remarque V.1.8. Toutes les surfaces ne sont pas définies via une équation implicite. On verra plus loin que le ruban de Möbius par exemple ne peut pas être défini par submersion. Il est néanmoins très facile quand cela est possible de vérifier qu'une surface est régulière en utilisant cette proposition. Pour la sphère par exemple $f : (x, y, z) \mapsto x^2 + y^2 + z^2 - R^2$ a pour gradient $2OM$ et donc a pour norme $2R$ sur la sphère, il ne s'annule pas : la sphère est une surface régulière.

Les surfaces définies par une équation implicite sont des fermés comme image réciproque d'un fermé par une application continue. Ce qui permet pour $V = \mathbb{R}^3$ d'en déduire des propriétés

topologiques intéressantes, par exemple si la fonction f ne s'annule plus au delà d'un certain rayon, si elle tends vers l'infini par exemple on en déduit que la surface est bornée et donc que c'est un compact. Si S est régulière, la proposition V.1.3 implique que chaque point de la surface appartient à un voisinage relatif ouvert, on a alors $\text{Int}(S) = S$. Si l'on définit le "bord" par $\partial S = \bar{S} - \text{Int}(S) = \bar{S} - S$ cet ensemble est vide pour une surface définie implicitement puisque on aura alors S fermé de \mathbb{R}^3 et donc $\bar{S} = S$. C'est le cas de la sphère ou du tore par exemple.

Exercice 26. — Tore.

On définit le tore de rayon $a > 0$ et $b > 0$ par l'image de l'application lisse

$$\varphi : (u, v) \mapsto (a + b \cos v) \cos u \mathbf{e}_x + (a + b \cos v) \sin u \mathbf{e}_y + b \sin v \mathbf{e}_z$$

On pourrait penser que $f_{a,b} = (\sqrt{x^2 + y^2} - 2)^2 + z^2 - b^2$ pourrait convenir pour définir implicitement le tore. Or elle n'est pas différentiable sur l'axe $O\mathbf{e}_z$. Pour assurer son caractère lisse il va falloir éliminer cette racine carré. Commençons par $x(u, v)^2 + y(u, v)^2 + z(u, v)^2 = a^2 + b^2 + 2b \cos v$ et donc $(x(u, v)^2 + y(u, v)^2 + z(u, v)^2 - a^2 - b^2)^2 = 4b^2(1 - \sin^2 v) = 4(b^2 - z(u, v)^2)$, on pose alors $f_{a,b}(x, y, z) = (x^2 + y^2 + z^2 - a^2 - b^2)^2 - 4a^2(b^2 - z^2)$. On montre alors que le gradient de $f_{a,b}$ est égal à $(4(x^2 + y^2 + z^2 - a^2 - b^2)x; 4(x^2 + y^2 + z^2 - a^2 - b^2)y; 4(x^2 + y^2 + z^2 - a^2 - b^2)z)$. Ses points d'annulations sont $(0, 0, 0)$ (qui appartient au tore si et seulement si $a = b$), l'intersection du plan $z = 0$ avec la sphère de rayon (qui n'appartient pas au tore car $|b - a| < \sqrt{a^2 + b^2} < a + b$ puisque $a > 0$ et $b > 0$), enfin il y a les points $(0, 0, \pm\sqrt{b^2 - a^2})$ (qui existent et appartiennent au tore si et seulement si $b \geq a$). Les tores pour lesquels $b < a$ sont donc des surfaces régulières. Pour les autres on montre que ce ne sont même pas des nappes paramétrée puisque φ n'est pas injective.

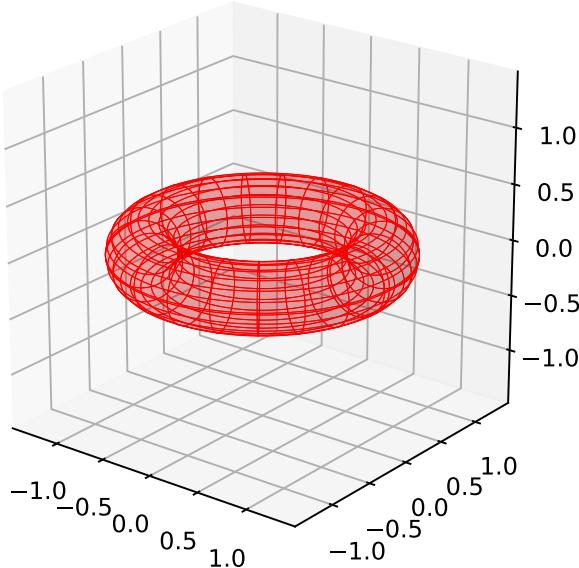


FIGURE V.5 – A gauche représentation d'un tore $a = 1$ et $b = 0.3$.

Nous résumons les différents points de vue équivalents pour définir une surface géométrique régulière :

Proposition V.1.9. Les propriétés suivantes sont localement équivalentes :

- (i) \mathcal{S} est une surface paramétrée régulière ;
- (ii) $\mathcal{S} \subset \mathbb{R}^3$ est le graphe d'une fonction de deux variables réelles ;
- (iii) $\mathcal{S} = \{(x, y, z) \in V \mid f(x, y, z) = 0\}$ avec $f : V \rightarrow \mathbb{R}$ une fonction lisse dont la différentielle est surjective.

Preuve. La première implication $i) \Rightarrow ii)$ résulte de la Proposition V.1.5 tandis que $iii) \Rightarrow i)$ résulte de la Proposition V.1.7. Il reste à vérifier $ii) \Rightarrow iii)$. Supposons donc que $\mathcal{S} = \{(x, y, z) \mid z = h(x, y)\}$ est localement le graphe d'une fonction lisse. Alors $\mathcal{S} = f^{-1}(\{0\})$ avec $f(x, y, z) = z - h(x, y)$ submersive, puisque $\nabla f = (-\partial_x h, -\partial_y h, 1)$. ■

V.1.4 Les Quadriques

Les quadriques sont les ensembles définis implicitement par $S = \{(x, y, z) \in \mathbb{R}^3 \mid P(x, y, z) = 0\}$ où $P \in \mathbb{R}_2[x, y, z]$.

Or tout polynôme de degré deux s'écrit comme la somme d'une forme quadratique, d'une forme linéaire et d'une constante : $P(x, y, z) = Q(x, y, z) + L(x, y, z) + R$. Puisque toute forme quadratique est "diagonalisable" dans une base orthonormale, il existe une base orthonormale directe (e_x, e_y, e_z) et en particulier une base duale associée $\ell_X = \langle e_x | ? \rangle, \ell_Y = \langle e_y | ? \rangle, \ell_Z = \langle e_z | ? \rangle$ (forme coordonnées), et 3 scalaires a_X, a_Y, a_Z tels que $Q(x, y, z) = a_X \ell_X(x, y, z)^2 + a_Y \ell_Y(x, y, z)^2 + a_Z \ell_Z(x, y, z)^2$. Puisque ℓ_X, ℓ_Y, ℓ_Z est une base de l'espace vectoriel des formes linéaires (famille libre puisque orthonormale dans un espace de dimension 3) il existe donc 3 scalaires b_X, b_Y, b_Z tel que $L(x, y, z) = b_X \ell_X(x, y, z) + b_Y \ell_Y(x, y, z) + b_Z \ell_Z(x, y, z)$. On a donc en toute généralité

$$P(x, y, z) = a_X \ell_X(x, y, z)^2 + a_Y \ell_Y(x, y, z)^2 + a_Z \ell_Z(x, y, z)^2 + b_X \ell_X(x, y, z) + b_Y \ell_Y(x, y, z) + b_Z \ell_Z(x, y, z) + R$$

Pour aller plus loin concernant l'étude des quadriques il va falloir distinguer les 3 cas, les quadriques non dégénérées, celles qui le sont une fois et celles qui le sont deux fois (celles qui le sont 3 fois sont nulles et donc la surface étudiée est un plan affine si $L \neq 0$, l'ensemble vide ou l'espace tout entier sinon).

Quadriques non dégénérées

Si la forme quadratique est non dégénérée alors il existe $A, B, C, d, e, f \in \mathbb{R}$ tels que : $P(x, y, z) = \pm \frac{1}{A^2} [\ell_X(x, y, z)^2 + 2d\ell_X(x, y, z)] \pm \frac{1}{B^2} [\ell_Y(x, y, z)^2 + 2e\ell_Y(x, y, z)] \pm \frac{1}{C^2} [\ell_Z(x, y, z)^2 + 2f\ell_Z(x, y, z)] + R$. On fait apparaître ainsi des formes "pré-canonical". On pose alors les formes affines $X(x, y, z) = \ell_X(x, y, z) + d, Y(x, y, z) = \ell_Y(x, y, z) + e$ et $Z(x, y, z) = \ell_Z(x, y, z) + f$, elles correspondent à un changement de repère (après la rotation correspondant à l'introduction des formes ℓ_X, ℓ_Y, ℓ_Z , le passage à X, Y, Z correspond à une translation), on a alors $P(x, y, z) = \pm \frac{1}{A^2} [X(x, y, z)]^2 \pm \frac{1}{B^2} [Y(x, y, z)]^2 \pm \frac{1}{C^2} [Z(x, y, z)]^2 + R \mp \frac{d^2}{A^2} \mp \frac{e^2}{B^2} \mp \frac{f^2}{C^2}$. On pose alors $\xi = |R \mp \frac{d^2}{A^2} \mp \frac{e^2}{B^2} \mp \frac{f^2}{C^2}|$, et donc à nouveau deux sous-cas qui se dégagent :

1. Soit $\xi = 0$ et donc $P(x, y, z) = \pm \frac{1}{A^2} [X(x, y, z)]^2 \pm \frac{1}{B^2} [Y(x, y, z)]^2 \pm \frac{1}{C^2} [Z(x, y, z)]^2$, deux sous-sous-cas se dégagent à nouveau :

- a) L'ensemble en question est réduit à un point si la forme quadratique est définie (positive ou négative) ;
- b) Sinon il s'agit d'un cône elliptique, il ne possède qu'un point non-régulier, c'est le point $X = Y = Z = 0$. C'est une surface, connexe, réglée, non bornée.

2. Soit $\xi \neq 0$, on obtient $P(x, y, z) = \xi \left[\pm \frac{1}{a^2} (X(x, y, z))^2 \pm \frac{1}{b^2} (Y(x, y, z))^2 \pm \frac{1}{c^2} (Z(x, y, z))^2 \pm 1 \right]$ où on a posé $a^2, b^2, c^2 = \xi A^2, \xi B^2, \xi C^2$. L'étude de P/ξ est équivalente à celle de P . Ici on est amené à distinguer quatre sous-sous-cas :

- a) Si la forme quadratique est définie du même signe que le reste $\pm \xi$ alors $S =$;
- b) Si la forme quadratique est définie du signe opposé au reste $\pm \xi$ alors S est un ellipsoïde. C'est une surface régulière, connexe, la seule quadrique compacte. Son équation implicite cartésienne est

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} - 1 = 0$$

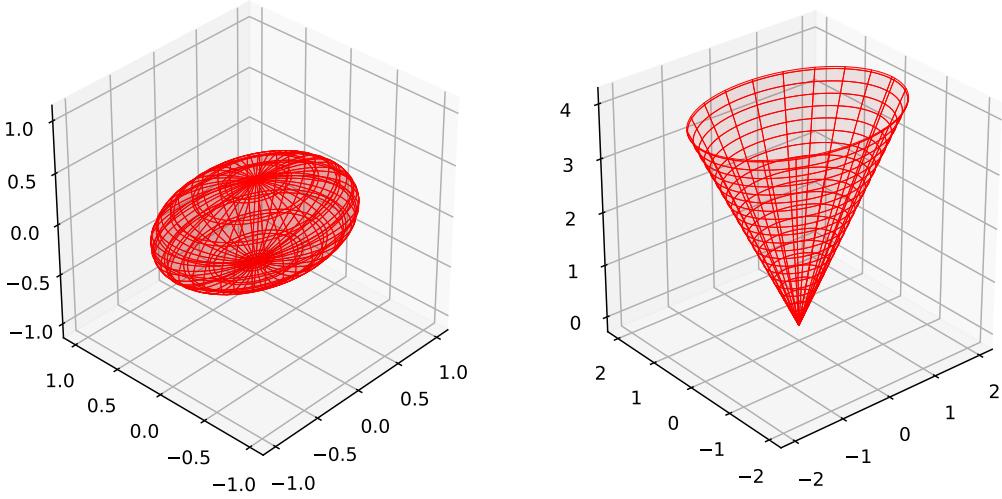


FIGURE V.6 – À gauche représentation d'un ellipsoïde d'axe $a = 1$, $b = 0.7$ et $c = 0.4$, à droite un cône elliptique d'axe $a = 1$, $b = 0.7$ et $c = 2.0$.

- c) Si la forme quadratique n'est pas définie et la trace de la signature est d'un signe opposé au reste, alors on obtient un hyperboloïde à une nappe. C'est une surface régulière, connexe, réglée non bornée. Son équation cartésienne est

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} - \frac{Z^2}{c^2} - 1 = 0$$

- d) Si la forme quadratique n'est pas définie et la trace de la signature est du même signe que le reste, alors on obtient un hyperboloïde à deux nappe. C'est une surface régulière, composée de deux morceaux connexes, non bornés. Son équation implicite cartésienne est

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} - \frac{Z^2}{c^2} + 1 = 0$$

Quadrique dégénérée une fois

Si la quadrique est une fois dégénérée on peut supposer sans perte de généralité qu'il s'agit de $a_Z = 0$. Alors il existe $A, B, d, e, F \in \mathbb{R}$ tels que

$$P(x, y, z) = \pm \frac{1}{A^2} [\ell_X(x, y, z)^2 + 2d\ell_X(x, y, z)] \pm \frac{1}{B^2} [\ell_Y(x, y, z)^2 + 2e\ell_Y(x, y, z)] + F\ell_Z(x, y, z) + R \mp \frac{d^2}{A^2} \mp \frac{e^2}{B^2}$$

On pose alors $X(x, y, z) = \ell_X(x, y, z) + d$, $Y(x, y, z) = \ell_Y(x, y, z) + e$, et $r = \pm R \mp \frac{d^2}{A^2} \mp \frac{e^2}{B^2}$, on est amené à considérer 2 sous-cas

1. Si $F \neq 0$ on pose $Z(x, y, z) = \ell_Z(x, y, z) + (R \mp \frac{d^2}{A^2} \mp \frac{e^2}{B^2})/F$ et on est alors ramené à $P(x, y, z) = P \left[\pm \frac{1}{a^2} (X(x, y, z))^2 \pm \frac{1}{b^2} (Y(x, y, z))^2 + Z(x, y, z) \right]$. On est amené à considérer 2 sous-sous-cas

- a) La forme quadratique restant est définie, l'ensemble est alors un paraboloïde elliptique, c'est une surface régulière, connexe, non borné. Son équation cartésienne est

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} - Z = 0$$

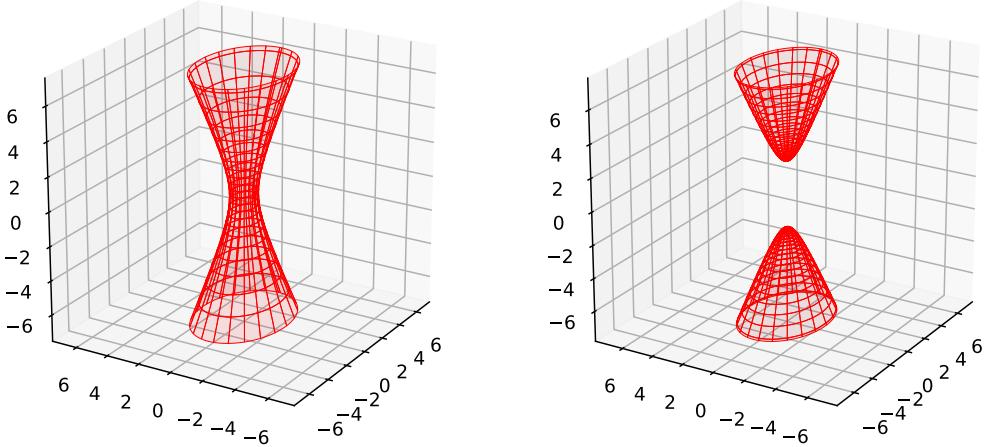


FIGURE V.7 – À gauche représentation d'un hyperboïde à une nappe d'axe $a = 1$, $b = 0.7$ et $c = 2.0$, à droite un hyperboïde à deux nappes d'axe $a = 1$, $b = 0.7$ et $c = 2.0$.

- b) La forme quadratique restant n'est pas définie, l'ensemble est alors un paraboloïde hyperbolique, c'est une surface régulière connexe, non bornée. Son équation cartésienne est

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} - Z = 0$$

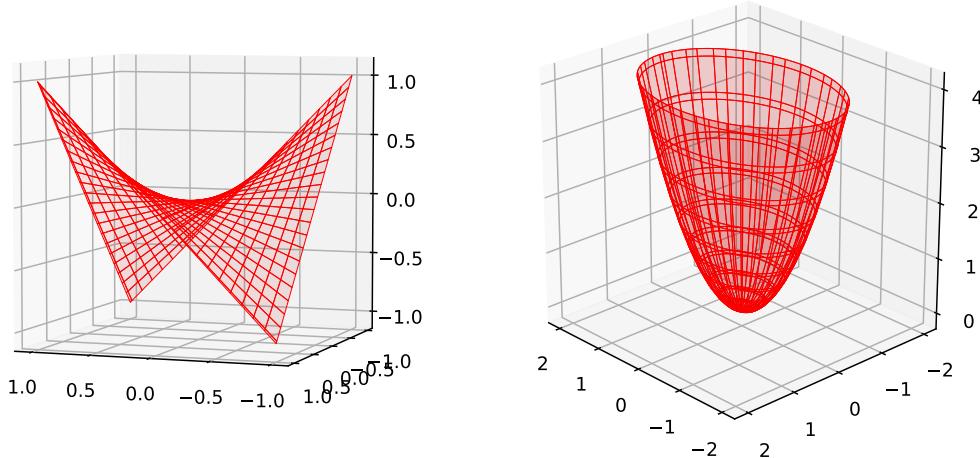


FIGURE V.8 – À gauche représentation d'un paraboloïde hyperbolique d'axe $a = 1$ et $b = 0.7$, à droite un paraboloïde elliptique d'axe $a = 1$ et $b = 0.7$.

2. Si $F = 0$ on pose $Z(x, y, z) = \ell_Z(x, y, z)$ et on est alors ramené à $P(x, y, z) = [\pm \frac{1}{a^2} (X(x, y, z))^2 \pm \frac{1}{b^2} (Y(x, y, z))^2 \pm r]$. On est à nouveau amené à considérer 3 sous-sous-cas,

- a) La forme quadradique restante est définie et du même signe que le reste qui est non nul, l'ensemble en question est vide ;
- b) La forme quadradique restante est définie et du signe opposé (ou nul) au reste. L'ensemble est un cylindre elliptique (la droite $X = Y = 0$). Dans les cas non réduit à la droite, c'est une surface régulière non bornée et connexe par arc. Son équation cartésienne est

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} - 1 = 0$$

- c) La forme quadratique restante n'est pas définie alors l'ensemble S est un cylindre hyperbolique, il est composé de deux nappes connexes, c'est une surface régulière non bornée. Son équation cartésienne est

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} - 1 = 0$$

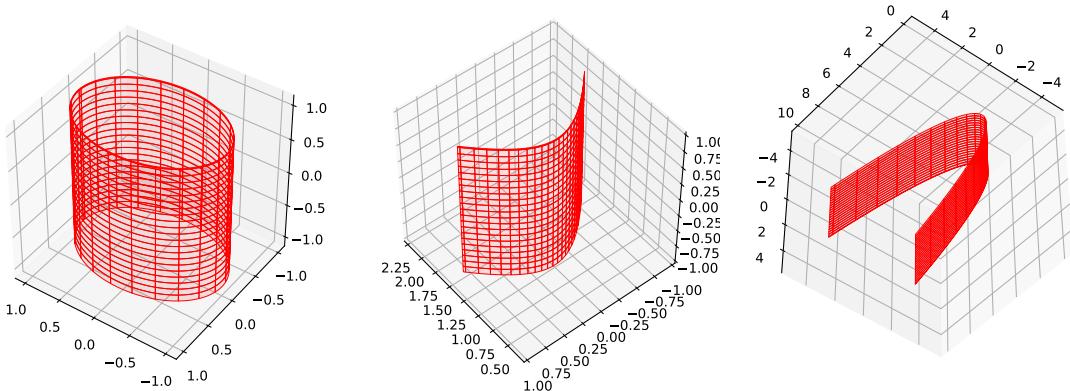


FIGURE V.9 – À gauche représentation d'un cylindre elliptique d'axe $a = 1$ et $b = 0.7$, au centre une cylindre parabolique $a = 1$, à droite un cylindre hyperbolique d'axe $a = 1$ et $b = 0.7$.

Quadrique dégénérée deux fois

On se retrouve dans ce cas avec des polynômes de la forme $a_X \ell_X(x, y, z)^2 + b_X \ell_X(x, y, z) + b_Y \ell_Y(x, y, z) + b_Z \ell_Z(x, y, z) + R$ qu'on peut réduire à l'aide de deux formes affines orthogonales X et Y de la manière suivante : $\frac{X^2}{a^2} + FY + R$. On est amené à distinguer trois sous-cas :

1. Si $F = 0$ et R est strictement positif alors l'ensemble est vide
2. Si $F = 0$ et R est négatif (ou nul) on a la réunion de deux plans parallèles (confondus), c'est une surface régulière, non bornée, à deux sous-ensembles connexes ;
3. Si $F \neq 0$ alors on travaille avec un cylindre parabolique. Son équation cartésienne est

$$\frac{X^2}{a^2} - Y = 0$$

V.1.5 Plan tangent

On a vu qu'une surface était régulière autour d'un point si et seulement si la nappe paramétrée qui la représente était localement lisse, injective avec une différentielle de rang 2. L'espace affine engendré par ce point et l'image de cette différentielle est appelé *espace tangent*.

Définition V.1.10 (Espace tangent et espace normal). *Soit S une surface régulière alors $\forall P$, on note φ_P une nappe paramétrée sur un ouvert de \mathbb{R}^2 dans un voisinage ouvert relatif de P , alors on appelle espace tangent en $P = \varphi(m)$, qu'on note $T_P(S)$ l'espace affine $P + \text{im } d\varphi_m$. La droite $P + \text{im } d\varphi(m)^\perp$ est alors appelée espace normal, on le note $N_P(S)$.*

Remarque V.1.11. Une ambiguïté sera utilisée afin d'éviter d'alourdir les notations dans ce qui suivra, en effet tantôt $\mathcal{T}_P(\mathcal{S})$ désignera l'espace affines $P + \text{im } d\varphi_m$, tantôt il désignera l'espace vectoriel associé $\text{im } d\varphi_m$.

L'image de cette différentielle est indépendante du paramétrage choisi comme on l'a déjà vu. En effet, si $(\varphi, \mathcal{U}) \sim (\psi, \mathcal{V})$ alors $\exists f : \mathcal{V} \rightarrow \mathcal{U}$ un difféomorphisme tel que $\psi = \varphi \circ f$ et donc $\forall m \in \mathcal{V}$ on a $d\psi_m = d\varphi_{f(m)} \circ df_m$ et donc $\text{im } d\psi_m = \text{im } (d\varphi_{f(m)} \circ df_m) = d\varphi_{f(m)} \circ df_m(\mathbb{R}^2)$. Or df_m est un isomorphisme (puisque f est difféomorphisme) et donc $df_m(\mathbb{R}^2) = \mathbb{R}^2$ et donc $\text{im } d\psi_m = d\varphi_{f(m)}(\mathbb{R}^2) = \text{im } d\varphi_{f(m)}$.

On peut remarquer alors que pour chaque *courbe régulière tracée sur la surface régulière*, c'est-à-dire localement la composée d'une nappe paramétrée locale avec une courbe plane régulière passant dans son domaine de définition, le vecteur tangent de cette courbe est inclus dans le plan tangent en ce point. La réciproque est vraie.

Dans une une nappe paramétrée locale φ , il vient immédiatement que $\mathcal{T}_P(\mathcal{S}) = P + \text{vect}(\partial_u \varphi(m), \partial_v \varphi(m))$, ce qui justifie que $(\partial_u \varphi(m), \partial_v \varphi(m))$ soit qualifiée de *base naturelle* de l'espace tangent.

Typiquement le plan tangent a un contacte d'ordre 2 avec la surface, c'est-à-dire $\forall P \in S$, pour toute nappe paramétrée locale $\varphi : \mathcal{U} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dont la nappes géométrique contient $P = \varphi(p)$, pour toute courbe paramétrée régulière $\gamma : t \in I \rightarrow \mathcal{U}$ telle que $\gamma(0) = m$, la distance de $\Gamma = \varphi \circ \gamma$ au plan tangent croit comme $t^2 e(t)$ (où $e(t)$ est borné au voisinage de 0).

Proposition V.1.12. Le plan tangent à une surface régulière S en un point P est exactement constitué des vecteurs tangents en P aux courbes régulières tracées sur S (qui passent par P).

Preuve. Pour tout point $P \in S$ et pour tout $\mathbf{V} \in \mathcal{T}_P(S)$, $\exists a, b \in \mathbb{R}^2$ tel que $\mathbf{V} = a\partial_u \varphi(P) + b\partial_b \varphi(P)$. Il suffit de considérer, dans un paramétrage local $\varphi : \mathcal{U} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ à $P = \varphi(p)$, l'image Γ par φ de la courbe paramétrée définie pour $\epsilon > 0$ suffisamment petit, $\gamma : t \in [-\epsilon, \epsilon] \mapsto p + (at, bt) \in \mathcal{U}$. Sa position en $t = 0$ est $\Gamma(0) = \varphi \circ \gamma(0) = P$ et son vecteur tangent est alors égal à $\frac{d\Gamma}{dt}(0) = a\partial_u \varphi(P) + b\partial_b \varphi(P)$ qui appartient donc au plan tangent. ■

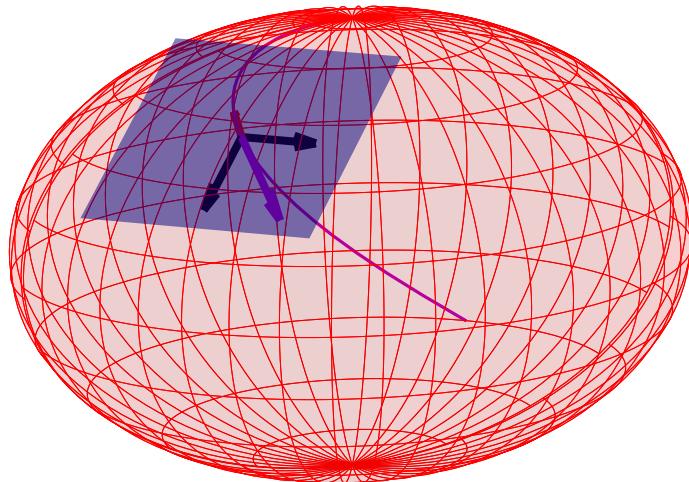


FIGURE V.10 – Représentation (rouge) d'un ellipsoïde d'axe $a = 3$ et $b = 2.4$ et $c = 2$, en bleu le plan tangent et en noir la base naturelle, l'espace normal est porté par le vecteur rouge, on remarque que le vecteur tangent à la courbe tracé sur la surface appartient au plan tangent.

V.1.6 Continuité sur une surface

Comment s'assurer rapidement du caractère lisse, d'une application $f : S \rightarrow R^d$ définie sur une surface régulière, quand celle-ci n'est pas simplement issue d'une nappe géométrique ? Comment

calculer sa différentielle si l'application n'est plus définie sur l'ouvert d'un espace vectoriel ? Plus généralement comment caractériser les propriétés des applications qui étaient alors définies sur des ouverts de \mathbb{R}^2 ? Pour cela on verra qu'il suffit de faire intervenir un paramétrage local φ , ramenant l'étude et les attributs de f au attributs de $f \circ \varphi$, définie sur un ouvert de \mathbb{R}^2 . La notion d'équivalence entre nappes montrera que les attributs alors accordés à f par la médiation du paramétrage sont indépendants du paramétrage équivalent utilisé.

Définition V.1.13 (Caractère lisse). *Une application $f : S \rightarrow \mathbb{R}^d$ est lisse si et seulement si $\forall P \in S$ et pour toute nappe paramétrée locale $\varphi : \mathcal{U} \rightarrow \mathbb{R}^3$ (dont l'image contient P) $f \circ \varphi : \mathcal{U} \rightarrow \mathbb{R}^d$ est lisse sur \mathcal{U} .*

Remarque V.1.14. Si cette propriété est vérifiée pour une nappe paramétrée locale alors elle est vérifiée pour toute autre nappe paramétrée locale équivalente. En effet considérons $(\varphi, \mathcal{U}) \sim (\psi, \mathcal{V})$ tel que $f \circ \varphi$ soit lisse. En effet posons le difféomorphisme $\rho = \psi^{-1} \circ \varphi : \mathcal{U} \rightarrow \mathcal{V}$ alors $f \circ \psi = f \circ \psi \circ I_{\mathcal{V}} = f \circ \psi \circ \rho \circ \rho^{-1} = f \circ \psi \circ \psi^{-1} \circ \varphi \circ \rho^{-1} = f \circ \varphi \circ \rho^{-1}$: il est lisse par composition.

Les applications lisses sont naturellement dérivables mais sur quel espace de départ s'applique la différentielle ? Si l'on considère deux points M et $M + dM$ infiniment voisins sur une surface on peut alors supposer qu'ils sont atteints tous deux par le même paramétrage local φ (on a $\varphi(m) = M$ et $M + dM = \varphi(m + h)$ où $h \in \mathbb{R}^2$), alors $M M' = d\varphi(h) + \|h\|^2 \mathbf{f}(h)$ (où \mathbf{f} est borné au voisinage de 0) qui appartient au plan tangent à $\|h\|^2 \mathbf{f}(h)$ près. Cela nous conduit à vouloir que la différentielle d'une fonction définie sur une surface s'applique sur des vecteurs de l'espace tangent.

Définition V.1.15 (Différentiel sur une surface). *Soit S une surface régulière et une application $f : S \rightarrow \mathbb{R}^d$ qui est lisse. Alors sa différentielle en $P \in S$ est la différentielle de l'application f composée avec une nappe paramétrée locale $\varphi : \mathcal{U} \rightarrow \mathbb{R}^3$ (dont l'image contient P). Si on note $P = \varphi_p$ alors on a*

$$\begin{aligned} df_P &: \mathcal{T}_P(S) = \text{im } d\varphi_p \rightarrow \mathbb{R}^d \\ x &\mapsto d(f \circ \varphi)_p((d\varphi_p)^{-1}(x)) \end{aligned}$$

En effet $d\varphi_p$ est un isomorphisme entre \mathbb{R}^2 et la direction du plan tangent en P . Si on note u, v les coordonnées sur \mathcal{U} alors on note du, dv tel que $d\varphi_p)^{-1}(x) = (du(x), dv(x))$ les formes coordonnées sur l'espace tangent¹. On a alors

$$\begin{aligned} df_P &: \mathcal{T}_P(S) = \text{im } d\varphi_p \rightarrow \mathbb{R}^d \\ x &\mapsto \partial_u(f \circ \varphi)_p du(x) + \partial_v(f \circ \varphi)_p dv(x) \end{aligned}$$

Remarque V.1.16. Cette définition est cohérente, et sa valeur est indépendante du paramétrage choisi. En effet considérons $(\varphi, \mathcal{U}) \sim (\psi, \mathcal{V})$ deux nappes paramétrées équivalentes contenant $P = \varphi(m) = \psi(n)$ tel qu'on ai $\forall x \in \mathcal{T}_P(S)$, l'équation $Df_p(x) = D(f \circ \varphi)_m((d\varphi_m)^{-1}(x))$, alors $D(f \circ \psi)_n((d\psi_n)^{-1}(x)) = D(f \circ \psi)_n \circ (d\psi_n)^{-1}(d\varphi_m((d\varphi_m)^{-1}(x)))$, on définit comme précédemment le difféomorphisme $\rho = \psi^{-1} \circ \varphi : \mathcal{U} \rightarrow \mathcal{V}$ et sa réciproque $\rho^{-1} = \varphi \circ \psi^{-1} : \mathcal{V} \rightarrow \mathcal{U}$ et puisque qu'il s'agit de la composition de deux applications définies sur des ouverts on a

$$\begin{aligned} D(f \circ \psi \circ \rho \circ \rho^{-1})_n &= D(f \circ \varphi \circ \rho^{-1})_n = D(f \circ \varphi)_{\rho^{-1}(n)} \circ D\rho_n^{-1} \\ &= D(f \circ \varphi)_m \circ D\rho_n^{-1} \\ &= D(f \circ \varphi)_m \circ D(\varphi^{-1} \circ \psi)_n \end{aligned} .$$

Et donc

$$D(f \circ \psi)_n((d\psi_n)^{-1}(x)) = D(f \circ \varphi)_m \circ D(\varphi^{-1} \circ \psi)_n \circ (d\psi_n)^{-1} \circ d\varphi_m((d\varphi_m)^{-1}(x))$$

ce qui implique la cohérence de la définition si et seulement si

$$D(\varphi^{-1} \circ \psi)_n \circ (d\psi_n)^{-1} \circ d\varphi_m = I_{\mathcal{U}}$$

ce qui est équivalent à $D(\varphi^{-1} \circ \psi)_n \circ (d\psi_n)^{-1} = (d\varphi_m)^{-1}$. Or φ est un difféomorphisme entre un ouvert de \mathbb{R}^2 et un ouvert relatif de S contenant P , il est cohérent de définir sur le plan tangent

1. Ce sont les formes coordonnées, on vérifie que pour tout $a, b \in \mathbb{R}^2$ on a $a\partial_u \varphi(P) + b\partial_b \varphi(P) \in \mathcal{T}_P(S)$ et $du(a\partial_u \varphi(P) + b\partial_b \varphi(P)) = a$ et $dv(a\partial_u \varphi(P) + b\partial_b \varphi(P)) = b$

$d\varphi_P^{-1} = (d\varphi_m)^{-1}$ et donc $(d\varphi_m)^{-1} \circ d\psi_n = d\varphi_{P=\psi(n)}^{-1} \circ d\psi_n = D(\varphi^{-1} \circ \psi)_n$ et donc notre définition est cohérente.

On montre en particulier que pour $F : \mathbb{R}^3 \rightarrow R^d$ la définition implique que la restriction à S de $f = F|_S$ a pour différentielle

$$Df_P = D(F|_S)_P = (DF_P)|_{T_P(S)}$$

Proposition V.1.17. Pour toute application $f : S \rightarrow R^d$ lisse, $\forall \mathbf{V} \in T_P(S)$ il existe $\epsilon > 0$ et $\gamma :]-\epsilon, \epsilon[\rightarrow S$ tels que $\gamma(0) = P$ et $\dot{\gamma}(0) = \mathbf{V}$ et alors on a

$$df_P(v) = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} = \lim_{t \rightarrow 0} \frac{f(\gamma(t)) - f(P)}{t}$$

V.1.7 Champs de vecteurs et fibré tangent

Les vecteurs de la base naturelle sont des champs de vecteurs qui appartiennent en tout point à l'espace tangent.

Définition V.1.18 (Champs de vecteurs tangent sur une surface régulière). *Soit \mathcal{S} une surface régulière, alors le champ de vecteurs $V : \mathcal{S} \rightarrow \mathbb{R}^3$ lisse est dit champs de vecteur tangent si et seulement si $\forall P \in \mathcal{S}, V(P) \in T_P(\mathcal{S})$.*

Ainsi on montre rapidement que les champs de vecteurs tangents sont définis localement par la donnée de leurs composantes (lisses) sur la base naturelle.

Définition V.1.19 (Fibré tangent). *On appelle fibré tangent d'un surface régulière \mathcal{S} , qu'on note $T(\mathcal{S})$ le sous espace de \mathbb{R}^6*

$$T(\mathcal{S}) = \{(P, T) \in \mathcal{S} \times \mathbb{R}^3 \mid T - P \in T_P(\mathcal{S})\}$$

ou ce qui est équivalent

$$T(\mathcal{S}) = \{(P, \mathbf{v}) \mid P \in \mathcal{S}, \mathbf{v} \in T_P(\mathcal{S})\}$$

Il vient immédiatement que $T_P(\mathcal{S}) = P + \text{vect}(\partial_u \varphi(m), \partial_v \varphi(m))$, ce qui justifie que $(\partial_u \varphi(m), \partial_v \varphi(m))$ soit qualifiée de *base naturelle* de l'espace tangent.

V.2 Première et seconde formes fondamentales

Comment faire des mesures de distance d'un arc, d'aire d'une sous-partie, comment se situe une surface vis-à-vis de son plan tangent, comment varie le plan tangent quand on parcourt la surface, quelle est la "courbure de Gauss" d'une surface ? C'est à toutes ces questions que répondra cette section.

V.2.1 Définitions première forme fondamentale

Le produit scalaire euclidien de \mathbb{R}^3 induit naturellement un produit scalaire entre vecteurs du même plan tangent à S , surface régulière.

Définition V.2.1 (Métrique ou première forme fondamentale). *Soit $S \subset \mathbb{R}^3$ une surface régulière. Alors $\forall P \in S$, on note*

$$\begin{aligned} g_P = I_P &: T_P(S) \times T_P(S) &\rightarrow \mathbb{R} \\ \mathbf{U}, \mathbf{V} &&\mapsto \mathbf{U} \cdot \mathbf{V} \end{aligned}$$

la forme bilinéaire symétrique qu'on nommera première forme fondamentale (en mathématiques) où métrique (en physique).

Remarque V.2.2. La première forme fondamentale reflète la façon dont S hérite de la structure Euclidienne de \mathbb{R}^3 . On conçoit bien qu'avec cet outil, on sera à même de calculer des longueurs d'arc de courbes régulières tracées sur la surface régulière, ainsi que des angles entre vecteurs de l'espace tangent mais aussi l'aire de domaines de S .

Définition V.2.3 (Produit tensoriel). *Soit f et g deux formes sur un \mathbb{R} -espace vectoriel E , alors on appelle $f \otimes g$ la forme bilinéaire, définie $\forall x, y \in E$ par $f \otimes g(x, y) = f(x)g(y)$*

Définition V.2.4 (Expression dans une paramétrisation, éléments de lignes de la première forme). Soit S une surface régulière et $\varphi : \mathcal{U} \rightarrow O \cap S$ une nappe paramétrée locale de nappe géométrique de S , alors si on note du : $\mathcal{T}_P(S) \rightarrow \mathbb{R}$ et $dv : \mathcal{T}_P(S) \rightarrow \mathbb{R}$ les formes linéaires correspondant à la base duale de $(\partial_u \varphi, \partial_v \varphi)$ ($\forall P \in O \cap S$) alors on a

$$\begin{aligned} I_p(.,.) &= (\partial_u \varphi \cdot \partial_u \varphi) du \otimes du + 2(\partial_u \varphi \cdot \partial_v \varphi) du \otimes dv + (\partial_v \varphi \cdot \partial_v \varphi) dv \otimes dv \\ &= E(u, v) du \otimes du + 2F(u, v) du \otimes dv + G(u, v) dv \otimes dv \end{aligned}$$

On appelle alors éléments de ligne associés à φ de la première forme fondamentale les fonctions E, F, G définies sur \mathcal{U}

Remarque V.2.5. Les éléments de ligne reflètent la façon dont S et le paramétrage φ héritent de la structure euclidienne de \mathbb{R}^3 . On va voir qu'enormément de choses peuvent être calculées à l'aide des éléments de ligne, on dit qu'il contiennent la géométrie *intrinsèque* de la surface.

Par définition les éléments de ligne de la première forme sont les coefficients de la matrice de Gramm des vecteurs $(\partial_u \varphi, \partial_v \varphi)$. En effet on a,

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} = \begin{pmatrix} \partial_u \varphi \\ \partial_v \varphi \end{pmatrix} \cdot (\partial_u \varphi \partial_v \varphi) = \text{Gram}(\partial_u \varphi, \partial_v \varphi).$$

On rappelle que le déterminant de la matrice $\text{Gram}(\partial_u \varphi, \partial_v \varphi)$ est le carré de la surface du parallélogramme engendré par les deux vecteurs. C'est un exercice classique d'algèbre que nous laisserons le lecteur motivé vérifier.

On va pouvoir mesurer la longueur d'arc de courbe régulière tracée sur la surface régulière. En effet considérons, $\Gamma : I \rightarrow \mathbb{R}^3$ une courbe paramétrée incluse dans une surface S , et O un voisinage ouvert de $\Gamma(t_0 \in I)$ ouvert de \mathbb{R}^3 tel que $O \cap S$ soit la nappe géométrique de nappe paramétrée associée $\varphi : \mathcal{U} \rightarrow O \cap S$, alors localement la portion $O \cap \Gamma$ d'une courbe paramétrée Γ incluse dans la surface est représentée par une courbe plane de \mathcal{U} , que nous notons $\gamma : I' \subset I \rightarrow \mathcal{U}$ $t \mapsto (u(t), v(t))$, et donc la longueur de l'arc est :

$$\ell(\Gamma \cap O) = \int_{I'} \left\| \frac{d\varphi \circ \gamma}{dt} \right\| dt = \int_{I'} \sqrt{E(\gamma(t)) \left(\frac{du}{dt} \right)^2 + 2F(\gamma(t)) \frac{du}{dt} \frac{dv}{dt} + G(\gamma(t)) \left(\frac{dv}{dt} \right)^2} dt$$

Cela permet de définir la longueur intrinsèque entre deux points.

Définition V.2.6. On suppose que notre surface régulière est une nappe géométrique c'est-à-dire qu'elle peut être représentée à l'aide d'une seule nappe paramétrée. Dans ce cas là, la distance entre tout point P, Q dans la surface peut être définie ainsi

$$d_S(P, Q) = \inf \{ \ell(\Gamma) | \gamma : [0, 1] \rightarrow \mathcal{U} \text{ tel que } \gamma(0) = \varphi^{-1}(P) \text{ et } \gamma(1) = \varphi^{-1}(Q) \}$$

Remarque V.2.7. La courbe paramétrée Γ minimisant la longueur est appelée *géodésique*, une caractérisation en termes d'équation différentielle sera développé ultérieurement.

Définition V.2.8 (Isométrie entre surface). Soit S_1, S_2 deux surfaces régulières. On dit qu'elles sont isométriques si et seulement si il existe un difféomorphisme $f : S_1 \rightarrow S_2$ qui préserve les distances intrinsèques (on dit que f est une isométrie).

On dit qu'elles sont localement isométriques si tout point de S_1 admet un voisinage relatif ouvert isométrique à un sous-ensemble de S_2 (et réciproquement).

Nous renvoyons le lecteur curieux à la thèse de AUDOLY 2000, s'il souhaite en savoir plus sur une notion plus pratique de l'isométrie : l'isométrie infinitésimale.

Proposition V.2.9. Deux surfaces sont localement isométriques si et seulement si pour tout point de $P_1 \in S_1$ il existe deux voisinage relatif ouvert $\tilde{S}_1 \subset S_1$ et $\tilde{S}_2 \subset S_2$ deux nappes paramétrées $\varphi : \mathcal{U} \rightarrow \tilde{S}_1$ et $\psi : \mathcal{U} \rightarrow \tilde{S}_2$, tels que les éléments de ligne de la première forme dans ces paramétrisations sont égaux.

Exercice 27. — Élément de lignes.

1. Calculer les éléments de ligne d'un plan dans un paramétrage cartésien, cylindrique.

2. Calculer les éléments de ligne d'un Cylindre droit.

Définition V.2.10 (Isométrie entre surfaces). Soient S_1, S_2 deux surfaces régulières. On dit qu'elles sont isométriques si et seulement si il existe un difféomorphisme $f : S_1 \rightarrow S_2$ qui préserve les distances intrinsèques (on dit que f est une isométrie).

On dit qu'elles sont localement isométriques si tout point de S_1 admet un voisinage relatif ouvert isométrique à un sous ensemble de S_2 (et réciproquement).

Proposition V.2.11. Deux surfaces sont localement isométriques si et seulement si pour tout point de $P_1 \in S_1$ il existe deux voisinages relatifs ouverts $\tilde{S}_1 \subset S_1$ et $\tilde{S}_2 \subset S_2$ deux nappes paramétrées $\varphi : \mathcal{U} \rightarrow \tilde{S}_1$ et $\psi : \mathcal{U} \rightarrow \tilde{S}_2$, tels que les éléments de ligne associés à φ de la première forme fondamentale soient égaux aux éléments de ligne de la première forme fondamentale associé à ψ .

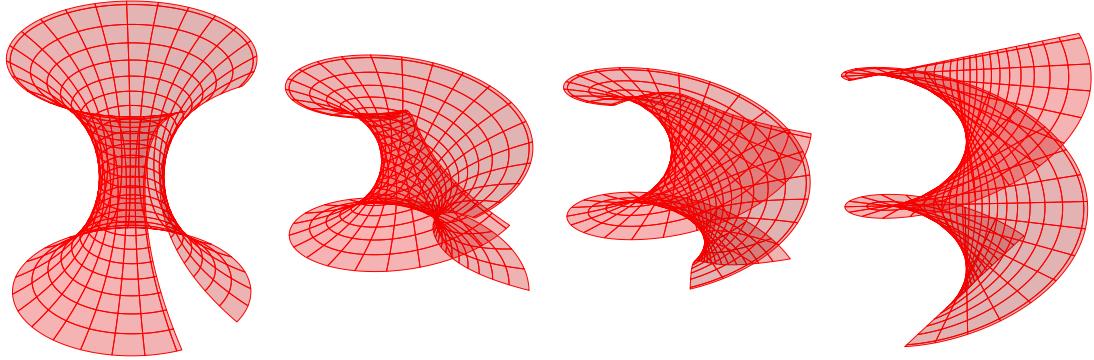


FIGURE V.11 – Transformation isométrique de la Caténoïde en Hélicoïde

Exercice 28. — Caténoïde et Hélicoïde.

On considère $\forall \alpha \in [0, \frac{\pi}{2}]$, la famille d'application

$$\begin{aligned} \psi_\alpha & :]-u_0, u_0[\times]\epsilon, 2\pi - \epsilon[\longrightarrow \mathbb{R}^3 \\ (u, v) & \mapsto \begin{cases} x_\alpha(u, v) = \cos \alpha \cos v \cosh u + \sin \alpha \sin v \sinh u \\ y_\alpha(u, v) = \cos \alpha \sin v \cosh u + \sin \alpha \cos v \sinh u \\ z_\alpha(u, v) = \cos \alpha u + \sin \alpha v \end{cases} \end{aligned}$$

1. On admet que pour ϵ suffisamment élevé ψ_α est injective, montrer que la nappe géométrique associée est régulière.
2. Calculer les éléments de ligne de la première forme de la surface $S_\alpha = \psi_\alpha (]-u_0, u_0[\times]\epsilon, 2\pi - \epsilon[)$ en déduire que $\forall \alpha_1, \alpha_2 \in [0, \frac{\pi}{2}]$, S_{α_1} et S_{α_2} sont isométriques.

On peut également définir un angle non orienté entre deux courbes incluses dans la surface.

Définition V.2.12 (Angles). Soit deux courbes paramétrées $\alpha : I \rightarrow$ et $\beta : I \rightarrow$ incluses dans une surface S , qui s'intersectent en $P = \alpha(t_0) = \beta(t_0)$ pour $t = t_0$, alors $\frac{d\alpha}{dt}(t_0)$ et $\frac{d\beta}{dt}(t_0) \in \mathcal{T}_P(S)$. Alors on définit l'angle $\theta \in [0, \pi]$ par

$$\cos \theta = \frac{I_P(\frac{d\alpha}{dt}(t_0), \frac{d\beta}{dt}(t_0))}{\sqrt{I_P(\frac{d\alpha}{dt}(t_0), \frac{d\alpha}{dt}(t_0)) I_P(\frac{d\beta}{dt}(t_0), \frac{d\beta}{dt}(t_0))}}$$

Définition V.2.13 (Difféomorphisme conforme). Soit S_1 une surface régulière et $f : S_1 \rightarrow f(S_1) = S_2$ un difféomorphisme, ce difféomorphisme est dit conforme si et seulement si pour toute nappe paramétrée locale φ de S_1 , les éléments de ligne de la première forme (E, F et G) de φ sont

proportionnels à ceux E_f, F_f et G_f de $f \circ \varphi$, c'est-à-dire, qu'il existe Ω^2 , positive, lisse et qui ne s'annule pas (définie sur le même ouvert que φ) tel que

$$\begin{cases} E_f = \Omega E \\ F_f = \Omega F \\ G_f = \Omega G \end{cases}$$

Ω est appelé le facteur conforme. L'angle de deux courbes incluses dans S_1 qui s'intersectent en P_1 est le même que celui des deux courbes transportées (et donc incluses dans $f(S_1)$) qui s'intersectent en $f(P_1)$.

Le facteur conforme s'interprète comme une facteur de proportionnalité local permettant de passer de la longueur d'un segment élémentaire à la longueur de ce segment transformé par f .

Exercice 29. — Transformation de Mercator.

On peut paramétriser la sphère, privée d'un arc et des pôles par S_1

$$\varphi : \begin{cases} x(\theta, \phi) = \cos \theta \cos \phi \\ y(\theta, \phi) = \cos \theta \sin \phi \\ z(\theta, \phi) = \sin \theta \end{cases}$$

On pose f l'application de Mercator qui va de la sphère dans le cylindre de révolution droit et de rayon 1 et d'axe \mathbf{u}_z et qui s'exprime comme suit si $\psi = f \circ \varphi : (\theta, \phi) \mapsto \lambda(\theta)\mathbf{u}_z + \mathbf{u}(\phi)$, où $\lambda :]-\pi/2, \pi/2[\rightarrow \mathbb{R}$ est une application qu'on déterminera et $\mathbf{u}(\phi) = \cos \phi \mathbf{u}_x + \sin \phi \mathbf{u}_y$.

1. Sur quel ensemble de définition maximal peut-on définir φ afin que l'application reste injective ?
2. Montrer que φ est une nappe paramétrée régulière.
3. Quelle est la valeur de l'application λ permettant à la transformation de Mercator d'être une transformation conforme ? Commenter.

Correction :

1. Définie sur $] -\pi/2, \pi/2[\times]0, 2\pi[$, φ est injective.
2. φ est lisse, il suffit donc de vérifier que $\partial_\theta \varphi$ et $\partial_\phi \varphi$ sont libres. On a $\partial_\theta \varphi = (\cos \theta \cos \phi, \cos \theta \sin \phi, \sin \theta)$ et $\partial_\phi \varphi = (-\sin \theta \sin \phi, \sin \theta \cos \phi, 0)$ qui sont libres.
3. Les éléments de ligne associées à φ de la première forme fondamentale sont $E(\theta, \phi) = 1$, $F(\theta, \phi) = 0$ et $G(\theta, \phi) = \cos^2 \theta$. On a donc $E_f(\theta, \phi) = \left(\frac{d\lambda}{d\theta}\right)^2$, $F_f(\theta, \phi) = 0$ et encore $G_f(\theta, \phi) = 1$. Puisque $G_f(\theta, \phi) = \frac{G(\theta, \phi)}{\cos^2 \theta}$, pour que la transformation soit conforme il faut que $\frac{d\lambda}{d\theta} = \frac{1}{\cos \theta}$ et donc on pose $\lambda(\theta) = \int_0^\theta \frac{d\tilde{\theta}}{\cos \tilde{\theta}} = \operatorname{arctanh} \sin \theta$, qui est bien définie sur $] -\pi/2, \pi/2[$. Contrairement à ce qui est souvent affirmé, la projection de Mercator, n'est pas la projection directe sur le cylindre droit d'axe \mathbf{u}_z de la sphère car dans ce cas, on aurait $\lambda(\theta) = \frac{1}{\sin \theta}$ et la transformation ne serait pas conforme. Le facteur de conformité de cette transformation est égal à $\Omega(\theta, \phi) = \frac{1}{\cos \theta}$, les distances sont donc assez bien représentées au niveau de $\theta = 0$ ce qui correspond à l'équateur.

Définissons maintenant, si elle existe, l'aire d'une portion de surface.

Théorème (et définition) V.2.14 (Aire). Soit $\Omega \subset S$ (où S est une surface régulière) tel que Ω soit inclus dans l'image d'une paramétrisation régulière $\varphi : \mathcal{U} \rightarrow S$, alors si elle est définie, l'aire de la portion Ω est :

$$\text{Aire}(\Omega) = \iint_{\phi^{-1}(\Omega)} \|\partial_u \varphi \wedge \partial_v \varphi\| du dv \iint_{\phi^{-1}(\Omega)} \sqrt{E(u, v)G(u, v) - F^2(u, v)} du dv.$$

cette quantité est indépendante du choix de paramétrage.

Preuve. Montrons que cette définition est indépendante du choix du paramétrage, prenons donc (φ, \mathcal{U}) et (ψ, \mathcal{V}) nappes équivalentes et $f : \mathcal{V} \rightarrow \mathcal{U}$ le difféomorphisme entre les ouverts de \mathbb{R}^2 . On note u, v les coordonnées cartésiennes sur \mathcal{U} et s, t les coordonnées sur \mathcal{V} , on note $m \in \mathcal{U}$, $n = f^{-1}(m) \in \mathcal{V}$ et $f(n) = (f_s(n), f_t(n))$, on a donc

$$\varphi : \begin{cases} \partial_s \psi(n) = \partial_s f_u(n) \partial_u \varphi(m) + \partial_s f_v(n) \partial_v \varphi(m) \\ \partial_t \psi(n) = \partial_t f_u(n) \partial_u \varphi(m) + \partial_t f_v(n) \partial_v \varphi(m) \end{cases}$$

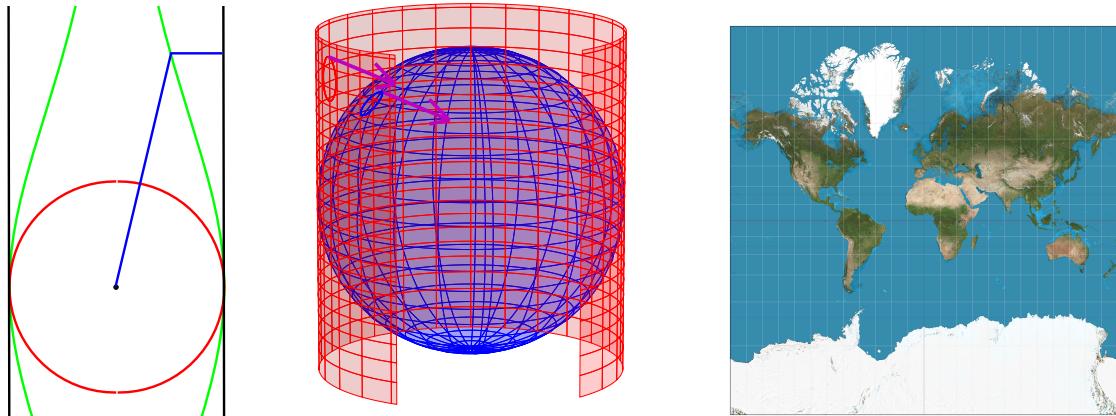


FIGURE V.12 – À gauche schématisation de la transformation de Mercator, on commence par projeter la sphère sur la surface de révolution de nappe paramétrée $(\theta, \phi) \mapsto \lambda(\theta)\mathbf{u}_z + \lambda(\theta) \tan \theta \mathbf{u}(\phi)$ (vert), puis on projette sur le cylindre droit en noir. Au centre, la transformation de Mercator en rouge d'un petit chemin inclus dans la sphère en bleu. À droite, représentation de la carte terrestre dans la transformation de Mercator.

De sorte qu'on a $\|\partial_s \psi(n) \wedge \partial_t \psi(n)\| = |\det df_n| \|\partial_u \varphi(m) \wedge \partial_v \varphi(m)\|$. On note alors $\tilde{U} = \psi^{-1}(\Omega)$ et $\tilde{V} = f^{-1}(\tilde{U})$ et on applique la formule du changement de variable

$$\begin{aligned} \iint_{\phi^{-1}(\Omega)} \|\partial_u \varphi \wedge \partial_v \varphi\| dudv &= \iint_{\tilde{U}=f(\tilde{V})} \|\partial_u \varphi(u, v) \wedge \partial_v \varphi(u, v)\| dudv \\ &= \iint_{\tilde{V}} \|\partial_u \varphi(f(s, t)) \wedge \partial_v \varphi(f(s, t))\| |\det df(s, t)| dt ds \\ &= \iint_{\tilde{V}} \|\partial_s \psi(s, t) \wedge \partial_t \psi(s, t)\| dt ds \end{aligned}$$

■

Remarque V.2.15. Pour des $\Omega \subset S$ qui ne peuvent être inclus dans une nappe géométrique régulière, on comprend qu'il va falloir découper Ω en morceau qui le sont. Si on a $\Omega = \bigcup_{i \in \mathcal{I}} \Omega_i$ tel que $\forall i \neq j \in \mathcal{I}$ on a $\Omega_i \cup \Omega_j = \Omega$ et Ω_i est inclus dans une nappe géométrique régulière. Alors on pose

$$\text{Aire}(\Omega) \doteq \sum_{i \in \mathcal{I}} \text{Aire}(\Omega_i)$$

Exercice 30. — Aire du tore.

Montrer que l'aire du tore de grand cercle de rayon a et de petit cercle b est $4\pi^2 ab$

V.2.2 Du problème du plongement

Considérons que les éléments de ligne E, F et G soient donnés et vérifient $F^2(u, v) < E(u, v)G(u, v)$ (inégalités de Schwarz) sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$. Existe-t-il une nappe paramétrée $\varphi : \mathcal{U} \rightarrow \mathbb{R}^3$ telle que E, F et G correspondent au éléments de ligne de la première forme fondamentale attachée à φ ? C'est ce qu'on appelle le *problème du plongement isométrique*. Si on note $\varphi(u, v) = O + x(u, v)\mathbf{e}_x + y(u, v)\mathbf{e}_y + z(u, v)\mathbf{e}_z$ (où $(O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ est un repère orthonormal quelconque), il s'agit de trouver une solution au système non linéaire d'équations aux dérivées partielles :

$$\begin{cases} (\partial_u x)^2 + (\partial_u y)^2 + (\partial_u z)^2 = E(u, v) \\ (\partial_u x \cdot \partial_v x) + (\partial_u y \cdot \partial_v y) + (\partial_u z \cdot \partial_v z) = F(u, v) \\ (\partial_v x)^2 + (\partial_v y)^2 + (\partial_v z)^2 = G(u, v) \end{cases}$$

Si la coordonnée v est ignorable ($\partial_v E = 0$ et $\partial_v G = 0$) et $F = 0$, et que l'ouvert en question est du type $I \times]0, 2\pi[$, alors on renomme $v = \phi$ et on peut chercher le plongement sous la forme d'une nappe paramétrée de révolution $\varphi(u, \phi) = \rho(u)u(\phi) + z(u)\mathbf{u}_z$. Alors le système à résoudre se

simplifie en :

$$\begin{cases} \left(\frac{d\rho}{du}\right)^2 + \left(\frac{dz}{du}\right)^2 = E(u) \\ \rho^2(u) = G(u) \end{cases}$$

Exercice 31. — Horizon d'un trou noir.

À un facteur multiplicatif près, la section $t = \text{Cst}$ d'horizon d'un trou noir a pour éléments de ligne

$$I_P = [1 - \beta^2 \sin^2 \theta] d\theta \otimes d\theta + \frac{\sin^2 \theta}{1 - \beta^2 \sin^2 \theta} d\phi \otimes d\phi$$

où le $\beta^2 = \frac{a^2}{1+\sqrt{1-a^2}}$ et a est le spin adimensionné du trou noir. Montrer qu'un plongement avec une nappe paramétrée de révolution existe pour $\beta^2 \geq 1/4$ ($a \geq \sqrt{7}/4$) (ce résultat fut établi par Smarr en 1978).

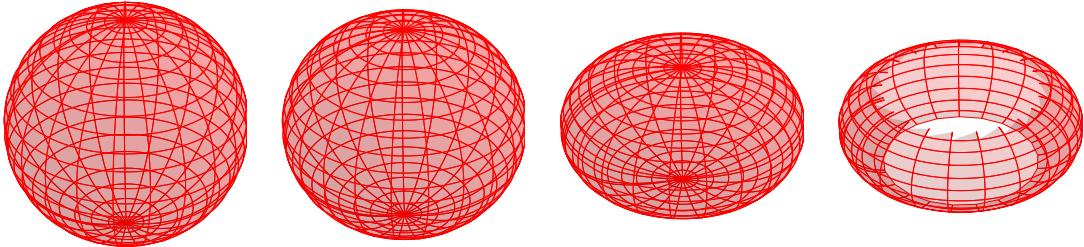


FIGURE V.13 – Évolution du plongement dans \mathbb{R}^3 d'une section $t = \text{Cst}$ d'horizon d'un trou noir à mesure que le spin $a = 0, \frac{\sqrt{7}}{8}, \frac{\sqrt{7}}{4}, 1.05 \frac{\sqrt{7}}{8}$ grandit, on obtient la valeur limite pour la troisième figure en partant de la gauche dans ce cas les pôles sont des points planaires (voir plus loin). Pour des valeurs supérieures seules des parties de l'horizon sont plongeables.

Remarque V.2.16. Le mathématicien J. Nash prouva en 1954 et 1956 à l'aide de deux théorèmes magnifiques l'existence du plongement dans des espaces de dimension supérieure. Par exemple pour les surfaces, les équations du plongement admettent des solutions C^3 dans \mathbb{R}^2 . Néanmoins le problème du plongement est souvent difficile en pratique. Gromov, mathématicien Russe introduira la théorie de l'intégration convexe ce qui aide en pratique à résoudre de tels problèmes. Néanmoins les surfaces obtenues sont d'une régularité comprise entre C^1 et C^2 : le plan tangent existe en tout point mais il est discontinu en tout point ! Mentionnons le plongement du tore plat (surface à la topologie du tore), mais dont les éléments de lignes de la première forme fondamentale sont égaux à ceux du plan euclidien : $E = 1, F = 0, G = 1$. Une illustration de cet objet se situe sur la couverture de ce cours.

V.2.3 Application de Gauss et seconde forme fondamentale

Jusqu'ici nous avons traité de caractéristiques intrinsèques construites à partir de la première forme fondamentale contenant des informations sur la différentielle des paramétrages. Nous allons maintenant nous intéresser à la façon dont varie le plan tangent. Il s'agit d'informations sur les dérivées secondes des paramétrages.

On considérera donc de façon générique $S \in \mathbb{R}^3$ une surface régulière, $P \in S$ et $\varphi : \mathcal{U} \rightarrow S_P$ une nappe paramétrée locale dont la nappe géométrique S_P contient P . On peut ainsi construire la base naturelle $\partial_u \varphi$ et $\partial_v \varphi$ et également

Définition V.2.17 (Application de Gauss). *L'application,*

$$\begin{aligned} \mathbf{n} &: S_P &\longrightarrow & S_1 \\ M &\mapsto & \frac{\partial_u \varphi(\varphi^{-1}(M)) \wedge \partial_v \varphi(\varphi^{-1}(M))}{\|\partial_u \varphi(\varphi^{-1}(M)) \wedge \partial_v \varphi(\varphi^{-1}(M))\|} \end{aligned}$$

est un vecteur unitaire génératrice de l'espace normal (S_1 est la sphère unité). Par construction, cette application est lisse sur S_P . L'application $\mathbf{n} : S_P \rightarrow S_1$ qui à tout point associe un vecteur normal unitaire s'appelle application de Gauss associée au paramétrage φ .

Remarque V.2.18. Cette application dépend pour l'instant du paramétrage et ne peut en particulier être étendue de façon continue sur toute la surface, on peut la construire de manière à être lisse sur des voisinages ouverts relatifs de tout point P de S . Nous verrons plus loin que nous pouvons en donner une définition invariante par changement de paramétrage et lisse sur la surface tout entière dans le cas des surfaces dites orientables.

Exercice 32. —

Calculer l'application de Gauss sur un Hélicoïde.

Lemme V.2.19. Pour tout $\mathbf{V} \in \mathcal{T}_P(S)$, la dérivée directionnelle le long de \mathbf{v} de l'application de Gauss, $D_{\mathbf{V}}\mathbf{n}(P) = (\mathbf{d}\mathbf{n})_P(\mathbf{V})$, est un vecteur tangent à S en P .

Preuve. On note $\mathbf{V} = V^u \partial_u \varphi + V^v \partial_v \varphi$, on pose alors $m = \varphi^{-1}(P)$ et on choisit $\epsilon > 0$ tel que $\alpha : t \in]-\epsilon, \epsilon[\mapsto m + (tv^u, tv^v) \in \mathcal{U}$. Alors la courbe incluse $\gamma : t \mapsto \varphi(\alpha(t))$ vérifie $\gamma(0) = P$ et $\frac{d\gamma}{dt}(0) = \mathbf{V}$ et donc puisque $t \mapsto \|\mathbf{n}(\gamma(t))\|^2 = 1$

$$0 = \frac{d\|\mathbf{n} \circ \gamma\|^2}{dt}(0) = 2I_P(\mathbf{n} \circ \gamma(0), \frac{d\mathbf{n} \circ \gamma}{dt}(0)) = 2I_P(\mathbf{n}(P), (\mathbf{d}\mathbf{n})_P(\mathbf{V}))$$

Ce qui montre que $d\mathbf{n}_P(\mathbf{V}) \in \mathcal{T}_P(S)$. ■

On définit ainsi

Définition V.2.20 (Opérateur de forme ou application de Weingarten). *L'application suivante est l'opérateur appelé opérateur de forme ou application de Weingarten.*

$$\begin{aligned} \mathbf{F}_P &: \mathcal{T}_P(S) &\rightarrow \mathcal{T}_P(S) \\ \mathbf{V} &\mapsto -(\mathbf{d}\mathbf{n})_P(\mathbf{V}) \end{aligned}$$

qui est linéaire par construction.

Cette application est symétrique en effet,

Proposition V.2.21. L'application \mathbf{F}_P est symétrique c'est-à-dire $\forall \mathbf{U}, \mathbf{V} \in \mathcal{T}_P(S)$, on a $I_P(\mathbf{F}_P(\mathbf{U}), \mathbf{V}) = I_P(\mathbf{F}_P(\mathbf{V}), \mathbf{U})$.

Preuve. Commençons par remarquer que $\mathbf{U}, \mathbf{V} \in \mathcal{T}_P(S) \mapsto I_P(\mathbf{F}_P(\mathbf{U}), \mathbf{V})$ est bilinéaire par construction. Il suffit donc de vérifier la propriété pour les vecteurs de la base naturelle. On pose $\mathbf{U} = \partial_u \varphi$ et $\mathbf{V} = \partial_v \varphi$ on a donc en dérivant $\mathbf{n} \cdot \mathbf{V} = 0$ par rapport à u :

$$\partial_u \mathbf{n} \cdot \mathbf{V} = -\mathbf{n} \cdot \partial_{uv}^2 \varphi$$

Puisque $\partial_u \mathbf{n} = (\mathbf{d}\mathbf{n})_P(\mathbf{U})$ et que le théorème de Schwarz s'applique à φ , cela implique que

$$I_P(\mathbf{F}_P(\mathbf{U}), \mathbf{V}) = \mathbf{n} \cdot \partial_{uv}^2 \varphi = \mathbf{n} \cdot \partial_{vu}^2 \varphi = -(\mathbf{d}\mathbf{n})_P(\mathbf{V}) \cdot \mathbf{U} = I_P(\mathbf{F}_P(\mathbf{V}), \mathbf{U})$$

La linéarité et la dimension 2 de l'espace tangent permettent de conclure. ■

Exercice 33. —

Montrer que si $\forall P \in S$, $\mathbf{F}_P = 0$ alors la surface est localement plane.

On peut enfin définir la seconde forme fondamentale

Définition V.2.22. La seconde forme fondamentale est la forme bilinéaire liée à l'opérateur de forme par

$$\begin{aligned} II_P &: \mathcal{T}_P(S) \times \mathcal{T}_P(S) &\rightarrow \mathbb{R} \\ (\mathbf{U}, \mathbf{V}) &\mapsto I_P(\mathbf{F}_P(\mathbf{U}), \mathbf{V}) \end{aligned}$$

Remarque V.2.23. En utilisant les éléments de la démonstration de la proposition précédente on montre que dans la base naturelle associée à un paramétrage, on a

$$II_P = \begin{pmatrix} n \cdot \partial_{uu}^2 \varphi & n \cdot \partial_{uv}^2 \varphi \\ n \cdot \partial_{vu}^2 \varphi & n \cdot \partial_{vv}^2 \varphi \end{pmatrix}$$

Cela explique également la raison du signe moins dans la définition de l'opérateur de forme.

L'application II_P est également caractéristique de la position de la surface vis-à-vis de son plan tangent. En effet posons $P = \varphi(m)$, alors pour $m + h \in \mathcal{U}$ on a

$$(\varphi(m)\varphi(m+h)) \cdot \mathbf{n}(P) = II_P(h, h) + \|h\|^3 \phi(h)$$

où $\phi(h)$ est borné au voisinage de 0. On a donc une surface localement "au dessus" (respectivement au dessous) de son plan tangents pour II_P définie positive (respectivement négative), ou en forme de point selle pour une seconde forme de signature $(+, -)$. On ne peut rien dire de plus si II_P est dégénérée.

V.2.4 Orientation

On dit qu'une surface régulière $S \subset \mathbb{R}^3$ est *orientable* quand on peut construire de façon continue l'application de Gauss sur toute la surface. Alors, il est possible de fixer de façon continue l'orientation des plans tangents. Pour que cela soit possible, il faut et il suffit que pour tout recouvrement de S par des nappes géométriques, les nappes paramétrées qui y sont associées soient compatibles. Qu'entend-t-on par là ?

Pour fixer les idées, prenons $\varphi : \mathcal{U} \rightarrow S$ une paramétrisation de S au voisinage d'un point P , on fixe alors l'orientation en postulant que la base $(\partial_u \varphi, \partial_v \varphi)$ soit directe. Si P est également voisin d'une paramétrisation $\psi : \mathcal{V} \rightarrow S$, alors la base $(\partial_u \psi, \partial_v \psi)$ est directe (respectivement indirecte) si le jacobien de $\varphi^{-1} \circ \psi$ est positif (respectivement négatif).

Définition V.2.24 (Orientabilité). *Une surface S est orientable si il est possible de la recouvrir par une famille de paramétrisation telle que les jacobiens des changements de paramétrisations qui se recoupent sont positifs. Le choix d'une telle famille est appelé une orientation de S . Lorsqu'un tel choix n'est pas possible, on dis que S est non orientable.*

Toute nappe géométrique est bien entendu orientable, ce qui implique que toute surface est localement orientable. Il s'agit donc d'un problème global : à l'échelle de la surface. La preuve de ce genre de propriétés est donc parfois assez complexe alors que ce qu'on cherche à vérifier est très visuel. Il est en général assez facile de prouver l'orientabilité d'une surface soit en utilisant les propriétés qui vont suivre soit en exhibant une orientation de la surface. La non-orientabilité est plus difficile à prouver.

Proposition V.2.25. Une surface régulière S est orientable si et seulement si $\exists \mathbf{N} : S \rightarrow \mathcal{S}_1$ application continue telle qu'en tout point $P + \mathbb{R}\mathbf{N}(P)$ est l'espace normal à l'espace tangent.

Preuve. Supposons S orientable, alors on la recouvre par des paramétrisations locales $(\varphi_i, \mathcal{U}_i)_{i \in I}$ telles que les jacobiens des changements de paramétrisations (là où ils existent) sont positif. On considère (φ, \mathcal{U}) et (ψ, \mathcal{V}) deux de ces paramétrisations telles que $\varphi(\mathcal{U}) \cap \psi(\mathcal{V}) \neq \emptyset$. Alors pour tout point $P \in \varphi(\mathcal{U}) \cap \psi(\mathcal{V})$ un rapide calcul montre que les applications de Gauss associées à ces paramétrisations sont pour tout point du recouvrement $\varphi(\mathcal{U}) \cap \psi(\mathcal{V})$,

$$\mathbf{n}_V(P) = \frac{\partial_u \psi \wedge \partial_v \psi}{\|\partial_u \psi \wedge \partial_v \psi\|} = \frac{\det \varphi^{-1} \circ \psi}{\|\det \varphi^{-1} \circ \psi\|} \frac{\partial_u \varphi \wedge \partial_v \varphi}{\|\partial_u \varphi \wedge \partial_v \varphi\|} = \mathbf{n}_U(P)$$

puisque les jacobiens de changements de paramétrisations sont positifs. Ainsi on peut définir une application $P \mapsto \mathbf{N}$ continue sur $\varphi(\mathcal{U}) \cup \psi(\mathcal{V})$, puisque les paramétrages recouvrent S , alors l'application \mathbf{N} peut être construite et étendue par continuité sur tout S .

Réciproquement, supposons qu'une telle application existe, alors pour chaque paramétrisation locale $(\varphi_i, \mathcal{U}_i)_{i \in I}$ qui recouvre S telle que les \mathcal{U}_i soient connexes, on pose à nouveau $\mathbf{n}_{\mathcal{U}_i}$ l'application de Gauss liée à cette paramétrisation. Alors on peut définir sur $\varphi_i(\mathcal{U}_i)$ l'application $P \mapsto \mathbf{N}(P) \cdot \mathbf{n}_{\mathcal{U}_i} \in \{-1, 1\}$ qui est continue sur \mathcal{U}_i connexe et donc constante sur \mathcal{U}_i . Quitte à remplacer φ_i par $\varphi_i \circ f$ avec $f : (x, y) \mapsto (x, -y)$, on peut supposer que pour tout i , $\mathbf{N}(P) \cdot \mathbf{n}_{\mathcal{U}_i} = 1$. Il s'ensuit que les jacobiens de changements de coordonnées sont tous positifs. En effet, en prenant dans la famille de paramétrisations deux d'entre elles qui se recouvrent et multipliant scalairement par \mathbf{N} l'équation

$$\mathbf{n}_V(P) = \frac{\partial_u \psi \wedge \partial_v \psi}{\|\partial_u \psi \wedge \partial_v \psi\|} = \frac{\det \varphi^{-1} \circ \psi}{\|\det \varphi^{-1} \circ \psi\|} \frac{\partial_u \varphi \wedge \partial_v \varphi}{\|\partial_u \varphi \wedge \partial_v \varphi\|} = \frac{\det \varphi^{-1} \circ \psi}{\|\det \varphi^{-1} \circ \psi\|} \mathbf{n}_U(P)$$

on aboutit au résultat escompté sur les Jacobiens. ■

Corrolaire V.2.26. Une surface régulière définie par submersion est orientable

Exercice 34. — Pour les Oufs.

Montrer que toute surface orientable compacte de \mathbb{R}^3 est une submersion. (Il faudra construire la submersion, ne pas oublier la propriété de Bolzano-Weierstrass, la preuve de ce théorème est trouvable dans CARMO 1976 page 109-114).

Remarque V.2.27. Toute nappe géométrique est orientable. Tout graphe est orientable.

Exercice 35. —

1. Monter que l'ellipsoïde est orientable.
2. Montrer que l'hélicoïde est orientable.

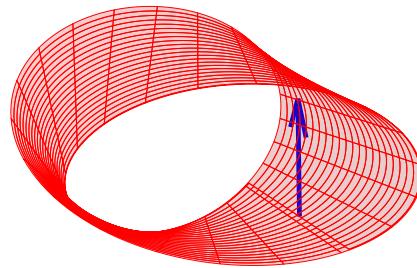


FIGURE V.14 – Représentation d'un ruban de Möbius

Exercice 36. — Möbius.

On pose $\mathbf{u} : \theta \mapsto \cos \theta \mathbf{e}_x + \sin \theta \mathbf{e}_y$ et $\mathbf{w} : \theta \mapsto \cos \frac{\theta}{2} \mathbf{u}(\theta) + \sin \frac{\theta}{2} \mathbf{e}_z$, alors l'application

$$\begin{aligned} \varphi &: \mathbb{R} \times]-1/2; 1/2[\rightarrow \mathbb{R}^3 \\ (\theta, s) &\mapsto O\mathbf{u}(\theta) + s\mathbf{w}(\theta) \end{aligned}$$

permet de définir le *Ruban de Möbius* $S = \text{im } \varphi$. On définit également $\forall \theta_0 \in \mathbb{R} \quad \mathcal{U}_{\theta_0} = [\theta_0 - \pi; \theta_0 + \pi] \times]-1/2; 1/2[$ et ψ_{θ_0} la restriction de φ sur \mathcal{U}_{θ_0} . Enfin on posera $\mathbf{N}_{\theta_0} : (\theta, s) \mathcal{U}_{\theta_0} \mapsto$

$$\begin{aligned} \mathbf{N}_{\theta_0} &: \mathcal{U}_{\theta_0} \rightarrow \mathcal{S}_1 \\ (\theta, s) &\mapsto \frac{\partial_u \psi_{\theta_0} \wedge \partial_v \psi_{\theta_0}}{\|\partial_u \psi_{\theta_0} \wedge \partial_v \psi_{\theta_0}\|} \end{aligned}$$

1. Monter que $\forall \theta_0 \in \mathbb{R}$, ψ_{θ_0} est injective.
2. Monter que $\forall \theta_0 \in \mathbb{R}$, ψ_{θ_0} est lisse et que sa différentielle est de rang 2, en déduire que S est une surface régulière.
3. On suppose que la surface est orientable et donc qu'il existe \mathbf{N} la normale continue sur S , montrer que l'application $\epsilon : \mathcal{U}_0 \rightarrow \{-1, 1\}$ définie par $\epsilon(\theta, s) = \mathbf{N} \circ \psi_0(\theta, s) \cdot \mathbf{N}_0(\theta, s)$ est constante sur \mathcal{U}_0 .
4. On pose $\gamma : \theta \in]0, 2\pi[\mapsto \varphi(\theta, 0)$. En étudiant la continuité en π de l'application continue par composition $N \circ \gamma$ montrer qu'il y a une contradiction : le ruban de Möbius n'est pas orientable.

V.3 Theorema Egregium de Gauss

Nous allons ici nous concentrer sur des quantités caractéristiques locales de la seconde forme fondamentale : son déterminant définira la *courbure de Gauss*, ses valeurs propres et vecteurs propres définiront les *directions et courbures principales*, la trace nous donnera la *courbure moyenne*. Le *théorème Egregium* de Gauss montre que la courbure de Gauss peut être exprimée uniquement à partir des éléments de ligne de la première forme de la surface, elle est donc une quantité propre à la "géométrie intrinsèque" alors qu'elle est construite à partir de la seconde forme fondamentale et donc de la géométrie extrinsèque.

V.3.1 Courbure(s)

Définition V.3.1 (Direction et Courbure principal). *L'opérateur \mathbf{F}_P est symétrique. Il est donc diagonalisable dans une base orthogonale de l'espace tangent. On appelle courbures principales ses valeurs propres $k_1(P)$ et $k_2(P)$ et les vecteurs propres (qu'on normalisera) $e_1(P)$ et $e_2(P)$ correspondants s'appellent les directions principales, elles sont orthogonales (car \mathbf{F}_P est symétrique). Une courbe paramétrée incluse dans la surface $\gamma : I \rightarrow S$ est une ligne de courbure si et seulement si ses vecteurs tangents sont tous des directions principales.*

Proposition V.3.2. Soit k_1, k_2 les courbures principales et e_1, e_2 deux vecteurs propres orthogonaux associés, alors pour tout vecteur \mathbf{V} de l'espace tangent,

$$\exists \theta \text{ tel que } \mathbf{V} = \|\mathbf{V}\| (\cos \theta e_1 + \sin \theta e_2)$$

et

$$II_P(v, v) = \|\mathbf{V}\|^2 (k_1 \cos^2 \theta + k_2 \sin^2 \theta)$$

Définition V.3.3 (Courbure de Gauss et courbure principale). *Le produit des courbures principales s'appelle la courbure de Gauss.*

$$K_P = \det \mathbf{F}_P = k_1(P)k_2(P)$$

C'est donc également le déterminant de \mathbf{F}_P . La moyenne des courbures est alors appelée courbure moyenne.

$$H_P = \frac{\text{trace} \mathbf{F}_P}{2} = \frac{k_1(P) + k_2(P)}{2}$$

Remarque V.3.4. On donnera plus loin une interprétation des courbures principales, en effet la courbure d'une courbe paramétrée unitairement sur la surface peut se décomposer en deux courbures, la courbure géodésique (dont nous verrons la définition plus loin) et la courbure normal $II_P(\mathbf{T}, \mathbf{T})$ qui est donc comprise entre les deux courbures principales.

Pour faire des calculs concrets, il est utile d'écrire l'endomorphisme de Weingarten dans une paramétrisation. Les dérivées de l'application de Gauss locales sont parfois difficiles à calculer à cause de la normalisation, alors que les dérivées premières et secondes de la paramétrisation et donc les éléments de ligne des formes fondamentales sont souvent plus faciles à calculer. Il est donc intéressant de montrer que les éléments de lignes de l'opérateur de Weingarten peuvent s'exprimer en fonction de ceux de la première forme et de la seconde forme fondamentales.

Théorème (et définition) V.3.5. Dans une paramétrisation locale φ , si les matrices de la première forme fondamentale et de la seconde fondamentale s'écrivent ainsi

$$I_P = \begin{pmatrix} \partial_u \varphi \cdot \partial_u \varphi & \partial_u \varphi \cdot \partial_v \varphi \\ \partial_v \varphi \cdot \partial_u \varphi & \partial_v \varphi \cdot \partial_v \varphi \end{pmatrix} \hat{=} \begin{pmatrix} E(u, v) & F(u, v) \\ F(u, v) & G(u, v) \end{pmatrix} \quad II_P = \begin{pmatrix} n \cdot \partial_{uu}^2 \varphi & n \cdot \partial_{uv}^2 \varphi \\ n \cdot \partial_{vu}^2 \varphi & n \cdot \partial_{vv}^2 \varphi \end{pmatrix} \hat{=} \begin{pmatrix} P & Q \\ Q & R \end{pmatrix}$$

alors la matrice de l'endomorphisme de Weingarten s'écrit en ces termes

$$\mathbf{F}_P = \begin{pmatrix} E(u, v) & F(u, v) \\ F(u, v) & G(u, v) \end{pmatrix}^{-1} \begin{pmatrix} P(u, v) & Q(u, v) \\ Q(u, v) & R(u, v) \end{pmatrix}$$

Les fonctions E, F, G (éléments de ligne de la première forme fondamentale) et P, Q, R (éléments de ligne de la seconde forme fondamentale) sont définis sur le même ouvert que la paramétrisation. Ainsi la courbure de Gauss et moyennne s'expriment

$$K_P = \frac{PR - Q^2}{EG - F^2} \quad H_P = \frac{1}{2} \frac{ER + GP - 2FQ}{EG - F^2}$$

Preuve. On confondra les endomorphismes et leurs matrices dans la base naturelle associée à la paramétrisation. Donc pour tout $\mathbf{U}, \mathbf{V} \in \mathcal{T}_P(S)$,

$$II_P(\mathbf{U}, \mathbf{V}) = {}^t \mathbf{U} II_P \mathbf{V} = I_P(\mathbf{F}_P(\mathbf{U}), \mathbf{V}) = I_P(\mathbf{F}_P \mathbf{V}, \mathbf{U}) = {}^t \mathbf{V} I_P \mathbf{F}_P \mathbf{U}$$

On en déduit donc $II_P = \mathbf{F}_P I_P$. Les autres égalités s'en déduisent aisément. ■

Exercice 37. —

On considère la surface "selle" définie par le graphe $(x, y) \mapsto (x, y, xy)$. Calculer les éléments de ligne de la première et seconde forme fondamentale. En déduire la matrice de l'application de Weingarten.

Définition V.3.6 (Caractérisation des points d'une surface). *Soit un point $P \in S$ d'une surface régulière. Alors*

- *P est un point planaire si et seulement si $k_1(P) = k_2(P) = 0$*
- *P est un point ombilic si et seulement si $k_1(P) = k_2(P) \neq 0$*
- *P est un point parabolique si et seulement si $K_P = k_1(P)k_2(P) = 0$ (non planaire)*
- *P est un point elliptique si et seulement si $K_P > 0$*
- *P est un point hyperbolique si et seulement si $K_P < 0$*

Remarque V.3.7. Une démonstration calculatoire et technique montre qu'un ellipsoïde à 3 axes a 4 point ombilics. On laissera les étudiants motivés en apporter une démonstration. L'ensemble des points étant elliptiques.

Sur un tore, l'ensemble des points "extérieurs" sont elliptiques tandis que les points "intérieurs" sont hyperboliques, les points sur les deux grands cercles extrémants z sont paraboliques.

Au début de 1805, Laplace montre que la surface d'une interface entre deux milieux fluides est soumise à une force surfacique proportionnelle à la courbure moyenne de l'interface. Dans le cas d'une bulle de savon suffisamment fine, on peut négliger les forces de gravité. Une seule force s'applique alors au film qui représente la bulle : cette tension superficielle. La forme stationnaire de la bulle est donc caractérisée par l'annulation de la courbure moyenne.

Définition V.3.8 (Surface minimal). *Une surface est dite minimale lorsque sa courbure moyenne est identiquement nulle sur toute la surface.*

Remarque V.3.9. On en déduit facilement qu'une surface minimale est une surface dont les points sont soit hyperboliques, soit planaires.

Exercice 38. — Surface de Enneper.

La surface d'Enneper est la nappe géométrique de la nappe paramétrée,

$$\varphi(u, v) = \begin{pmatrix} u - \frac{u^3}{3} + uv^2 \\ v - \frac{v^3}{3} + u^2v \\ u^2 - v^2 \end{pmatrix}$$

Montrer qu'il s'agit d'une surface minimale

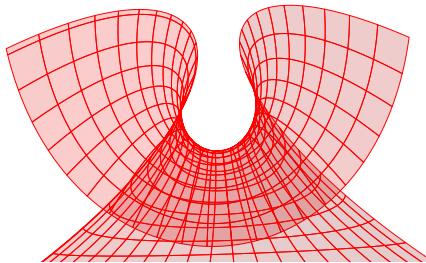


FIGURE V.15 – Surface d'Enneper

On peut ordonner les courbures principales de manière à avoir $k_1(P) \geq k_2(P)$

Proposition V.3.10. Les courbures principales sont des applications continues sur S . Les courbures principales sont des applications lisses au voisinage de tout point non ombilic.

Preuve. On a vu lors du problème précédent que la courbure de Gauss et la courbure moyenne étaient des applications lisses sur S . En effet, puisque les éléments de ligne des deux formes fondamentales sont lisses et que le déterminant de la première forme fondamentale ne s'annule

jamais ($EG - F^2 \neq 0$), les formules du théorème V.3.5 assurent le caractère lisse de la courbure de Gauss et de la courbure moyenne.

Les courbures principales sont les racines du polynômes $X^2 - 2H_P X + K$, d'où

$$\begin{cases} k_1(P) = H_P + \sqrt{H_P^2 - K_P} \\ k_2(P) = H_P - \sqrt{H_P^2 - K_P} \end{cases}$$

La continuité est assurée partout, par contre le caractère C^1 est remis en cause pour tous les points vérifiant $H_P^2 = K_P$: les ombilics.

Ce résultat s'étend à la base orthonormée des directions principales.

Proposition V.3.11. Il existe deux vecteurs tangents $P \mapsto \mathbf{v}_1(P), \mathbf{v}_2(P)$ continus sur S et lisses une base orthonormée $\mathbf{v}_1(P), \mathbf{v}_2(P)$ de $\mathcal{T}_P(S)$ constitué des directions principales de \mathbf{F}_P .

Preuve. Il suffit de noter que la matrice d'une endomorphisme symétrique est symétrique dans une base orthonormale. On construit cette base à partir de la base naturelle, en normalisant $\partial_u \varphi$ ce qui donne un vecteur tangent $\mathbf{u}_1(P)$ (application lisse car $\partial_u \varphi$ est lisse et $\|\partial_u \varphi(P)\|$ ne s'annule pas), puis en utilisant une orthonormalisation de Gram-Shmidt, c'est-à-dire qu'on pose $\mathbf{v}_2(P) = \partial_v \varphi(P) - (\mathbf{e}_1(P) \cdot \partial_v \varphi(P)) \mathbf{e}_1(P)$ qui est lisse par construction et ne s'annule pas sur la surface (surface régulière). On peut donc orthonormaliser $\mathbf{v}_2(P)$ pour obtenir un vecteur $\mathbf{u}_2(P)$ lisse. La matrice de passage $B(P)$ entre la base naturelle et la base orthonormale ainsi construite est lisse et sont inverse l'est aussi (car $\forall P \in S, B(P)$ est inversible). Ainsi la matrice de \mathbf{F}_P dans cette base est symétrique et c'est une application lisse sur la surface

$$\begin{aligned} M &: S \rightarrow \mathcal{M}_2(\mathbb{R}) \\ P &\mapsto \begin{pmatrix} a(P) & c(P) \\ c(P) & d(P) \end{pmatrix} \end{aligned}$$

En dehors des points ombilics ($a(P) = d(P)$ et $c(P) = 0$), un petit calcul conduit à montrer que les directions principales $\mathbf{e}_1, \mathbf{e}_2$ s'écrivent dans la base $(\mathbf{u}_1(P), \mathbf{u}_2(P))$:

$$\begin{aligned} \mathbf{e}_1 &= \frac{1}{\sqrt{4c^2 + [(a-d) - \sqrt{4c^2 + (a-d)^2}]^2}} \left[2c\mathbf{u}_1 + \left\{ (a-d) - \sqrt{4c^2 + (a-d)^2} \right\} \mathbf{u}_2 \right] \\ \mathbf{e}_2 &= \frac{1}{\sqrt{4c^2 + [(a-d) + \sqrt{4c^2 + (a-d)^2}]^2}} \left[2c\mathbf{u}_1 + \left\{ (a-d) + \sqrt{4c^2 + (a-d)^2} \right\} \mathbf{u}_2 \right] \end{aligned}$$

où a, c et d sont des fonctions scalaires lisses sur la surface. Les vecteurs propres sont donc lisses en dehors des points ombilics et forment une base orthonormée. ■

Remarque V.3.12. Il y a plusieurs cas où la différentiabilité des courbures principales et de la base des directions principales peut tout de même être assurée, dans le cas de portions de sphère par exemple.

V.3.2 Dérivation et symbole de Christoffel

Afin de démontrer le *Theorema Egregium*, nous allons définir une nouvelle série de scalaires sur la surface enregistrant les variations de la base naturelle. On calculera enfin ces quantités en fonction des éléments de lignes des deux formes fondamentales.

Definitions

Définition V.3.13 (Symbole de Christoffel). *Exprimons les dérivées seconde (ou les dérivées partielles de la base naturelle) dans la base constituée de la base naturelle et de la normale locale (définie par l'application de Gauss), ainsi que les dérivées partielles de l'application de Gauss :*

$$\begin{cases} \partial_{uu}^2 \varphi = \Gamma_{uu}^u \partial_u \varphi + \Gamma_{uu}^v \partial_v \varphi + L_u \mathbf{n} \\ \partial_{uv}^2 \varphi = \Gamma_{uv}^u \partial_u \varphi + \Gamma_{uv}^v \partial_v \varphi + \ell_v \mathbf{n} \\ \partial_{vu}^2 \varphi = \Gamma_{vu}^u \partial_u \varphi + \Gamma_{vu}^v \partial_v \varphi + \ell_u \mathbf{n} \\ \partial_{vv}^2 \varphi = \Gamma_{vv}^u \partial_u \varphi + \Gamma_{vv}^v \partial_v \varphi + L_v \mathbf{n} \end{cases}$$

$$\begin{cases} \partial_u \mathbf{n} = a_u^u \partial_u \varphi + a_u^v \partial_v \varphi \\ \partial_v \mathbf{n} = a_v^u \partial_u \varphi + a_v^v \partial_v \varphi \end{cases}$$

les coefficients Γ_{ij}^k sont appellés symboles de Christoffel. La définition de la seconde forme fondamentale et de ses éléments de ligne assure alors que

$$L_u = P \quad \ell_u = \ell_v = Q \quad L_v = R$$

Le théorème V.3.5 assure également alors que

$$\begin{cases} a_u^u = \frac{QF - PG}{EG - F^2} & a_u^v = \frac{PF - QE}{EG - F^2} \\ a_v^u = \frac{RF - QG}{EG - F^2} & a_v^v = \frac{QF - RE}{EG - F^2} \end{cases}$$

Courbure géodésique et normale d'une courbe paramétrée sur une surface

Ces définitions permettent d'exprimer la dérivée directionnelle d'un champ de vecteurs tangents à la surface. Celle-ci se décompose sur l'espace tangent (dérivée covariante) et l'espace normal.

Théorème (et définition) V.3.14 (Dérivée covariante). On définit la *dérivée covariante d'un champ de vecteurs $t : S \rightarrow \mathcal{T}_P(S)$ le long d'un autre champ de vecteurs tangents $w : S \rightarrow \mathcal{T}_P(S)$ par la projection de sa dérivée directionnelle $dt_P(w)$ sur le plan tangent. On la note $\nabla_w t$. On montre alors que :*

$$dt_P(w) = \nabla_w t(P) + II_P(w, t)\mathbf{n}(P)$$

C'est un autre champ de vecteurs tangents dont les composantes sont :

$$\nabla_w t(P) = \begin{bmatrix} [t^u \partial_u t^u + t^v \partial_v t^u + \Gamma_{uu}^u w^u t^u + \Gamma_{uv}^u w^u t^v + \Gamma_{vu}^u w^v t^u + \Gamma_{vv}^u w^v t^v] \partial_u \varphi \\ [t^u \partial_u t^v + t^v \partial_v t^v + \Gamma_{uu}^v w^u t^u + \Gamma_{uv}^v w^u t^v + \Gamma_{vu}^v w^v t^u + \Gamma_{vv}^v w^v t^v] \partial_v \varphi \end{bmatrix}$$

Preuve. Un calcul soigneux montre qu'on a

$$\begin{aligned} & [t^u \partial_u t^u + t^v \partial_v t^u + \Gamma_{uu}^u w^u t^u + \Gamma_{uv}^u w^u t^v + \Gamma_{vu}^u w^v t^u + \Gamma_{vv}^u w^v t^v] \partial_u \varphi(P) \\ dt_P(w) &= [t^u \partial_u t^v + t^v \partial_v t^v + \Gamma_{uu}^v w^u t^u + \Gamma_{uv}^v w^u t^v + \Gamma_{vu}^v w^v t^u + \Gamma_{vv}^v w^v t^v] \partial_v \varphi(P) \\ & [Pw^u t^u + Qw^u t^v + Qw^v t^u + R w^v t^v] \mathbf{n}(P) \end{aligned}$$

Ce qui permet de conclure, après projection. ■

Nous allons essayer, en introduisant la notion de courbure géodésique et de courbure normale d'une courbe, de donner un sens à la notion de dérivée covariantes mais aussi de courbures principales. Que se passe-t-il pour une courbe incluse ? On considère une courbe régulière Γ paramétrée par son abscisse curviligne incluse dans la surface. Localement on a $\Gamma = \varphi(\gamma(s))$. Alors

$$\mathbf{T}(s) = \frac{d\gamma^u}{ds} \partial_u \varphi + \frac{d\gamma^v}{ds} \partial_v \varphi$$

On admettra qu'on peut prolonger \mathbf{T} en dehors de la courbe de manière lisse de façon à définir un champ de vecteurs tangents $\mathbf{T} : \mathcal{U} \rightarrow S$ sur l'ouvert de la paramétrisation tel que $\mathbf{T}(\gamma(s)) = \mathbf{T}(s)$, ce qui nous permet d'écrire

$$\mathbf{T}(s) = \frac{d\gamma^u}{ds} \partial_u \varphi + \frac{d\gamma^v}{ds} \partial_v \varphi = T^u(\gamma(s)) \partial_u \varphi + T^v(\gamma(s)) \partial_v \varphi = T^\mu(\gamma(s)) \partial_\mu \varphi$$

en adoptant la convention d'Einstein sur la sommation des indices. Il vient

$$\begin{aligned} \frac{d\mathbf{T}}{ds} &= \partial_\mu T^\lambda(\gamma(s)) \frac{d\gamma^\mu}{ds} \partial_\lambda \varphi(\gamma(s)) + T^\lambda(\gamma(s)) \partial_{\mu\lambda}^2 \varphi(\gamma(s)) \frac{d\gamma^\mu}{ds} \\ &= T^\mu(\gamma(s)) \partial_\mu T^\lambda(\gamma(s)) \partial_\lambda \varphi(\gamma(s)) + T^\lambda(\gamma(s)) T^\mu(\gamma(s)) \Gamma_{\mu\lambda}^\kappa \partial_\kappa \varphi \\ &+ [PT^u T^u + QT^u T^v + QT^v T^u + RT^v T^v] \mathbf{n}(P) \end{aligned}$$

Ce qui permet enfin d'écrire

$$\frac{d\mathbf{T}}{ds} = \nabla_{\mathbf{T}} \mathbf{T}(\gamma(s)) + II_P(\mathbf{T}, \mathbf{T}) \mathbf{n}(P) = \boldsymbol{\kappa} \quad (\text{V.1})$$

Remarque V.3.15. Cela permet de définir deux courbures : la courbure géodésique $\|\nabla_T T(\gamma(s))\|$ qui est la courbure de l'arc projeté sur le plan tangent et $II_P(T, T)$, la courbure normale correspondant à la direction normale à la surface. En effet la courbure de la courbe gauche Γ , correspond à la norme du vecteur $\kappa = \frac{d^2\Gamma}{ds^2}$, la courbe étant incluse dans la surface, ce vecteur est décomposable dans $\mathbb{R}^3 = \mathcal{T}_{\Gamma(s)}(S) \oplus \mathbb{R}\mathbf{n}(\Gamma(s))$. L'expression de la courbure normale implique que toute courbe incluse dans la surface et de vecteur tangent unitaire $T(s)$ aura une courbure totale supérieur à sa courbure normale. Si l'on note $\mathcal{C}_P = \{\mathbf{V} \in \mathcal{T}_P(S) \mid \|\mathbf{V}\| = 1\}$, alors $\mathbf{V} \in \mathcal{C}_P \mapsto II_P(\mathbf{V}, \mathbf{V})$ est l'application qui indique le minimum de courbure atteint par une courbe incluse dans la surface passant par P ayant une direction \mathbf{V} dans le plan tangent. On montrera ultérieurement qu'une courbe plongé géodésique est une courbe qui vérifie en tout point $\nabla_T T(\gamma(s)) = 0$. Alors sa courbure est uniquement normale. Notons qu'on peut orienter la courbure géodésique à l'aide de l'orientation de la surface.

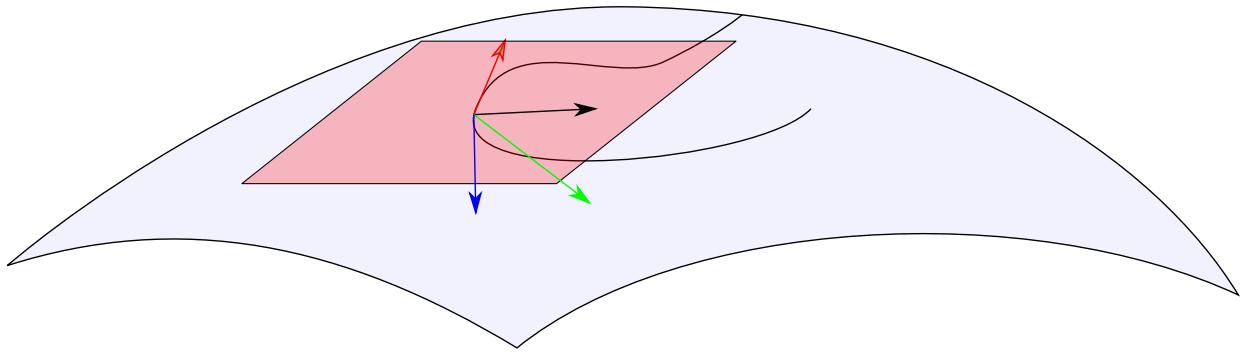


FIGURE V.16 – Représentation de la courbure normale et géodésique. Le vecteur tangent \mathbf{T} est en rouge, tandis que le vecteur courbure globale κ est en vert. La courbure géodésique $\nabla_T T(\gamma(s))$ est en noir et la courbure normale $II_P(\mathbf{T}, \mathbf{T})\mathbf{n}(P)$ en bleu.

Proposition V.3.16. Les symboles de Christoffel s'expriment en fonction des éléments de ligne de la première forme fondamentale et de leurs dérivées partielles. On a en particulier

$$\left\{ \begin{array}{l} \Gamma_{uu}^u = \frac{1}{2} \frac{G\partial_u E + F\partial_v E - 2F\partial_u F}{EG - F^2} \\ \Gamma_{uu}^v = \frac{1}{2} \frac{2E\partial_u F - E\partial_v E - F\partial_u E}{EG - F^2} \\ \Gamma_{vv}^u = \frac{1}{2} \frac{G\partial_v F - F\partial_u G}{EG - F^2} \\ \Gamma_{vv}^v = \frac{1}{2} \frac{E\partial_u G - F\partial_v E}{EG - F^2} \\ \Gamma_{uv}^u = \frac{1}{2} \frac{2G\partial_v F - G\partial_u G - F\partial_v G}{EG - F^2} \\ \Gamma_{uv}^v = \frac{1}{2} \frac{E\partial_v G + F\partial_u G - 2F\partial_v F}{EG - F^2} \end{array} \right.$$

Preuve. Inverser les systèmes obtenus en multipliant scalairement les dérivées partielles secondes avec la base naturelle. Utiliser Schwarz. ■

V.3.3 Le théorème remarquable et les équations de Mainardi-Coddazzi

Théorème V.3.17 (*Théorème Egregium*, ou théorème remarquable). La courbure de Gauss est entièrement déterminée par les éléments de lignes de la première forme fondamentale et de leurs dérivées partielles.

Preuve. On commence par calculer les dérivées troisièmes croisées (on ne calcule pas $\partial_{uuu}^3\varphi$ et $\partial_{vvv}^3\varphi$), en utilisant les dérivées secondes. On obtient :

$$\left\{ \begin{array}{l} \partial_v \partial_{uu}^2 \varphi = \left[\partial_v \Gamma_{uu}^u + \Gamma_{uu}^u \Gamma_{uv}^u + \Gamma_{uu}^v \Gamma_{vv}^u + P \frac{RF - QG}{EG - F^2} \right] \partial_u \varphi \\ \quad + \left[\partial_v \Gamma_{uu}^v + \Gamma_{uu}^u \Gamma_{uv}^v + \Gamma_{uu}^v \Gamma_{vv}^v + P \frac{QF - RE}{EG - F^2} \right] \partial_u \varphi \\ \quad + [\partial_u P + Q\Gamma_{uu}^u + R\Gamma_{uu}^v] N \\ \partial_u \partial_{uv}^2 \varphi = \left[\partial_u \Gamma_{uv}^u + \Gamma_{uv}^u \Gamma_{uu}^u + \Gamma_{uv}^v \Gamma_{vu}^u + Q \frac{QF - PG}{EG - F^2} \right] \partial_u \varphi \\ \quad + \left[\partial_u \Gamma_{uv}^v + \Gamma_{uv}^u \Gamma_{uu}^v + \Gamma_{uv}^v \Gamma_{vv}^v + Q \frac{PF - QE}{EG - F^2} \right] \partial_u \varphi \\ \quad + [\partial_u Q + P\Gamma_{uv}^u + Q\Gamma_{uv}^v] N \end{array} \right.$$

et

$$\left\{ \begin{array}{l} \partial_v \partial_{uv}^2 \varphi = \left[\partial_v \Gamma_{uv}^u + \Gamma_{uv}^u \Gamma_{uv}^u + \Gamma_{uv}^v \Gamma_{vv}^u + Q \frac{RF - QG}{EG - F^2} \right] \partial_u \varphi \\ \quad + \left[\partial_v \Gamma_{uv}^v + \Gamma_{uv}^u \Gamma_{uv}^v + \Gamma_{uv}^v \Gamma_{vv}^v + Q \frac{QF - RE}{EG - F^2} \right] \partial_u \varphi \\ \quad + [\partial_v P + Q\Gamma_{uv}^u + R\Gamma_{uv}^v] N \\ \partial_u \partial_{vv}^2 \varphi = \left[\partial_u \Gamma_{vv}^u + \Gamma_{vv}^u \Gamma_{uu}^u + \Gamma_{vv}^v \Gamma_{uv}^u + R \frac{QF - PG}{EG - F^2} \right] \partial_u \varphi \\ \quad + \left[\partial_u \Gamma_{vv}^v + \Gamma_{vv}^u \Gamma_{uu}^v + \Gamma_{vv}^v \Gamma_{vv}^v + R \frac{PF - QE}{EG - F^2} \right] \partial_u \varphi \\ \quad + [\partial_u R + P\Gamma_{vv}^u + Q\Gamma_{vv}^v] N \end{array} \right.$$

Schwarz assure que $\partial_v \partial_{uu}^2 \varphi = \partial_u \partial_{uv}^2 \varphi$ et $\partial_v \partial_{uv}^2 \varphi = \partial_u \partial_{vv}^2 \varphi$ en identifiant les composantes $\partial_u \varphi$ des deux premiers termes et en remaniant un peu, on arrive à

$$G \frac{PR - Q^2}{EG - F^2} = [\partial_u \Gamma_{uv}^u - \partial_v \Gamma_{uu}^u + \Gamma_{uv}^v \Gamma_{vu}^u - \Gamma_{uu}^v \Gamma_{vv}^u]$$

Or les résultats de la proposition V.3.5 impliquent

$$GK_P = [\partial_u \Gamma_{vv}^u - \partial_v \Gamma_{uv}^u + \Gamma_{vv}^v \Gamma_{uu}^u - \Gamma_{uv}^u \Gamma_{vv}^u]$$

et la proposition V.3.16 ainsi que la non annulation de G permettent de conclure quand au fait que K_P n'est l'expression que des éléments de lignes de la première forme fondamentale. ■

Les résultats des équations précédentes sont plus riches que l'expression du *théorème Egregium*. En effet nous obtenons une formule qui bien que compliquée permet le calcul de K_P à partir des éléments de ligne. Des autres identifications possibles on tire deux autres équations équivalentes nous donnant la courbure mais aussi des équations liant les éléments de ligne des deux formes fondamentales et en particulier exprimant les variations des éléments de ligne de la seconde forme fondamentale : les équations de *Mainardi-Coddazzi*. On obtient après un long calcul :

Proposition V.3.18. L'équation de *Gauss*,

$$K_P = \frac{1}{(EG - F^2)^2} \left[\begin{vmatrix} -\frac{1}{2} \partial_{vv}^2 E + \partial_{vv}^2 F - \frac{1}{2} \partial_{uu}^2 G & \frac{1}{2} \partial_u F & \partial_u F - \frac{1}{2} \partial_v E \\ \partial_v F - \frac{1}{2} \partial_u G & E & F \\ \frac{1}{2} \partial_v G & F & G \end{vmatrix} - \begin{vmatrix} 0 & \frac{1}{2} \partial_v E & \frac{1}{2} \partial_u G \\ \frac{1}{2} \partial_v E & E & F \\ \frac{1}{2} \partial_u G & F & G \end{vmatrix} \right]$$

Et les équations de *Mainardi-Coddazzi*

$$\left\{ \begin{array}{l} \partial_u Q - \partial_v P = -P\Gamma_{uv}^u + Q(\Gamma_{uu}^u - \Gamma_{uv}^v) + R\Gamma_{uu}^v \\ \partial_u R - \partial_v Q = -P\Gamma_{vv}^u + Q(\Gamma_{uv}^u - \Gamma_{vv}^v) + R\Gamma_{uv}^v \end{array} \right.$$

Remarque V.3.19. Deux surfaces localement isométriques ont la même courbure de Gauss.

Deux surfaces localement isométriques ont les mêmes géodésiques.

Deux surfaces peuvent avoir les mêmes éléments de ligne pour la première forme fondamentale, mais avoir des éléments de lignes de seconde forme fondamentale différents. Exemple : plan, cylindre, cône, surface développable.

Quand on cherche la nappe paramétrée à partir de l'expression des éléments de lignes, on dit qu'on cherche un plongement localement isométrique. Le théorème de Nash assure son existence pour les surfaces dans un espace Euclidien de dimension au pire $3 \times 17\dots$

Néanmoins,

Proposition V.3.20. Deux surfaces ayant les mêmes éléments de ligne pour la première et la seconde forme fondamentales sont liés entre elles par une isométrie globale de R^3 (c'est-à-dire que la seconde surface peut être obtenue par translation-rotation voire symétrie de la première surface).

Nous ne démontrerons pas ce théorème difficile et renvoyons le lecteur au livre de CARMO 1976 p.311-314.

V.3.4 Équation de Gauss et de Codazzi-Mainardi dans la base principale

Nous allons exprimer les résultats précédents en utilisant non plus la base naturelle, mais une base orthonormée directe (vis à vis de l'application de Gauss $\mathbf{n}(P)$) de l'espace tangent constitué des directions principales. On l'appellera *base principale*. Ces résultats permettront une meilleure compréhension en termes géométriques. En effet, la base naturelle est liée à une paramétrisation et donc n'a rien de lié à la structure géométrique locale d'ordre deux (courbure) de la surface.

On a vu lors de la démonstration de la proposition V.3.11 qu'une telle base existe et peut être considérée comme continue partout et lisse en dehors des points ombilics. Dans ce qui suivra, on se placera au voisinage de points d'une surface régulière où cette base pourra être considérée comme lisse (voisinage d'un point non-ombilic, ou sur une portion de sphère). On notera $P \in S \mapsto \mathbf{e}_1, \mathbf{e}_2 \in \mathcal{T}_P(S)$ cette base, on introduit $\mathbf{e}_N(P) = \mathbf{e}_1(P) \wedge \mathbf{e}_2(P) = \mathbf{n}(P)$. La matrice de la seconde forme fondamentale est, dans cette base la diagonale

$$\text{Mat}_{\mathbf{e}_1(P), \mathbf{e}_2(P)}(II_P) = \begin{pmatrix} -(\mathbf{d}\mathbf{e}_N)_P(\mathbf{e}_1) \cdot \mathbf{e}_1 & -(\mathbf{d}\mathbf{e}_N)_P(\mathbf{e}_1) \cdot \mathbf{e}_2 \\ -(\mathbf{d}\mathbf{e}_N)_P(\mathbf{e}_2) \cdot \mathbf{e}_1 & -(\mathbf{d}\mathbf{e}_N)_P(\mathbf{e}_2) \cdot \mathbf{e}_2 \end{pmatrix} = \begin{pmatrix} (\mathbf{d}\mathbf{e}_1)_P(\mathbf{e}_1) \cdot \mathbf{e}_N & (\mathbf{d}\mathbf{e}_2)_P(\mathbf{e}_1) \cdot \mathbf{e}_N \\ (\mathbf{d}\mathbf{e}_1)_P(\mathbf{e}_2) \cdot \mathbf{e}_N & (\mathbf{d}\mathbf{e}_2)_P(\mathbf{e}_2) \cdot \mathbf{e}_N \end{pmatrix} = \begin{pmatrix} k_1(P) & 0 \\ 0 \cdot v_1 & k_2(P) \end{pmatrix}$$

On introduit alors les quantités équivalentes au symboles de Christoffel de la base naturelle, mais cette fois pour la base principale :

$$\begin{aligned} (\mathbf{d}\mathbf{e}_1)_P(\mathbf{e}_i) \cdot \mathbf{e}_1 &= (\mathbf{d}\mathbf{e}_2)_P(\mathbf{e}_i) \cdot \mathbf{e}_2 = 0 \\ (\mathbf{d}\mathbf{e}_1)_P(\mathbf{e}_i) \cdot \mathbf{e}_2 &= (\mathbf{d}\mathbf{e}_2)_P(\mathbf{e}_i) \cdot \mathbf{e}_1 = \Gamma_i \end{aligned}$$

où on a déjà utilisé le fait que cette base reste orthonormale. On est donc à même de calculer les différentielles des vecteurs de la base principale.

$$\begin{aligned} (\mathbf{d}\mathbf{e}_1)_P(\mathbf{e}_1) &= \Gamma_1 \mathbf{e}_2 + k_1 \mathbf{e}_N \\ (\mathbf{d}\mathbf{e}_2)_P(\mathbf{e}_1) &= -\Gamma_2 \mathbf{e}_1 \\ (\mathbf{d}\mathbf{e}_1)_P(\mathbf{e}_2) &= \Gamma_1 \mathbf{e}_2 \\ (\mathbf{d}\mathbf{e}_2)_P(\mathbf{e}_2) &= -\Gamma_2 \mathbf{e}_1 + k_2 \mathbf{e}_N \end{aligned}$$

On rappelle que pour toute application $f : S \rightarrow R^d$ lisse, $\forall \mathbf{v} \in \mathcal{T}_P(S)$ il existe $\epsilon > 0$ et $\gamma :]-\epsilon, \epsilon[\rightarrow S$ tel que $\gamma(0) = P$ et $\dot{\gamma}(0) = \mathbf{v}$, alors

$$(\mathbf{d}f)_P(\mathbf{v}) = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} = \lim_{t \rightarrow 0} \frac{f(\gamma(t)) - f(P)}{t}$$

Dans ce qui suit, pour tout champ de vecteurs $\mathbf{V} \in \mathcal{T}_P(S)$, on notera $D_{\mathbf{V}} f : P \mapsto \mathbf{d}f_P(\mathbf{V})$. Les équations de Gauss et de Codazzi-Mainardi proviennent du théorème de Schwarz des ordres 3 d'une paramétrisation, ce qui revient au théorème de Schwarz appliqué à l'ordre deux sur la base naturelle. Nous devons donc explorer la "dérivée croisée" de la base principale. En effet, on va chercher la valeur de $D_{\mathbf{e}_1}(D_{\mathbf{e}_2} f)(P) - D_{\mathbf{e}_2}(D_{\mathbf{e}_1} f)(P)$.

Pour cela construisons à partir du point P , les chemins

- Posons $\varphi : \mathcal{U} \rightarrow S \cap O$ une paramétrisation locale dans un voisinage O de P et posons pour tout $M \in S \cap O$ de ce voisinage $\tilde{e}_1 = (\mathrm{d}\varphi^{-1})_M^{-1}(e_1) = [(\mathrm{d}\varphi)_{\varphi^{-1}M}]^{-1}(e_1)$ et $\tilde{e}_2 = (\mathrm{d}\varphi)_M^{-1}(e_2) = [(\mathrm{d}\varphi)_{\varphi^{-1}M}]^{-1}(e_2)$. On définit ainsi sur \mathcal{U} deux champs de vecteurs. Posons $m = \varphi^{-1}(P)$, on construit ainsi $c_1, c_2 : I \rightarrow \mathcal{U}$ solution de l'équation différentielle $\frac{\mathrm{d}c_1}{\mathrm{d}t} = \tilde{e}_1(c_1(t))$ et $\frac{\mathrm{d}c_2}{\mathrm{d}t} = \tilde{e}_2(c_2(t))$ avec pour condition initiale $c_1(0) = c_2(0) = P$. On pose alors $C_1(t) = \varphi(c_1(t))$ et donc $\forall t$ où $c_1(t)$ est définie, on a $\dot{C}_1(t) = (\mathrm{d}\varphi)_{c_1(t)}\left(\frac{\mathrm{d}c_1}{\mathrm{d}t}\right) = (\mathrm{d}\varphi)_{c_1(t)}[(\mathrm{d}\varphi_{c_1(t)})^{-1}(e_1(C_1(t)))] = e_1(C_1(t))$. De même on montre que $\dot{C}_2(t) = e_2(C_2(t))$. On a construit de cette manière deux courbes paramétrées $C_1 :]-\delta_1, \delta_1[\rightarrow S$ et $C_2 :]-\delta_2, \delta_2[\rightarrow S$ incluses dans la surface et dont la tangente est partout colinéaire à e_1 (pour C_1) ou e_2 (pour C_2).
- On peut $\forall \epsilon_1 \epsilon_2$ (suffisamment petits) répéter cette opération et construire de la même manière une courbe incluse $\tilde{C}_2 :]-\tilde{\delta}_2, \tilde{\delta}_2[\rightarrow S$ (tangentes à e_2) sauf qu'on prendra pour condition initiale $\tilde{C}_2(0) = C_1(\epsilon_1)$. On effectue à nouveau la même opération en suivant cette fois-ci le flots de e_1 et on construit $\tilde{C}_1 :]-\tilde{\delta}_1, \tilde{\delta}_1[\rightarrow S$ avec pour condition initiale $\tilde{C}_1(0) = C_2(\epsilon_1)$

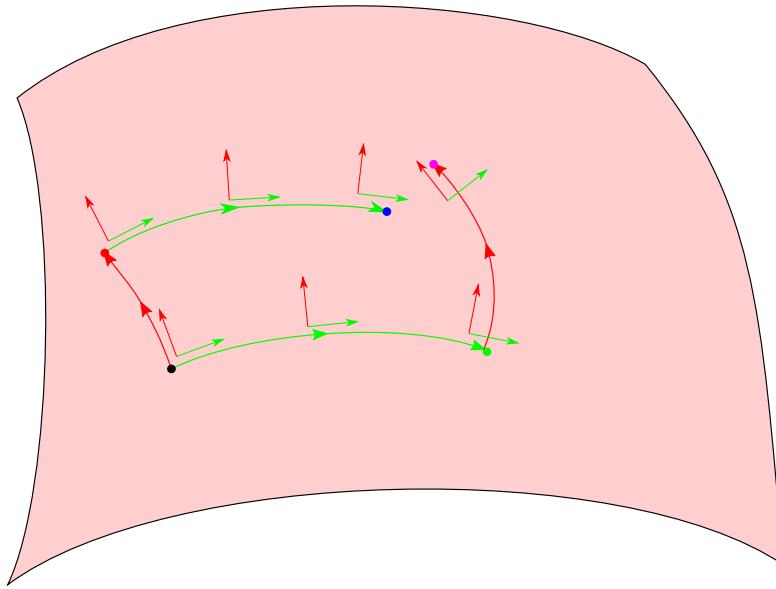


FIGURE V.17 – Représentation des courbes paramétrées construites, on a P en noir, $C_1(\epsilon_1)$ en vert, $\tilde{C}_2(\epsilon_2)$ en mauve, $C_2(\epsilon_1)$ en rouge et $\tilde{C}_1(\epsilon_2)$ en bleu.

On a donc,

$$D_{e_1}(D_{e_2}f)(P) = \frac{\mathrm{d}}{\mathrm{d}t} ((D_{e_2}f)(C_1(t))) \Big|_{t=0} = \lim_{\epsilon_1 \rightarrow 0} \frac{(D_{e_2}f)(C_1(\epsilon_1)) - (D_{e_2}f)(P)}{\epsilon_1}$$

et puisque

$$\begin{aligned} (D_{e_2}f)(C_1(\epsilon_1)) &= \lim_{\epsilon_2 \rightarrow 0} \frac{f(\tilde{C}_2(\epsilon_2)) - f(C_1(\epsilon_1))}{\epsilon_2} \\ (D_{e_2}f)(P) &= \lim_{\epsilon_2 \rightarrow 0} \frac{f(C_2(\epsilon_1)) - f(P)}{\epsilon_2} \end{aligned}$$

On a

$$D_{e_1}(D_{e_2}f)(P) = \lim_{\epsilon_1 \rightarrow 0} \lim_{\epsilon_2 \rightarrow 0} \frac{f(\tilde{C}_2(\epsilon_2)) - f(C_1(\epsilon_1)) - f(C_2(\epsilon_2)) + f(P)}{\epsilon_1 \epsilon_2}$$

et donc en intervertissant 1 et 2

$$D_{e_2}(D_{e_1}f)(P) = \lim_{\epsilon_2 \rightarrow 0} \lim_{\epsilon_1 \rightarrow 0} \frac{f(\tilde{C}_1(\epsilon_2)) - f(C_2(\epsilon_2)) - f(C_1(\epsilon_1)) + f(P)}{\epsilon_1 \epsilon_2}$$

Puisque cette limite existe, on peut intervertir les limites. On obtient donc au final

$$(D_{e_1}D_{e_2} - D_{e_2}D_{e_1})f(P) = \lim_{(\epsilon_1, \epsilon_2) \rightarrow (0,0)} \frac{f(\tilde{C}_2(\epsilon_2)) - f(\tilde{C}_1(\epsilon_2))}{\epsilon_1 \epsilon_2} = \tau(\epsilon_1, \epsilon_2)$$

Or τ est continue en $(0, 0)$ ce qui implique que

$$(D_{\mathbf{e}_1} D_{\mathbf{e}_2} - D_{\mathbf{e}_2} D_{\mathbf{e}_1})f(P) = \lim_{\epsilon \rightarrow 0} \frac{f(\tilde{C}_2(\epsilon)) - f(\tilde{C}_1(\epsilon))}{\epsilon^2}$$

Pour calculer cette valeur il va falloir avoir une approximation à l'ordre 2 des bouts de chemin $\tilde{C}_2(\epsilon)$ et $\tilde{C}_1(\epsilon)$. On a $C_1(t) = P + \mathbf{e}_1(P)t + O(t^2)$ et donc

$$\dot{C}_1(t) = \mathbf{e}_1(P + \mathbf{e}_1(P)t + O(t^2)) = \mathbf{e}_1(P) + D_{\mathbf{e}_1} \mathbf{e}_1 t + O(t^2) = \mathbf{e}_1(P) + \Gamma_1 t \mathbf{e}_2 + O(t^2)$$

et donc

$$C_1(t) = P + \mathbf{e}_1(P)t + \frac{\Gamma_1(P)t^2}{2} \mathbf{e}_2(P) + O(t^3)$$

Continuons pour le chemin allant du point vert au point mauve, on a

$$\dot{\tilde{C}}_2(t) = \mathbf{e}_2(C_1(\epsilon) + t \mathbf{e}_2(C_1(\epsilon)) + O(t^2)) = \mathbf{e}_2(C_1(\epsilon)) + t D_{\mathbf{e}_2} \mathbf{e}_2(C_1(\epsilon)) + O(t^2)$$

ce qui implique

$$\begin{aligned} \tilde{C}_2(t) &= C_1(\epsilon) + t \mathbf{e}_2(C_1(\epsilon)) + \frac{t^2}{2} [-\Gamma_2(C_1(\epsilon)) \mathbf{e}_1(C_1(\epsilon)) + k_2(C_1(\epsilon)) v_N(C_1(\epsilon))] + O(t^3) \\ \tilde{C}_2(t) &= P + \epsilon \mathbf{e}_1(P) + \Gamma_1(P) \frac{\epsilon^2}{2} + t [\mathbf{e}_2(P) - \Gamma_1 \epsilon \mathbf{e}_1(P)] + \frac{t^2}{2} [-\Gamma_2(P) \mathbf{e}_1(P) + k_2(P) v_N(P)] + O(t^3) \end{aligned}$$

et donc finalement

$$\tilde{C}_2(\epsilon) = P + \epsilon (\mathbf{e}_1(P) + \mathbf{e}_2(P)) + \frac{\epsilon^2}{2} (\Gamma_1 \mathbf{e}_2 - \Gamma_2 \mathbf{e}_1 - 2\Gamma_1 \mathbf{e}_1 + (k_1 + k_2) \mathbf{e}_N) + O(\epsilon^3)$$

De la même manière on exprime $\tilde{C}_1(\epsilon)$,

$$\tilde{C}_1(\epsilon) = P + \epsilon (\mathbf{e}_1(P) + \mathbf{e}_2(P)) + \frac{\epsilon^2}{2} (\Gamma_1 \mathbf{e}_2 - \Gamma_2 \mathbf{e}_1 + 2\Gamma_2 \mathbf{e}_2 + (k_1 + k_2) \mathbf{e}_N) + O(\epsilon^3)$$

on en déduit

$$\tilde{C}_1(\epsilon) \tilde{C}_2(\epsilon) = -\epsilon^2 (\Gamma_1 \mathbf{e}_1 + \Gamma_2 \mathbf{e}_2) + \epsilon^3 m(\epsilon).$$

Ce qui implique

$$\begin{aligned} (D_{\mathbf{e}_2} D_{\mathbf{e}_1} - D_{\mathbf{e}_1} D_{\mathbf{e}_2})f(P) &= \lim_{\epsilon \rightarrow 0} \frac{f(\tilde{C}_1(\epsilon)) - f(\tilde{C}_2(\epsilon))}{\epsilon^2} \\ &= \lim_{\epsilon \rightarrow 0} \frac{f(\tilde{C}_2(\epsilon) + \epsilon^2 (\Gamma_1 \mathbf{e}_1 + \Gamma_2 \mathbf{e}_2) + \epsilon^3 m(\epsilon)) - f(\tilde{C}_2(\epsilon))}{\epsilon^2} \\ &= \Gamma_1 D_{\mathbf{e}_1} f + \Gamma_2 D_{\mathbf{e}_2} f \end{aligned}$$

On a donc la formule, concernant la permutation des dérivées directionnelles de la base principale sur une fonction quelconque :

$$(D_{\mathbf{e}_2} D_{\mathbf{e}_1} - D_{\mathbf{e}_1} D_{\mathbf{e}_2})f(P) = \Gamma_1 D_{\mathbf{e}_1} f + \Gamma_2 D_{\mathbf{e}_2} f$$

Dans la partie précédente, on a vu que les équations de Gauss sont la traduction du théorème de Schwarz s'appliquant à la base naturelle. Ici le rôle des dérivées partielles, qui sont des dérivées directionnelles dans la direction des vecteurs de la base naturelle sont remplacées par les dérivées directionnelles dans la base principale et le Théorème de Schwarz est substitué par l'équation précédente. On applique donc la formule précédente au vecteurs \mathbf{e}_1

$$\begin{aligned} \Gamma_1 D_{\mathbf{e}_1} \mathbf{e}_1 + \Gamma_2 D_{\mathbf{e}_2} \mathbf{e}_1 &= (D_{\mathbf{e}_2} D_{\mathbf{e}_1} - D_{\mathbf{e}_1} D_{\mathbf{e}_2}) \mathbf{e}_1 \\ (\Gamma_1^2 + \Gamma_2^2) \mathbf{e}_2 + k_1 \Gamma_1 \mathbf{e}_N &= D_{\mathbf{e}_2} [\Gamma_1 \mathbf{e}_2 + k_1 \mathbf{e}_N] - D_{\mathbf{e}_1} [\Gamma_2 \mathbf{e}_2] \\ (\Gamma_1^2 + \Gamma_2^2) \mathbf{e}_2 + k_1 \Gamma_1 \mathbf{e}_N &= (D_{\mathbf{e}_2} \Gamma_1) \mathbf{e}_2 + \Gamma_1 (-\Gamma_2 \mathbf{e}_1 + k_2 \mathbf{e}_N) + (D_{\mathbf{e}_2} k_1) \mathbf{e}_N - k_1 k_2 \mathbf{e}_2 - (D_{\mathbf{e}_1} \Gamma_2) \mathbf{e}_2 + \Gamma_1 \Gamma_2 \mathbf{e}_1 \\ \implies & [(D_{\mathbf{e}_2} \Gamma_1) - (D_{\mathbf{e}_1} \Gamma_2) - \Gamma_1^2 - \Gamma_2^2 - k_1 k_2] \mathbf{e}_2 + [(D_{\mathbf{e}_2} k_1) - \Gamma_1 (k_1 - k_2)] \mathbf{e}_N = 0 \end{aligned}$$

De même pour \mathbf{e}_2 , on obtient,

$$[(D_{\mathbf{e}_2} \Gamma_1) - (D_{\mathbf{e}_1} \Gamma_2) - \Gamma_1^2 - \Gamma_2^2 - k_1 k_2] \mathbf{e}_1 + [(D_{\mathbf{e}_1} k_2) - \Gamma_1 (k_1 - k_2)] \mathbf{e}_N = 0$$

Ce qui est l'expression de l'équations de Gauss et de Coddazzi-Mainardi dans la base principale :

$$\boxed{\begin{cases} K = k_1 k_2 = (D_{\mathbf{e}_2} \Gamma_1) - (D_{\mathbf{e}_1} \Gamma_2) - \Gamma_1^2 - \Gamma_2^2 & \text{Gauss} \\ (D_{\mathbf{e}_2} k_1) = \Gamma_1 (k_1 - k_2) \quad \text{et} \quad (D_{\mathbf{e}_1} k_2) = \Gamma_2 (k_1 - k_2) & \text{Coddazzi-Mainardi} \end{cases}}$$

Remarque V.3.21. Les équations précédentes sont bien plus compactes que celles utilisant les base naturelles associées à un paramétrage, cela traduit le caractère "géométriquement adapté" de la base principale. On remarque que dans le cas d'un point ombilic, les équations de Codazzi-Mainardi impliquent que le gradient de k_1 est aligné avec e_1 et que le gradient de k_2 est aligné avec e_2 .

V.4 Classification des surfaces compactes à bord

La notion de surface compacte à bord est intéressante notamment pour préciser des notions nécessaires pour aborder le théorème de Stokes bi-dimensionnel. Nous en effectuerons une classification quand celles-ci seront compactes. En effet, nous montrerons que le bord de telles surfaces est la réunion d'un nombre fini de courbes paramétrées homéomorphes au cercle.

Nous souhaitons remercier POL 2012 dont les documents de mémoire nous ont beaucoup aidé lors de la rédaction de cette section.

V.4.1 Définitions des surfaces à bord

Nous avons vu lors des définitions fondamentales des surfaces sans bords que celles-ci étaient difféomorphes au disque unité ouvert au voisinage de tout point. C'est-à-dire qu'à partir de chaque point on trouvera des points voisins dans toutes les directions. Quand on considère l'adhérence d'une surface sans bord, on voit que certains points de la frontière ne peuvent être traités de cette manière : une partie des directions sont interdites. Ce qui nous amène à comprendre la définition d'une surface à bord.

Définition V.4.1 (Surface à bord). *Une surface régulière S à bord (régulier) est une partie de \mathbb{R}^3 dans laquelle tout point $P \in S$ possède un voisinage ouvert O qui est difféomorphe au disque unité ouvert $h : \mathcal{D}_1 \subset \mathbb{R}^2 \rightarrow O \cap S$ (avec $h(0;0) = P$) ou (non exclusif à ce stade) homéomorphe (difféomorphe) au demi-disque unité $h : \mathcal{D}_{1/2} \subset \mathbb{R}^2 \rightarrow O \cap S$ (avec $h(0;0) = P$) où le demi-disque unité est l'ensemble*

$$\mathcal{D}_{1/2} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1 \text{ et } x \geq 0\} \subset \mathbb{R}^2$$

On appelle intérieur de S (noté $\text{int}S$) l'ensemble des points pour lesquels il existe un voisinage relatif ouvert difféomorphe au disque unité et bord de S l'ensemble des points (noté ∂S) pour lesquels il existe un voisinage relatif ouvert homéomorphe (difféomorphe) au demi-disque unité

Remarque V.4.2. Le carré du plan est une surface à bord (non régulier). Le disque unité fermé est une surface à bord régulier. La demi-sphère supérieure est une surface régulière à bord régulier.

La proposition V.1.3 assure qu'une surface sans bord est une surface dont le bord est vide.

V.4.2 Partition intérieur extérieur

Un point peut-il appartenir à la fois à l'intérieur et aux bords ? On montre qu'il y a partition entre ces points :

Proposition V.4.3. Soit S une surface régulière à bord $\forall P \in S$. Soit P est un point du bords, soit P est intérieur.

Preuve. Supposons qu'il existe $P \in S$ surface régulière à bord telle que $P \in \text{int}S \cap \partial S$. Alors

- $\exists O_I$ et $\exists h_I : \mathcal{D}_1 \rightarrow O_I \cap S$ difféomorphisme tel que $h_I(0,0) = P$
- $\exists O_B$ et $\exists h_B : \mathcal{D}_{1/2} \rightarrow O_B \cap S$ difféomorphisme tel que $h_B(0,0) = P$

On pose alors $O = O_I \cap O_B$ ouvert contenant P , $\mathcal{U}_{1/2} = h_B^{-1}(O \cap S)$ et $\mathcal{U}_1 = h_I^{-1}(O \cap S)$ et enfin

$$\begin{aligned} f &: \mathcal{U}_1 &\rightarrow \mathcal{U}_{1/2} \\ x &\mapsto h_B^{-1} \circ h_I(x) \end{aligned}$$

C'est par composition un homéomorphisme. Il vérifie donc $f(\mathring{\mathcal{U}}_1) \subset [f(\mathring{\mathcal{U}}_1)]$. Or \mathcal{U}_1 est un ouvert et donc $\mathring{\mathcal{U}}_1 = \mathcal{U}_1$ et donc $f(\mathring{\mathcal{U}}_1) = f(\mathcal{U}_1) = \mathcal{U}_{1/2} \subset f(\mathring{\mathcal{U}}_1) = \mathring{\mathcal{U}}_{1/2}$ et donc $\mathcal{U}_{1/2}$ est ouvert.

Or $P \in O \cap S$ et donc $(0,0) \in \mathcal{U}_1$ et $(0,0) \in \mathcal{U}_{1/2}$, et $(0,0) \in \text{Fr}(\mathcal{D}_{1/2})$ et donc $(0,0) \in \text{Fr}(\mathcal{U}_{1/2})$ ce qui en contradiction avec $\mathcal{U}_{1/2}$ ouvert. ■

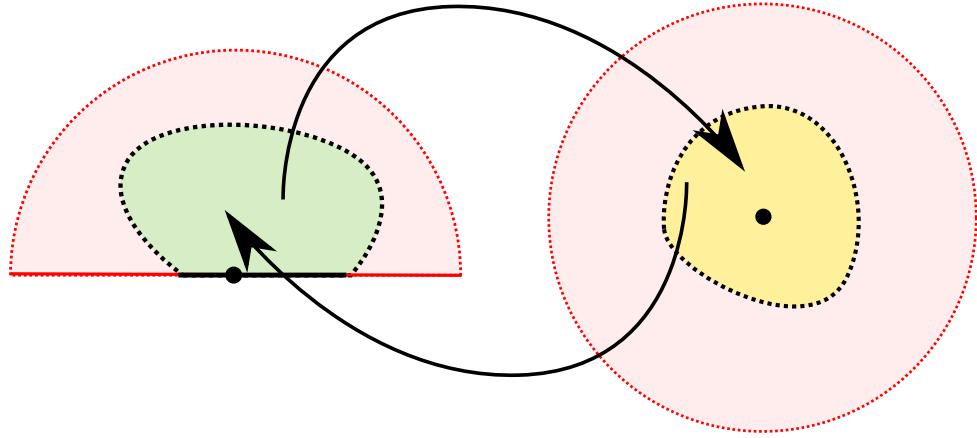


FIGURE V.18 – Représentation de l'homéomorphisme f construit lors de la preuve. Le demi-disque et le disque sont en rose pâle quand \mathcal{U}_1 est en jaune et $\mathcal{U}_{1/2}$ en vert. La frontière de \mathcal{U}_1 a été représentée en noir. Les points $(0, 0)$ sont les petits points noirs. Les flèches noires symbolisent le passage homéomorphe entre l'espace en jaune et l'espace en vert.

Ainsi les points ne peuvent appartenir au bord et à l'intérieur d'une surface. Cela provient du fait que les homéomorphismes conservent la structure topologique des espaces sur lesquels ils sont définis. On montre que si un point $M \in S$ est tel que quelque soit l'ouvert O contenant M , $O \cap \partial S \neq \emptyset$ alors ce point appartient au bord.

Proposition V.4.4. Soit S une surface régulière à bord. Alors $M \in S$ appartient à ∂S si et seulement si $\forall \epsilon > 0$, $B(M, \epsilon \cap \partial S \neq \emptyset$

Preuve. Le sens direct est évident si l'on suit les définitions. En effet prenons $M \in \partial S$, alors $\forall \epsilon > 0$, $B(M, \epsilon \cap \partial S)$ contient M et donc est non vide.

Le sens inverse se démontre par l'absurde. Supposons que $M \in \text{int}S$, alors il existe un ouvert O contenant M tel que $O \cap S$ soit difféomorphe à \mathcal{D}_1 , notons $h : \mathcal{D}_1 \rightarrow O \cap S$ ce difféomorphisme.

Puisque $M \in O$ est O ouvert, $\exists \epsilon > 0$ tel que $B(M, \epsilon) \subset O$. Puisque $B(M, \epsilon \cap \partial S \neq \emptyset$, on note $N \in B(M, \epsilon \cap \partial S)$ et considérons l'ouvert \tilde{O} tel que $\tilde{O} \cap S$ soit homéomorphe à $\mathcal{D}_{1/2}$, notons $\tilde{h} : \mathcal{D}_1 \rightarrow \tilde{O} \cap S$ cet homéomorphisme, vérifiant $\tilde{h}(0, 0) = N$.

Alors on considère $\Omega = B(M, \epsilon \cap \tilde{O})$ (qui par construction est inclus dans O et contient N), on peut donc construire $D_1 = h^{-1}(\Omega) \subset \mathcal{D}_1$ et $D_{1/2} = \tilde{h}^{-1}(\Omega) \subset \mathcal{D}_{1/2}$ et donc $f : D_1 \rightarrow D_{1/2}$ est un homéomorphisme. Ce qui est impossible puisque $D_{1/2} - D_{1/2} \ni (0, 0) \neq$ comme dans la démonstration précédente. ■

Ce qui a pour conséquence immédiate,

Corollaire V.4.5. Soit S une surface régulière à bord. Alors $M \in S$ appartient à l'intérieur si et seulement si il existe un rayon $\epsilon > 0$ tel que $B(M, \epsilon \cap \partial S) = \emptyset$

V.4.3 Bord localement difféomorphe à une courbe paramétrée

Les bords des exemples de surfaces à bords mentionnés plus haut (carré du plan, le disque unité fermé, la demi-sphère supérieure fermé) sont tous localement des courbes paramétrées, régulières dans le cas du disque ou de la demi-sphère. C'est une propriété générale du bord des surfaces régulières, elle sont des objets de dimension $2 - 1 = 1$. En effet,

Proposition V.4.6. Soit S une surface régulière à bords régulier, alors $\forall P \in \partial S$ il existe une courbe paramétrée régulière injective incluse dans le bord de S et passant par P . Autrement dit ∂S est localement une courbe géométrique.

Preuve. On considère $P \in \partial S$ et W l'ouvert de \mathbb{R}^3 tel qu'il existe un difféomorphisme $h : \mathcal{D}_{1/2} \rightarrow W \cap S$. On pose alors $\gamma : t \in [-1, 1] \mapsto h(t, 0)$, une courbe paramétrée régulière (h est un difféomorphisme et donc $\dot{\gamma} = \partial_x h(0, t) \neq 0$). On montre maintenant que pour tout $t \in [-1, 1]$, $\gamma(t) \in \partial S$.

En effet, pour tout $t \in [-1, 1]$, on pose $r = \frac{1}{2} \min(|1-t|, |1+t|)$ et $\tilde{D}_t = B((t, 0), r) \cap \mathcal{D}_{1/2}$ un ouvert relatif de l'espace de départ. Son image par un homéomorphisme est donc un ouvert relatif de $W \cap S$. Il existe donc un ouvert W_t de \mathbb{R}^3 tel que $h(\tilde{D}_t) = W_t \cap S$, W_t contient $\gamma(t)$. On pose alors $h_t : m \in \mathcal{D}_{1/2} \mapsto h(rm + (t, 0)) \in W_t \cap S$. Par construction h_t est un difféomorphisme entre $\mathcal{D}_{1/2}$ et $W_t \cap S$ de plus $h_t(0, 0) = h(t, 0) = \gamma(t)$. Et donc $\gamma(t) \in \partial S$. Par construction γ est injective. ■

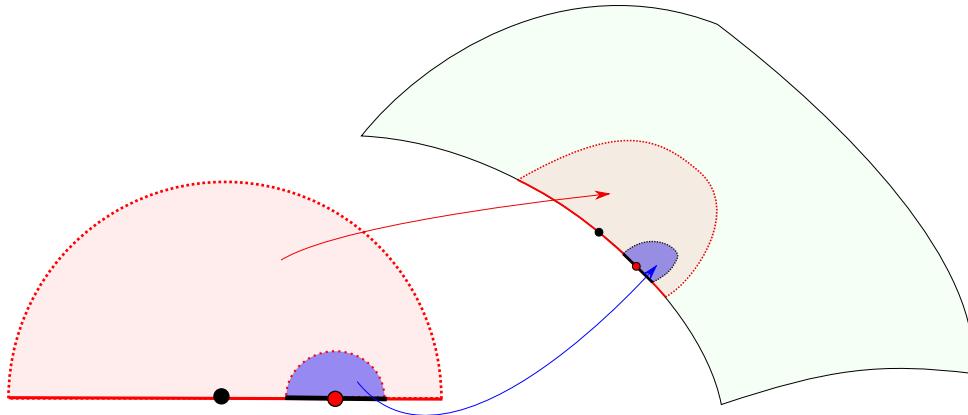


FIGURE V.19 – Représentation de l’homéomorphisme h (flèche rouge) et h_t (flèche bleue) construit lors de la preuve. Le demi-disque est en rose pâle $(t, 0)$ et $\gamma(t)$ sont représentés par des points rouge cerclés de noir. Les ensembles W_t et \tilde{D}_t sont en bleu. La surface S est représentée en vert.

V.4.4 Surface compacte et bord comme union de courbes paramétrées fermées simples

Que peut-on dire sur le bord d'une surface régulière à bord régulier ? La proposition précédente nous indique que le bord est localement une courbe paramétrée. Les bords des surfaces à bord mentionné sont tous des courbes fermées simples. Est-ce le cas de toutes ? L'exemple de la partie supérieure du tore suffit à nous convaincre qu'il est possible d'avoir l'union d'un nombre fini de courbes fermées simples. Si nous recollons bord à bord les chemins locaux, est-il obligatoire d'obtenir un arc fermé simple ?

Le contre-exemple de la bande infinie $(u, v) \in [-1, 1] \times \mathbb{R} \mapsto uu_x + vu_y \in \mathbb{R}^3$ montre qu'il faut ajouter des contraintes à la surface régulière pour être sûr que son bord soit constitué d'arcs fermés simples. S'il suffisait de contraindre l'aire de la surface, on a malheureusement l'exemple de la surface sous la courbe $x \mapsto \exp(-x^2)$ dont l'aire est $\sqrt{\pi}$, dont les bords sont également constitué de deux courbes paramétrées de longueurs infinies. On peut regarder ce qu'il en est des surfaces bornées, malheureusement cette fois-ci encore la contrainte n'est pas suffisante. En effet, prenons la surface contenue entre les deux bras de la spirale $\rho : \theta \mapsto \theta / \sqrt{1 - \theta^2}$. Dans ce cas nous avons une surface régulière à bord régulier borné dont le bord est constitué d'une courbe paramétrée de longueur infinie. Cette surface est-elle fermée ? On se convainc qu'on est capable de créer des suites qui convergent vers un point du cercle unité qui n'appartient pas à la surface.

Nous allons voir que dans les cas des surfaces régulières à bord régulier qui sont fermées et bornées (c'est-à-dire compact) ont un bord qui est constitué d'un nombre fini de courbes fermées simples. En effet, la compacité, et en particulier la propriété de Borel-Lebesgue qui y est attachée permet souvent de contraindre globalement et de déduire des propriétés globales.

Théorème V.4.7. Le bord d'une surface compacte régulière à bord régulier est constitué d'une union finie de courbes paramétrées régulières fermées simples.

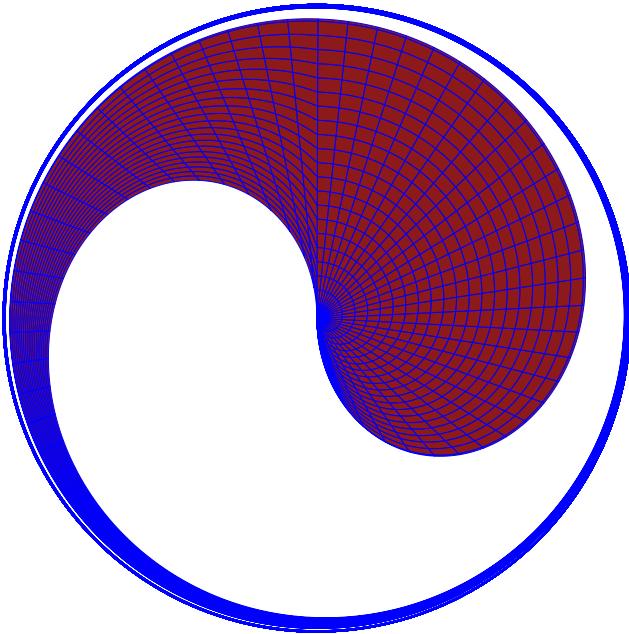


FIGURE V.20 – Surface comprise entre les bras spiraux définis par l'équation polaire $\rho : \theta \mapsto \theta/\sqrt{1 + \theta^2}$. Posons $\gamma : \theta \mapsto \rho(\theta)\mathbf{u}(\theta)$. La surface est l'image de l'application $\varphi : (t, \theta)[0, 1] \times [-\pi, +\infty[\rightarrow t\gamma(\theta) - (1-t)\gamma(\theta - \pi)$, on comprend que le bord est constitué d'une courbe paramétrée bornée de longueur infinie. L'ensemble $\bar{S} - \tilde{S}$ est constitué du bord et du cercle unité qui n'appartient pas à la surface : celle-ci n'est pas fermée donc pas compacte.

Preuve. Etape 1 : On a vu dans la proposition V.4.6 que localement, le bord d'une surface était une courbe paramétrée, le bord est donc la réunion de ces courbes. On va commencer par montrer qu'on peut se ramener à un nombre fini de courbe en utilisant la propriété de Borel-Lesbesgue.

Soit S une surface régulière à bord régulier, la proposition V.4.3 implique que pour tout point $P \in S$, soit $P \in \text{int}S$ soit $P \in \partial S$.

- $\forall P \in \text{int}S$, il existe un ouvert $U_P \ni P$ tel qu'il existe un difféomorphisme $h_{I,P} : \mathcal{D}_1 \rightarrow U_P \cap S$
- $\forall P \in \partial S$, il existe un ouvert $V_P \ni P$ tel qu'il existe un difféomorphisme $h_{B,P} : \mathcal{D}_{1/2} \rightarrow V_P \cap S$

On a donc

$$S = \bigcup_{P \in S} \{P\} = \bigcup_{P \in S} \{P\} = \left(\bigcup_{P \in \text{int}S} \{P\} = \bigcup_{P \in S} \{P\} \right) \cup \left(\bigcup_{P \in \partial S} \{P\} \right) = \left(\bigcup_{P \in \text{int}S} (U_P \cap S) \right) \cup \left(\bigcup_{P \in \partial S} (V_P \cap S) \right)$$

On a donc écrit S comme un recouvrement d'ouverts relatifs. La propriété de Borel-Lesbesgue nous permet d'en extraire un sous-recouvrement fini, c'est-à-dire qu'il existe un nombre fini de points $(P_i)_{i=1 \dots n} \in \text{int}S$ et un nombre fini de points $(Q_j)_{j=1 \dots k} \in \partial S$ tel que si $\forall i \in \{1 \dots n\}$ on note $U_{P_i} = U_i$ et $\forall j \in \{1 \dots k\}$ on note $V_{P_i} = V_i$ alors

$$S = \left(\bigcup_{i=1}^n (U_i \cap S) \right) \cup \left(\bigcup_{j=1}^k (V_j \cap S) \right)$$

On montre que $\forall P \in \text{int}S$, U_P ne contient pas de point du bord. En effet si un tel point M existait alors il serait possible d'avoir un ouvert V_M tel que $V_M \cap S$ soit difféomorphe au disque unité. Alors comme dans la preuve de V.4.3, on construirait $W = U_P \cap V_M \ni M$ qui est un ouvert non vide car contenant M et union de deux ouverts. Dès lors on poserait $\mathcal{U}_1 = h_{I,P}^{-1}(W \cap S)$ qui serait un ouvert relatif de \mathcal{D}_1 et $\mathcal{U}_{1/2} = h_{B,M}^{-1}(W \cap S)$ un ouvert relatif de $\mathcal{D}_{1/2}$, ouvert relatif qui contiendrait le point $(0, 0) = h_{B,M}^{-1}(M)$. Et enfin on considérerait l'homéomorphisme $f : x \in \mathcal{U}_1 \mapsto h_{B,P}^{-1} \circ h_{I,M}(x)$ ce qui est impossible car les structures topologiques de $\mathcal{D}_{1/2}$ et \mathcal{D}_1 ne sont pas les mêmes.

On a donc $\forall i \in \{1 \dots n\} U_i \cap \partial S =$ et donc $\partial S \subset \left(\bigcup_{j=1}^k (V_j \cap S) \right).$

Comme dans la preuve de V.4.6, $\forall j \in \{1 \dots k\} \gamma_j : t \in]-1, 1[\mapsto h_{B, Q_j}(t, 0)$ est à valeurs dans ∂S . Le bord est donc constitué d'au plus k courbes paramétrées. De plus la courbe géométrique de γ_j est $V_j \cap \partial S$ (il s'agit du même type d'argument que lors de la démonstration de V.4.3, si il existait dans $\mathcal{D}_{1/2}$ des antécédents de points du bord on serait à même de construire un homéomorphisme entre un ouvert et un ensemble non-ouvert ce qui n'est pas possible).

Etape 2 : On va maintenant montrer qu'il est possible de raccorder les courbes paramétrées. À partir de la liste $(V_j \cap \partial S)_{j=1 \dots k}$ on construit la liste minimal, c'est-à-dire la sous liste $(A_j = V_j \cap \partial S)_{j=1 \dots m}$ (où on a éventuellement réordonné les éléments et $m \leq k$) où si l'on retire le moindre élément, la liste ne recouvre plus ∂S .

On va montrer que $\forall j \in \llbracket 1, m \rrbracket, \exists i \neq j$ tel que $A_i \cap A_j \neq \emptyset$. Supposons $\exists j \in \llbracket 1, m \rrbracket, \exists i \neq j$ tel que $A_i \cap A_j = \emptyset$. Il est important de rappeler que γ_j est un homéomorphisme entre $] -1, 1[$ et A_j . Dans la suite on notera $h_{B, Q_j} = h_j$.

On considère alors

$$\begin{cases} u_n^+ = \gamma_j \left(1 - \frac{1}{2^n} \right) \\ u_n^- = \gamma_j \left(-1 + \frac{1}{2^n} \right) \end{cases}$$

Il se peut que ces suites ne convergent pas. Néanmoins, ces suites appartiennent à S qui est compacte et donc la propriété de Bolzano-Weirstrass assure qu'il existe une sous-suite qui converge. Notons $v_n^+ = u_{\varphi^+(n)}^+$ et $v_n^- = u_{\varphi^-(n)}^-$ ces sous-suites qui convergent ($\varphi^- : \mathbb{N} \rightarrow \mathbb{N}$ et $\varphi^+ : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissantes) et ℓ^-, ℓ^+ . Les limites ℓ^-, ℓ^+ appartiennent à S (S compact donc fermé) sont-elles dans ∂S ou dans $\text{int } S$? On a

$$\forall \epsilon > 0, \exists N_\epsilon^+ \text{ tel que } \forall n > N_\epsilon^+, v_n^+ \in B(\ell^+, \epsilon) \quad \forall \epsilon > 0, \exists N_\epsilon^- \text{ tel que } \forall n > N_\epsilon^-, v_n^- \in B(\ell^-, \epsilon)$$

Or $\forall n, v_n^+ = \gamma_j \left(1 - \frac{1}{2^{\varphi^+(n)}} \right)$ et donc $v_n^+ \in \partial S$, de même $v_n^- \in \partial S$ et donc $\forall \epsilon > 0, B(\ell^+, \epsilon) \cap \partial S \neq \emptyset$ et $B(\ell^-, \epsilon) \cap \partial S \neq \emptyset$. La proposition V.4.4 implique que $\ell^+ \in \partial S$ et $\ell^- \in \partial S$. Ces limites n'appartiennent pas à A_j .

En effet, si $\ell_+ \in A_j$ (on travaillera avec ℓ^+ mais le résultat sera valable avec ℓ^-), alors on note $t_+ \in]-1, 1[$ sont antécédent par γ_j , on a $\ell^+ = \gamma_j(t_+) = h_j(t_+, 0)$. On considère alors $\forall \epsilon > 0$ (suffisamment petit pour que $B(\ell^+, \epsilon) \subset V_j$), l'image réciproque $\mathcal{U}_{+, \epsilon} = h_j^{-1}(B(\ell^+, \epsilon) \cap S)$ qui est un ouvert connexe par arc (car la continuité préserve la connexité par arc). De plus pour tout $\epsilon > 0$, l'ensemble $\mathcal{U}_{+, \epsilon}$ contient un voisinage de $(t_+, 0)$ (car h_j continue et $\ell^+ = h_j(t_+, 0)$). De plus puisque $v_n^+ = h_j \left(1 - \frac{1}{2^{\varphi^+(n)}} \right) \rightarrow \ell^+$, $\exists N_\epsilon$ tel que $\forall n \geq N_\epsilon, v_n^+ \in B(\ell^+, \epsilon)$ et donc $\mathcal{U}_{+, \epsilon}$ contient tous les termes de la suite $x_n = \left(1 - \frac{1}{2^{\varphi^+(n)}}, 0 \right)$ d'indices supérieurs à N_ϵ et donc un voisinage ouvert de ces points. On considère alors l'arc de cercle

$$\mathcal{C} = \left\{ (x, y) \in \mathcal{D}_{1/2} \mid (x - t_0)^2 + y^2 = \left(\frac{1 + t_0}{2} \right)^2 \right\}$$

c'est un compact inclus dans $\mathcal{D}_{1/2}$, en considérant sur $\mathcal{C} \in c \mapsto d(\ell_+, h_j(c))$ la distance de ℓ_+ à $h_j(c \in \mathcal{C})$ qui est continue à valeurs dans \mathbb{R}_+ et donc atteint ses bornes puisque \mathcal{C} est compacte et donc on note $m = \min \{d(\ell_+, h_j(c)) \mid c \in \mathcal{C}\} = d(\ell_+, h_j(c_0))$ ($m \neq 0$ sinon $\ell_+ = h_j(c)$ ce qui impliquerait $c_0 = (t_0, 0)$ ce qui est impossible puisque $t_0 \in]-1, 1[$). On prend alors $\epsilon < m$ et donc il n'existe aucun point de \mathcal{C} qui appartient à $\mathcal{U}_{+, \epsilon}$. On a donc partitionné un ouvert connexe par arc en deux ouverts disjoints non vides ce qui est impossible. Et donc $\ell^+ \notin A_j$, de la même manière $\ell^- \notin A_j$.

Et donc il existe $k_+, k_- \neq j$ tel que $\ell_+ \in A_{k_+}$ et $\ell_- \in A_{k_-}$. De plus pour $\forall \epsilon > 0$ on a $\partial S \cap B(\ell_+, \epsilon)$ contient des points de A_j et A_{k_+} et donc puisque $\partial S \cap B(\ell_+, \epsilon)$ est localement une ligne paramétrée $A_j \cap A_{k_+} \neq$ de même $A_j \cap A_{k_-} \neq$.

Notons $A_{j, k_+} = A_j \cap A_{k_+}$ et $W_+ = \gamma_j^{-1}(A_{j, k_+})$

- 1er cas : $W_+ =]a_+, 1[$ alors $A_j \cup A_{k_+}$ est un courbe géométrique injective connexe et non fermée, on remplace alors A_j et A_{k_+} dans la liste des courbes géométriques $(A_k)_{k=1 \dots m-1}$
- 2em cas : $W_+ =]-1, b_+[\cup]a_+, 1[$ alors $A_j \cup A_{k_+}$ est une composante connexe fermée injective donc simple. On a trouvé une composante connexe, on retire ces deux courbes géométriques de la liste $(A_k)_{k=1 \dots m-2}$

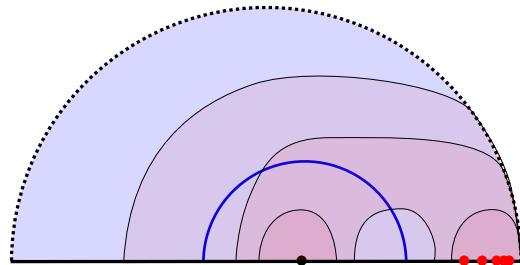


FIGURE V.21 – Représentation de la configuration de la démonstration, en mauve de plus en plus prononcé l’ensemble $\mathcal{U}_{+, \epsilon}$ à mesure que ϵ décroît. Le cercle \mathcal{C} est ici représenté en bleu, quand les termes de x_n sont en rouge. On remarque que pour $\epsilon < m$ l’ensemble $\mathcal{U}_{+, \epsilon}$ se retrouve forcément partitionné entre ses point à l’intérieur et ses point à l’extérieur du cercle. Il ne peut plus être connexe par arc.

Les autres cas ne sont pas possibles. Pour en savoir plus, voir (POL 2012). Dans tous les cas, le cardinal de la liste des courbes géométriques diminue à chaque étape. Quand il ne reste que deux éléments, on est forcément dans le second cas, car dans ce cas-là les limites ℓ^+ et ℓ^- ne peuvent appartenir qu’à l’élément restant de la liste. ■

CHAPITRE VI

Variétés

Sommaire

VI.1 Qu'est ce qu'une variété...	116
VI.2 Algèbre extérieure	129
VI.3 Dérivation	141
VI.4 Intégration	160
VI.5 Flots et dérivée de Lie	165

Les variétés sont des espaces différents, a priori, des espaces vectoriels. En effet, la notion de distance entre deux points ne peut plus s'appréhender par l'algèbre linéaire ni donc par la norme de vecteurs. Nous connaissons l'exemple des surfaces régulières : sur une sphère comme la Terre, tracer une ligne droite qui transperce le sol entre Paris et Moscou n'a pas grand intérêt pour savoir quelle longueur de rail il faudra pour relier ces deux villes. Cependant, localement, les rails sont droits : on peut faire de l'algèbre linéaire dans le plan tangent.

Ce genre de considération est encore plus important en relativité générale, car l'espace-temps est une variété intrinsèquement courbée : contrairement aux surfaces de \mathbb{R}^3 , l'espace-temps n'a pas à être considéré comme plongé dans un espace euclidien, sa métrique (ou sa première forme fondamentale) est un objet intrinsèque attaché à l'espace-temps. On ne peut donc pas "visualiser" la variété comme une surface plongée dans un espace plus gros puisque l'on vit littéralement dedans ! Tout ce que l'on peut faire (en relativité générale), c'est de prédire les résultats des expérimentations d'un observateur – ce qui est l'essence même de la physique.

Il existe aussi d'autres applications à la physique. Par exemple, dès lors qu'un système dynamique à n degrés de liberté a une intégrale première f , alors cela veut dire que si y est une solution de l'équation différentielle qui définit le système dynamique en question, nous avons $f(y) = cst$. Pour peu que f soit une fonction suffisamment régulière, cela définit une sous-variété par submersion. La plupart des équations différentielles n'étant pas solubles explicitement, les mathématiciens s'intéressent à donner des propriétés géométriques des solutions. Par exemple, savoir que les solutions vivent dans telle variété et non telle autre peut donner des contraintes intéressantes.

Nous allons donc donner quelques propriétés utiles sur ce type d'espace qu'on appelle variétés différentiables.

Pour rédiger ce chapitre, j'ai abondamment utilisé les ouvrages suivants que je recommande aux étudiants qui veulent approfondir les notions abordées :

1. L'excellent polycopié de cours d'Éric GOURGOULHON 2014 de relativité générale disponible ici <https://luth.obspm.fr/~luthier/gourgoulhon/fr/master/relatM2.pdf> ou sur sa page internet <https://luth.obspm.fr/~luthier/gourgoulhon/> Ce cours n'est pas un cours sur les tenseurs à proprement parler mais les objets sont introduits de façon intuitives et géométriques sans pour autant perdre trop de rigueur mathématique, bien que l'objet du cours traite plus de la physique de la relativité que des mathématiques. C'est tout le talent pédagogique d'Éric Gourgoulhon.
2. Le livre de "relativité restreinte" d'Éric GOURGOULHON 2010, aux éditions du CNRS. Ce livre exhaustif épouse absolument tout ce qu'on peut faire dans l'espace de Minkowski. La relativité restreinte y est traitée avec des systèmes de coordonnées quelconques, ce qui a poussé Éric Gourgoulhon à expliciter toutes les mathématiques nécessaires... y compris la notion

de connexion affine ! La plupart des notions abordées dans ce cours peuvent se retrouver dans le livre d'Éric Gourgoulhon. Même s'il se restreint au cas de l'espace de Minkowski, sa formulation est si élégante que la généralisation aux variétés différentielles quelconques ne pose aucun problème. Les développements du cours sur le champ électromagnétique sont tous tirés de cet ouvrage.

3. Le polycopié de cours de géométrie différentielle de HELEIN 2016 est relativement complet et bien rédigé. Il va plus loin que nous sur certaines notions (l'image réciproque des formes différentielles est plus approfondie). Il est disponible ici [https://webusers.imj-prg.fr/~frederic.helein/cours.html](https://webusers.imj-prg.fr/~frederic.helein/cours/sdem1.pdf).
4. Le livre très (trop ?) complet de relativité générale de STRAUMANN 1984, aux éditions Springer, qui contient des chapitres entiers sur les aspects mathématiques de la géométrie différentielle. Assez difficile à lire, je le recommande pour les lecteurs ambitieux.

Les deux derniers cours sont à un niveau plus avancé que l'introduction que nous proposons ici, nous encourageons les étudiants motivés à les consulter pour aller plus loin.

VI.1 Qu'est ce qu'une variété...

VI.1.1 Qu'est-ce qu'une variété ?

Une variété est un espace qui ressemble localement à \mathbb{R}^n . On dit alors que la variété est de dimension n . Pour éviter les pathologies, les mathématiciens utilisent souvent la définition suivante

Définition VI.1.1. Une variété différentielle \mathcal{M} de dimension n est un espace topologique muni d'un atlas, c'est-à-dire d'un ensemble de cartes locales $\phi_i : \mathcal{V}_i \subset \mathcal{M} \rightarrow \mathcal{U}_i \subset \mathbb{R}^n$ où $i \in I$ tel que :

VI.1.1.1. I est dénombrable.

VI.1.1.2. $\forall i \in I, \mathcal{V}_i \subset \mathcal{M}$ est ouvert et $\bigcup_{i \in I} \mathcal{V}_i = \mathcal{M}$.

VI.1.1.3. ϕ_i est un C^1 difféomorphisme. Donc \mathcal{U}_i est ouvert.

VI.1.1.4. Si $\mathcal{V}_{ij} = \mathcal{V}_i \cap \mathcal{V}_j \neq \emptyset$, alors $\phi_{ij} = (\phi_j \circ \phi_i^{-1})|_{\phi_i(\mathcal{V}_{ij})}$ est un C^1 difféomorphisme de $\phi_i(\mathcal{V}_{ij})$ dans $\phi_j(\mathcal{V}_{ij})$.

On dit que la variété est de classe C^k si les cartes et les applications de recollement ϕ_{ij} le sont.

Nous représentons symboliquement les applications de recollement figure VI.1 (à droite), ainsi que la notion d'atlas figures VI.1 (à gauche) et VI.2.

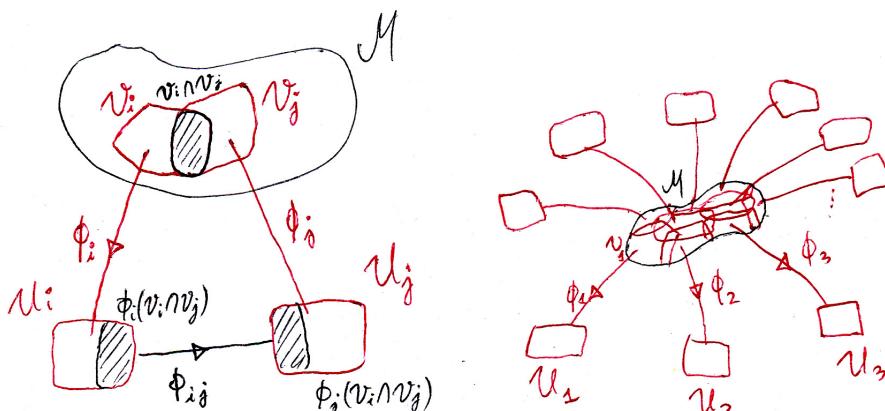


FIGURE VI.1 – A droite, application de recollement entre $\phi_i(\mathcal{V}_i \cap \mathcal{V}_j)$ et $\phi_j(\mathcal{V}_i \cap \mathcal{V}_j)$. A gauche, Représentation d'un atlas de \mathcal{M} .

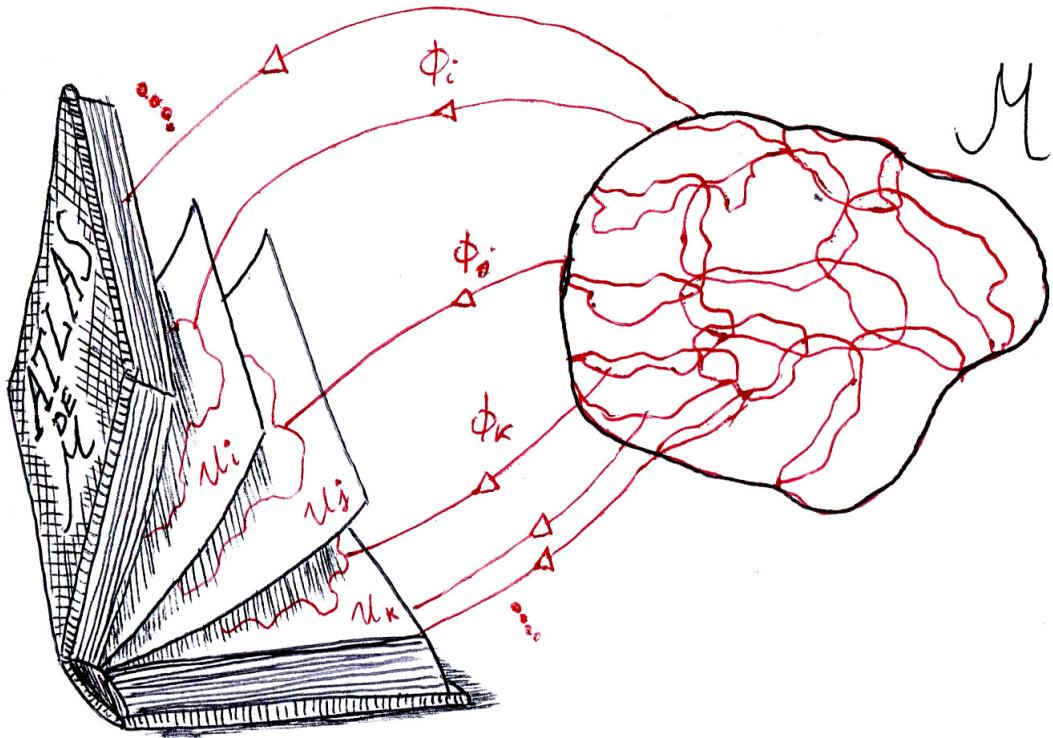


FIGURE VI.2 – Nous avons fait une représentation plus métaphorique pour illustrer le choix de ce nom. L'idée est de disposer d'une liste finie exhaustive de cartes qui recouvrent toute la variété. Si le nombre de cartes est infini mais dénombrables, il faut simplement imaginer un livre avec un nombre infini de pages.

VI.1.2 Espace tangent

Construction géométrique de l'espace tangent

Pour construire l'espace tangent en un point de la variété, il faut que les cartes soient différentiables et qu'on puisse passer d'une carte à l'autre de façon différentiable en tout point d'intersection ; alors seulement on dit que la variété est différentiable. Voyons comment construire l'espace tangent. Remarquons pour commencer que nous avons déjà rencontré des variétés différentiables dotées d'atlas : les surfaces régulières par immersion en sont un exemple explicite, et la définition des variétés différentiables n'est qu'une extension de cette définition. Par ailleurs, en raison de la courbure de la surface, les vecteurs tangents sortent de celle-ci, de telle sorte qu'il est impossible de relier deux points de la surface par un vecteur dans l'espace tangent. Nous avons besoin de courbes paramétrées $\gamma : I \rightarrow \mathcal{M}, \lambda \mapsto \gamma(\lambda) = \mathcal{P}(\lambda) \in \mathcal{M}$. Localement, on peut disposer d'une carte et donc d'un système de coordonnées pour repérer $\mathcal{P}(\lambda)$. Par abus de langage, on décrira donc localement γ par n fonctions différentiables dont on indexera le numéro de la composante par l'indice : $x^\mu = \gamma^\mu(\lambda)$. De façon plus rigoureuse, tant que $\gamma(\lambda) \in \mathcal{V}_i \subset \mathcal{M}$, on a $\gamma(\lambda) = \phi_i^{-1}(\gamma^1(\lambda), \dots, \gamma^n(\lambda))$. De même, très souvent, nous identifierons la carte ϕ_i par son système de coordonnées $(x^k)_{k \in \llbracket 1, n \rrbracket}$ que nous noterons plus simplement (x^k) . Dans cette notation, le fait que k varie de 1 à n , la dimension de la variété, est implicite.

Prenons un champ scalaire différentiable sur \mathcal{M} : c'est une fonction qui à un point de \mathcal{M} associe un réel. Voyons comment varie f au voisinage de \mathcal{P} le long de l'arc γ . On a

$$f(\gamma(\lambda + \varepsilon)) = f(\gamma(\lambda)) + \varepsilon \frac{d\gamma^j}{d\lambda}(\lambda) \frac{\partial f}{\partial x^j}(\gamma(\lambda)) + o(\varepsilon) \quad (\text{VI.1})$$

Désormais, nous adoptons la convention d'Einstein : à chaque fois qu'un indice se répète une fois en haut et une fois en bas, nous effectuons la sommation sur cet indice : $a^k b_k = \sum_{k=1}^n a^k b_k$ par

convention. Ici il faut donc faire implicitement

$$\frac{d\gamma^j}{d\lambda}(\lambda) \frac{\partial f}{\partial x^j}(\gamma(\lambda)) = \sum_{j=1}^n \frac{d\gamma^j}{d\lambda}(\lambda) \frac{\partial f}{\partial x^j}(\gamma(\lambda)) \quad (\text{VI.2})$$

Pour les deux équations précédentes, nous avons utilisé le fait que $f \circ \gamma = f \circ \phi_i^{-1} \circ \tilde{\gamma}$ où $\tilde{\gamma} = (\gamma^1, \dots, \gamma^n)$ sont les n fonctions qui permettent de représenter γ dans la carte ϕ_i . En vérité, les dérivées partielles de f de l'équation ci-dessus sont les dérivées partielles de la fonction $f \circ \phi_i^{-1}$. Lorsque cela ne présentera pas d'ambiguïté, nous identifierons γ et $\tilde{\gamma}$, ainsi que f et $f \circ \phi_i^{-1}$ dans les notations. En effet, comme les cartes sont des difféomorphismes, nous savons d'après le chapitre 3 que nous pouvons effectuer le changement de variables sans problème. Le terme différentiel de cette équation peut s'écrire

$$\frac{d\gamma^j}{d\lambda}(\lambda) \frac{\partial f}{\partial x^j}(\lambda) = \vec{v}(\mathcal{P}(\lambda)) \cdot \vec{\nabla} f(\mathcal{P}(\lambda)) \quad (\text{VI.3})$$

où l'on a posé $\vec{v}(\gamma(\lambda)) = \vec{\gamma}'(\lambda) = d\vec{\gamma}/d\lambda$. Considérons $\vec{v}(\mathcal{P}(\lambda))$ comme un objet agissant sur f :

$$\vec{v}(\mathcal{P}(\lambda))f = D_{\vec{v}}f = \left. \frac{df}{d\lambda} \right|_{\gamma} \quad (\text{VI.4})$$

c'est-à-dire que $\vec{v}(f)$ est la dérivée de f le long de l'arc γ dont la tangente est \vec{v} en $\mathcal{P} = \gamma(\lambda)$. Dans le système de coordonnées (x^k) cela devient tout simplement

$$\vec{v}(f) = v^k \partial_k f \quad (\text{VI.5})$$

Pour établir une base, naturellement adaptée à (x^k) , de l'espace vectoriel dans lequel vivent les vecteurs, voyons ce qui se passe si on prend les lignes définies par toutes les coordonnées constantes sauf une, par exemple la première : $x^1 = \lambda$, $x^k = \text{constante}$ pour tout $k \geq 1$. Alors l'action du vecteur vitesse associé à cette courbe sur un champ scalaire est

$$\vec{v}_1(f) = \partial_1 f \quad (\text{VI.6})$$

Nous notons par convention $\vec{\partial}_1$ ce vecteur, qui se définit par son action sur f : $\vec{\partial}_1(f) = \partial_1 f$. De même, on peut définir $\vec{\partial}_k$ pour tout k entre 1 et n . Cet ensemble, $(\vec{\partial}_k)$, est ce qu'on appelle la **base naturelle** de l'espace tangent, associée à la carte (x^k) . Nous avons représenté la base naturelle dans le cas d'une surface régulière figure VI.3. Notez que si l'on note $\psi = \phi_i^{-1}$, alors nous avons tout simplement

$$\vec{\partial}_k = \frac{\vec{\partial}\psi}{\partial x^k} \quad (\text{VI.7})$$

ce qui montre que la définition est cohérente avec celle des vecteurs de base de l'espace tangent d'une surface, associée à un paramétrage régulier. Le fait qu'on utilise un système de coordonnées, ou une carte, sous-entend toujours implicitement que nous disposons localement d'un paramétrage régulier qui établit un difféomorphisme entre un ouvert de \mathbb{R}^n et \mathcal{M} . Comme ce sera toujours le cas, sauf exception signalée, nous n'en ferons plus mention.

La base naturelle nous permet de définir l'espace tangent en tout point. Cependant c'est insatisfaisant car en ces termes, l'espace tangent n'apparaît pas comme un objet intrinsèque attaché à la variété en un point, c'est-à-dire indépendant du choix de la carte. Il faut revenir à une définition plus générale qui peut être la suivante, s'inspirant de la définition des vecteurs de l'équation (VI.4) qui elle, est indépendante du choix de la carte.

Définition VI.1.2. Soit \mathcal{M} une variété différentiable de dimension n . Soit $\mathcal{P} \in \mathcal{M}$. On définit l'espace tangent à \mathcal{M} en \mathcal{P} comme étant la classe d'équivalence des arcs réguliers qui passent par \mathcal{P} et qui donc admettent un vecteur tangent en \mathcal{P} .

Une représentation de l'espace tangent est faite figure VI.4.

Construction algébrique de l'espace tangent

Il existe un point de vue plus algébrique mais équivalent à celui-ci. Celui-ci consiste à voir les vecteurs en un point donné comme des applications linéaires sur les champs scalaires évalués en ce point. Plus précisément, on a :

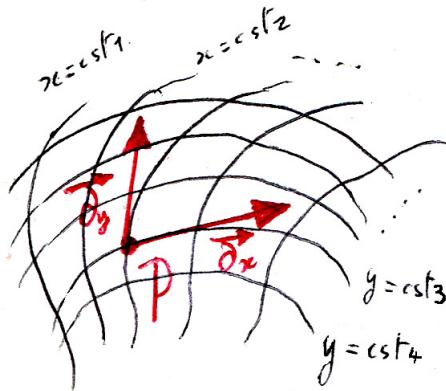


FIGURE VI.3 – Représentation de la base naturelle. On voit que les vecteurs de la base naturelle sont tangents au lignes des coordonnées.

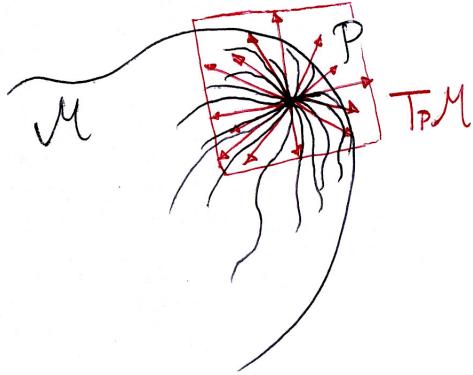


FIGURE VI.4 – Représentation de l'espace tangent en un point comme classe d'équivalence de courbes régulières admettant une tangente en un point.

Définition VI.1.3. Soit \mathcal{M} une variété différentiable, soit $\mathcal{P} \in \mathcal{M}$, soient f et g deux champs scalaires qui sont définis sur un voisinage de \mathcal{P} . On dit que \vec{v} est un vecteur de $T_{\mathcal{P}}\mathcal{M}$ s'il est une dérivation en $\mathcal{F}(\mathcal{P})$, l'ensemble des fonctions qui sont définies sur un voisinage de \mathcal{P} , c'est-à-dire que \vec{v} est une application linéaire de $\mathcal{F}(\mathcal{P})$ dans \mathbb{R} satisfait la règle de Leibniz

$$\vec{v}(fg)(\mathcal{P}) = (\vec{v}(f)(\mathcal{P}))g(\mathcal{P}) + f(\mathcal{P})\vec{v}(g)(\mathcal{P}) \quad (\text{VI.8})$$

À partir de cette définition, on voit immédiatement que les dérivations des fonctions constantes sont nulles : $\vec{v}(1) = \vec{v}(1 \times 1) = \vec{v}(1) + \vec{v}(1) = 2\vec{v}(1)$, d'où $\vec{v}(1) = 0$, puis par linéarité, $\vec{v}(c) = 0$. On montre ensuite que $\vec{\partial}_i$ est bien une dérivation. À partir de cette définition intrinsèque nous pouvons de plus montrer que c'est une base. Nous avons d'abord besoin d'un lemme.

Lemme VI.1.4. Soit f un champ scalaire défini sur un ouvert U de \mathbb{R}^n . Si f est différentiable sur une boule ouverte B centrée en $x_0 \in U$, alors il existe n fonctions f_i telles que

$$\forall x \in B, f(x) = f(x_0) + f_i(x)(x^i - x_0^i) \quad (\text{VI.9})$$

Preuve. Un calcul direct d'intégration montre que

$$f_i(x) = \int_0^1 \partial_i f(tx^1 + (1-t)x_0^1, \dots, tx^n + (1-t)x_0^n) dt \quad (\text{VI.10})$$

convient.

Théorème VI.1.5. $(\vec{\partial}_i)$ est une base de $T_{\mathcal{P}}\mathcal{M}$.

Preuve. C'est une famille libre : en effet, s'il existe n scalaires α^k tels que $\alpha^k \vec{\partial}_k = 0$, en appliquant ce vecteur à $f = x^j$, on obtient $\alpha^j = 0$ pour tout j .

C'est une famille génératrice : soit $\vec{v} \in T_{\mathcal{P}}\mathcal{M}$. On note $a^s = \vec{v}(x^s)$. Posons $\vec{y} = \vec{v} - a^k \vec{\partial}_k$. Soit un champ scalaire défini au voisinage de \mathcal{P} . Soit U une boule ouverte de $\phi_i(\mathcal{P}) = (x_{\mathcal{P}}^k)$. Alors dans cette boule ouverte, d'après le lemme, tout champ scalaire s'exprime ainsi : $f(x^k) = f(x_{\mathcal{P}}^k) + (x^k - x_{\mathcal{P}}^k)f_i(x^k)$. De cette sorte, on a

$$\vec{y}(f) = (x^k - x_{\mathcal{P}}^k)\vec{v}(f_k) + f_k \vec{v}(x^k - x_{\mathcal{P}}^k) - (x^k - x_{\mathcal{P}}^k)a^s \partial_s f_k - f_k a^s \partial_s(x^k - x_{\mathcal{P}}^k) \quad (\text{VI.11})$$

$$= (x^k - x_{\mathcal{P}}^k)\vec{v}(f_k) + f_k a^k - (x^k - x_{\mathcal{P}}^k)a^s \partial_s f_k - f_k a^k \quad (\text{VI.12})$$

que l'on doit évaluer en $x_{\mathcal{P}}^k$, ce qui donne 0. On en déduit $\vec{y} = \vec{0}$, soit $\vec{v} = a^k \vec{\partial}_k$, d'où le résultat.

On a donc construit de façon intrinsèque l'espace tangent à travers la notion de dérivation abstraite en un point, et en montrant que $(\vec{\partial}_k)$ était une base de l'ensemble des dérivations en \mathcal{P} , on a montré que ce point de vue était équivalent à celui de la classe d'équivalence des arcs paramétrés passant par \mathcal{P} . Cela montre d'ailleurs que l'espace tangent est de dimension n , la même dimension que \mathcal{M} , comme on pouvait s'y attendre.

Remarquons qu'en principe, à ce stade nous ne devrions pas noter $\vec{\partial}_k$ mais $\vec{\partial}_k(\mathcal{P})$ car les vecteurs sont toujours définis localement sur l'espace tangent en un point. Nous choisissons d'alléger les notations et de sous-entendre la dépendance en \mathcal{P} de $\vec{\partial}_k$. De toute manière, nous serons souvent amenés à considérer plutôt des champs de vecteurs que des vecteurs en un point : tout comme sur les surfaces régulières, on peut généraliser la notion de fibré tangent et considérer des champs de vecteurs.

Définition VI.1.6. Le fibré tangent de \mathcal{M} , noté $T\mathcal{M}$, est l'ensemble des champs de vecteurs sur \mathcal{M} :

$$T\mathcal{M} = \{(x, \vec{v}) | x \in \mathcal{M}, \vec{v} \in T_x\mathcal{M}\}. \quad (\text{VI.13})$$

Transformation contravariante des coordonnées par changement de carte

Les vecteurs sont des objets intrinsèques. Ils ne dépendent pas de la carte, puisqu'ils sont définis ou bien par un arc paramétré intrinsèque passant par un point, ou bien par leur action de dérivation sur les champs scalaires en un point donné (les deux points de vue étant équivalents). L'équivalence du point de vue algébrique et géométrique implique que pour établir des propriétés algébriques sur les vecteurs, on pourra toujours se référer à leur action sur un champ scalaire. Par exemple, voyons comment varient les coordonnées d'un vecteur lorsque l'on passe d'une carte à l'autre. Soient deux cartes (x^i) et (X^i) qui se recoupent sur \mathcal{M} . Comme \vec{v} est un objet intrinsèque, il doit pouvoir s'exprimer dans les deux cartes

$$\vec{v} = v^i \frac{\vec{\partial}}{\partial x^i} = V^i \frac{\vec{\partial}}{\partial X^i} \quad (\text{VI.14})$$

Nous savons que là où les cartes se recoupent, l'application de recollement est un difféomorphisme. Nous pouvons donc écrire la règle de la chaîne

$$\frac{\partial}{\partial X^i} = \frac{\partial X^i}{\partial x^j} \frac{\partial}{\partial x^j} \quad (\text{VI.15})$$

et par l'équivalence de la définition géométrique des vecteurs à celle de leur action sur les champs scalaires, nous pouvons en déduire

$$\frac{\vec{\partial}}{\partial X^i} = \frac{\partial x^j}{\partial X^i} \frac{\vec{\partial}}{\partial x^j} \quad (\text{VI.16})$$

soit

$$\vec{v} = V^i \frac{\partial x^j}{\partial X^i} \frac{\vec{\partial}}{\partial x^j} \quad (\text{VI.17})$$

Par unicité de la décomposition sur la base, il vient

$$v^i = V^j \frac{\partial x^i}{\partial X^j} \quad (\text{VI.18})$$

ceci est la loi de transformation des vecteurs par changement de carte. Comme elles varient de façon contraire aux coordonnées de la carte, on dit que ces coordonnées sont "contravariantes". Certains anciens ouvrages définissent les vecteurs comme des tableaux de chiffres qui se transforment selon cette loi. C'est le cas du (néanmoins merveilleux) livre de théorie classique des champs de Landau & Lifchitz. Le caractère géométrique et intrinsèque des vecteurs y est moins évident qu'ici, où la loi de transformation des coordonnées est déduite de la définition intrinsèque des vecteurs.

Exercices-exemples

1. \mathbb{R}^3 est une variété différentiable dont la carte constituée des coordonnées cartésiennes $(x^i) = (x, y, z)$ est un atlas. Les vecteurs de la base naturelle sont $\forall (x, y, z) \in \mathbb{R}^3, \vec{\partial}_x = (1, 0, 0), \vec{\partial}_y = (0, 1, 0), \vec{\partial}_z = (0, 0, 1)$. Comme \mathbb{R}^3 est plat et centré en 0, c'est un espace vectoriel et pour les coordonnées cartésiennes, les vecteurs de base ne varient pas. Il se trouve qu'on a ici $\vec{\partial}_i = \vec{e}_i$ où \vec{e}_i est la base canonique de \mathbb{R}^n . Bien entendu ça n'est pas le cas en général.
2. On définit les coordonnées sphériques sur $\mathbb{R}^3 \setminus (\mathbb{R} \vec{\partial}_z)$ (on modifie l'origine de l'ensemble de φ pour avoir tous les méridiens) :

$$x = r \sin \theta \cos \varphi \quad (\text{VI.19})$$

$$y = r \sin \theta \sin \varphi \quad (\text{VI.20})$$

$$z = r \cos \theta \quad (\text{VI.21})$$

pour $r > 0, 0 < \theta < \pi$, et alternativement $0 < \varphi < 2\pi$ puis $-\pi < \varphi < \pi$. Calculer les coordonnées de $\vec{\partial}_r, \vec{\partial}_\theta, \vec{\partial}_\varphi$ en fonction des vecteurs de la base naturelle des coordonnées cartésiennes, puis des vecteurs de la base orthonormée locale des coordonnées sphériques. Commenter.

Solution : on étudie l'action de chacun de ces vecteurs sur un champ scalaire quelconque. Ou, plus rapidement, on utilise la règle de la chaîne. Par exemple :

$$\partial_r = \partial_r x \partial_x + \partial_r y \partial_y + \partial_r z \partial_z = \sin \theta \cos \varphi \partial_x + \sin \theta \sin \varphi \partial_y + \cos \theta \partial_z \quad (\text{VI.22})$$

d'où

$$\vec{\partial}_r = \sin \theta \cos \varphi \vec{\partial}_x + \sin \theta \sin \varphi \vec{\partial}_y + \cos \theta \vec{\partial}_z \quad (\text{VI.23})$$

De même on a

$$\vec{\partial}_\theta = r \cos \theta \cos \varphi \vec{\partial}_x + r \cos \theta \sin \theta \vec{\partial}_y - r \sin \theta \vec{\partial}_z \quad (\text{VI.24})$$

$$\vec{\partial}_\varphi = -r \sin \theta \sin \varphi \vec{\partial}_x + r \sin \theta \cos \varphi \vec{\partial}_y \quad (\text{VI.25})$$

On remarque que

$$\vec{\partial}_r = \vec{e}_r, \quad \vec{\partial}_\theta = r \vec{e}_\theta, \quad \vec{\partial}_\varphi = r \sin \theta \vec{e}_\varphi \quad (\text{VI.26})$$

Ainsi la base naturelle n'est pas nécessairement orthonormée. Nous avons déjà rencontré cela dans les surfaces régulières, où les vecteurs de base de l'espace tangents formés par les dérivées partielles du paramétrage par rapport aux deux coordonnées n'étaient pas nécessairement orthonormés. On représente les coordonnées sphériques et les différentes bases figure VI.5.

VI.1.3 Dualité : formes linéaires et espace cotangent

Construction de l'espace dual

À tout espace vectoriel on peut faire correspondre un espace dual, l'ensemble des formes linéaires sur cet espace. L'espace tangent ne fait pas exception à cette règle, ainsi :

Définition VI.1.7. On note $T_x^* \mathcal{M}$ l'espace cotangent à \mathcal{M} en x , l'ensemble des formes linéaires sur $T_x \mathcal{M}$.

Et de même on peut définir des champs de formes linéaires sur \mathcal{M} .

Définition VI.1.8. Le fibré cotangent de \mathcal{M} , noté $T^* \mathcal{M}$, est l'ensemble des champs formes linéaires sur \mathcal{M} :

$$T^* \mathcal{M} = \{(x, \underline{\omega}) | x \in \mathcal{M}, \underline{\omega} \in T_x^* \mathcal{M}\} \quad (\text{VI.27})$$

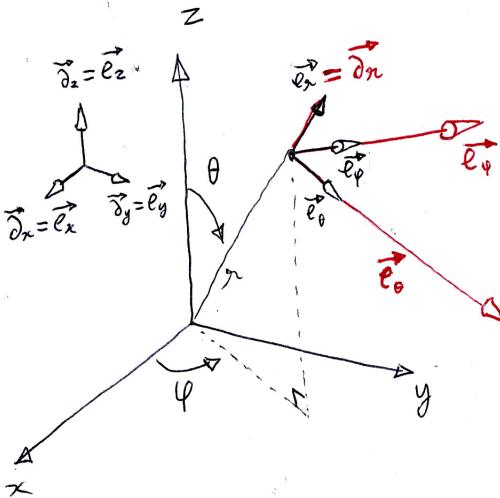


FIGURE VI.5 – Coordonnées sphériques. En rouge, les vecteurs de la base naturelle. On voit qu'ils ne sont pas égaux aux vecteurs de la base orthonormée locale.

Vous devez commencer à comprendre les choix de nos notations. Tous les éléments en gras sont des éléments du fibré. S'il y a une flèche au dessus, ce sont des vecteurs. S'ils sont soulignés, ce sont des formes linéaires.

Essayons de construire une base de $T_x^*\mathcal{M}$. La forme linéaire la plus élémentaire par rapport à la base naturelle a déjà été définie dans le chapitre sur les bases du calcul différentiel :

Définition VI.1.9. Soit $\vec{v} \in T_x\mathcal{M}$. On représente \vec{v} dans sa base naturelle : $\vec{v} = v^i \vec{\partial}_i$. On appelle k -ème projecteur, ou forme coordonnée k , et on note \underline{dx}^k , la forme linéaire dont le graphe est

$$\underline{dx}^k(\vec{v}) = v^k \quad (\text{VI.28})$$

Nous avons notamment $\underline{dx}^k(\vec{\partial}_s) = \delta_s^k$ qui vaut 1 si $k = s$, 0 sinon. Essayons d'en donner une interprétation géométrique. Voyons encore une fois comment varie un champ scalaire localement le long d'un arc. Si nous notons symboliquement

$$\gamma(\lambda + \varepsilon) = P + \delta \vec{P} + o(\|\delta \vec{P}\|) = \gamma(\lambda) + \varepsilon \vec{\gamma}'(\lambda) + o(\varepsilon) \quad (\text{VI.29})$$

nous avons

$$f(P + \delta \vec{P}) = f(P) + \varepsilon \gamma'(P) \partial_i f(P) + o(\varepsilon) \quad (\text{VI.30})$$

$$= f(P) + \varepsilon \partial_i f(P) \underline{dx}^i(\vec{\gamma}'(\lambda)) + o(\varepsilon) \quad (\text{VI.31})$$

$$= f(P) + \partial_i f(P) \underline{dx}^i(\delta \vec{P}) + o(\varepsilon) \quad (\text{VI.32})$$

nous retrouvons le résultat classique que pour un champ scalaire, $df = \partial_i f \underline{dx}^i$. Les formes coordonnées vont donc indiquer quelles sont les variations de la coordonnée considérée lorsqu'un point varie localement. De même que les vecteurs de la base naturelle, les formes coordonnées dépendent de la carte. Nous représentons figure VI.6 l'action des formes coordonnées sur la variation locale d'un arc paramétré.

Théorème VI.1.10. (\underline{dx}^k) est une base de $T^*\mathcal{M}$.

Preuve. C'est une famille libre. En effet, s'il existe n scalaires a_k tels que $a_k \underline{dx}^k = \underline{0}$, alors en appliquant cette forme à $\vec{\partial}_s$, on trouve $a_s = 0$. D'autre part on sait que la dimension de $T^*\mathcal{M}$ est n . (\underline{dx}^k) est une famille libre de n éléments de $T^*\mathcal{M}$, c'est donc une base.

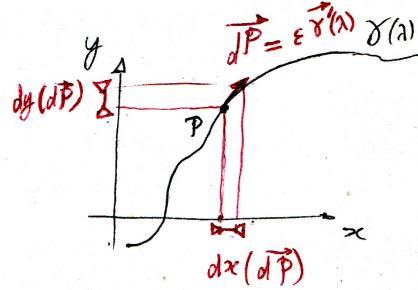


FIGURE VI.6 – Action des formes coordonnées sur la variation locale d'un arc paramétré. Au vecteur $\vec{d}P$, elles associent les variations induites sur chaque coordonnée de la carte choisie.

Pour toute forme linéaire $\underline{\omega}$, dès lors qu'on se donne une carte (x^k) , il existe n champs scalaires ω_k tels que $\underline{\omega} = \omega_k \underline{dx}^k$. Par exemple, la différentielle d'un champ scalaire est une forme linéaire : $df = \partial_i f \underline{dx}^i$. Son application sur un champ vectoriel est tout simplement la dérivée de f le long de ce champ, mais c'est aussi égal au champ vectoriel appliqué à ce champ : $df(\vec{v}) = D_{\vec{v}}f = \vec{v}(f)$. Attention cependant, la réciproque est fausse : toute forme linéaire n'est pas nécessairement la différentielle d'un champ scalaire. Nous y reviendrons au chapitre suivant pour voir à quelle condition une forme différentielle (objet plus général) est la différentielle d'une autre.

Transformation covariante des coordonnées par changement de carte

Il est naturel de se demander comment varient les composantes d'une forme linéaire lorsqu'on change de carte via l'application de recollement. Considérons deux cartes (x^k) et (X^k) dont les ensembles images ont une intersection non vide et ouverte. Sur ce domaine, l'application de recollement est un C^1 difféomorphisme, on peut donc considérer sans danger que chaque coordonnée X^i est une fonction différentiable de x^j et que la donnée des n fonctions $X^i(x^j)$ constitue bien un C^1 difféomorphisme qui n'est autre que l'application de recollement. Dans ce cas, on a

$$\underline{dX}^i(x^j) = \frac{\partial X^i}{\partial x_j} \underline{dx}^j \quad (\text{VI.33})$$

De même, on sait que toute forme linéaire se décompose de façon unique sur chaque base. Comme elle est définie de façon intrinsèque, on a nécessairement

$$\underline{\omega} = \omega_k \underline{dx}^k = \Omega_m \underline{dX}^m = \Omega_m \frac{\partial X^m}{\partial x_q} \underline{dx}^q \quad (\text{VI.34})$$

d'où l'on déduit

$$\omega_k = \Omega_m \frac{\partial X^m}{\partial x_k} \quad (\text{VI.35})$$

C'est ainsi que se transforment les composantes d'une forme linéaire. On constate que la loi est contraire à celle des vecteurs. Comme les composantes des formes linéaires varient en même temps que les coordonnées de la carte, on dit qu'elles sont "covariantes".

VI.1.4 Les tenseurs en toute généralité

Prélude : dual du dual

En guise de préliminaire, posons nous une question absurde en apparence : quel est l'espace dual de $T^*\mathcal{M}$? Quel est le dual du dual? En fait on peut identifier $T^{**}\mathcal{M}$ à $T\mathcal{M}$. En effet, un vecteur peut également se considérer comme forme linéaire sur l'espace dual en considérant le graphe de $\underline{\omega}$ à l'envers :

$$\begin{aligned} \vec{v} &\in T^{**}\mathcal{M} : T^*\mathcal{M} \rightarrow \mathbb{R} \\ \underline{\omega} &\mapsto \underline{\omega}(\vec{v}) \end{aligned} \quad (\text{VI.36})$$

Cela revient à écrire $\underline{\vec{v}} = \langle \cdot, \vec{v} \rangle$. On voit le lien avec les notations de la mécanique quantique qui identifient un "vecteur d'état" avec son action sur les formes linéaires : $\vec{\psi} = |\psi\rangle$. Il est évident que défini ainsi, $\underline{\vec{v}}$ a toutes les bonnes propriétés d'un objet de $T^{**}\mathcal{M}$ (on peut s'amuser à refaire toutes les démonstrations de l'algèbre linéaire classique "à l'envers"). On identifiera désormais $T^{**}\mathcal{M}$ et $T\mathcal{M}$. Rappelons les trois points de vue équivalents que nous avons vus jusqu'ici pour considérer un vecteur :

1. Un vecteur est la tangente d'une courbe en un point (point de vue géométrique).
2. Un vecteur se définit par son action sur un champ scalaire en un point (point de vue algébrique).
3. Un vecteur se définit par son action sur une forme linéaire en un point (point de vue dual du dual).

Fugue : construction des tenseurs

Définition VI.1.11. Une forme multilinéaire de rang s est une application de $(T_x\mathcal{M})^s$ dans \mathbb{R} linéaire par rapport à chacun de ses arguments :

$$\omega : T_x\mathcal{M} \times \dots \times T_x\mathcal{M} \rightarrow \mathbb{R}, (\vec{u}_1, \dots, \vec{u}_s) \mapsto \omega(\vec{u}_1, \dots, \vec{u}_s) \quad (\text{VI.37})$$

Et de là nous sommes armés pour définir ce qu'est un tenseur en toute généralité :

Définition VI.1.12. Un tenseur de type (r, s) , ou r fois contravariant et s fois covariant, est une forme multilinéaire définie sur $(T_x^*\mathcal{M})^r \times (T_x\mathcal{M})^s$. On note $T_x(r, s)\mathcal{M}$ l'ensemble des tenseurs de type (r, s) au point $x \in \mathcal{M}$, et tout simplement $T_s^r\mathcal{M}$ l'ensemble des champs de tenseurs de type (r, s) . $k + s$ est ce qu'on appelle le rang du tenseur.

Remarques

1. Un vecteur est un tenseur de type $(1, 0)$.
2. Une forme linéaire est un tenseur de type $(0, 1)$.
3. Par convention, un champ scalaire est un champ tensoriel de type $(0, 0)$.
4. Le terme "rang d'un tenseur" est ambigu. En effet, il ne faut pas confondre le rang d'un tenseur qui dépend de son type, avec le rang d'un tenseur au sens rang d'une application linéaire, c'est-à-dire la dimension de l'espace image. Par exemple, la forme linéaire nulle définie sur $(T_x\mathcal{M})^{10000}$ a un rang de 10000 au sens défini plus haut, mais comme son espace image est le singleton 0, son rang au sens de l'algèbre linéaire classique est 0.

Avant de trouver une base de l'ensemble des tenseurs, nous avons besoin de définir le produit tensoriel.

Définition VI.1.13. Le produit tensoriel de deux tenseurs de types (r, s) et (p, q) est un tenseur de type $(r + p, s + q)$ défini ainsi. Soient $\omega \in T_s^r\mathcal{M}$, $\mu \in T_q^p\mathcal{M}$. On note $\omega \otimes \mu$ le produit tensoriel de ω et μ qui vérifie

$$\begin{aligned} \omega \otimes \mu : (T^*\mathcal{M})^r \times (T\mathcal{M})^s \times (T^*\mathcal{M})^p \times (T\mathcal{M})^q &\rightarrow \mathbb{R} \\ (\underline{\omega}_1, \dots, \underline{\omega}_r, \vec{v}_1, \dots, \vec{v}_s, \underline{\omega}_{r+1}, \dots, \underline{\omega}_{r+p}, \vec{v}_{s+1}, \dots, \vec{v}_{s+q}) &\mapsto \omega(\underline{\omega}_1, \dots, \underline{\omega}_r, \vec{v}_1, \dots, \vec{v}_s) \mu(\underline{\omega}_{r+1}, \dots, \underline{\omega}_{r+p}, \vec{v}_{s+1}, \dots, \vec{v}_{s+q}) \end{aligned} \quad (\text{VI.38})$$

Remarques

1. Exemple : pour $\omega, \mu \in T^*\mathcal{M}$, on a

$$\forall \vec{u}, \vec{v} \in T\mathcal{M}, (\underline{\omega} \otimes \underline{\mu})(\vec{u}, \vec{v}) = \underline{\omega}(\vec{u}) \underline{\mu}(\vec{v}) \quad (\text{VI.39})$$

2. Autre exemple : le tenseur $\vec{v} \otimes \underline{u}$ est défini sur $T^*\mathcal{M} \times T\mathcal{M}$:

$$\forall (\underline{\alpha}, \vec{\beta}) \in T^*\mathcal{M} \times T\mathcal{M}, \vec{v} \otimes \underline{u}(\underline{\alpha}, \vec{\beta}) = \vec{v}(\underline{\alpha}) \underline{u}(\vec{\beta}) = \underline{\alpha}(\vec{v}) \underline{u}(\vec{\beta}) \quad (\text{VI.40})$$

Et on peut combiner ainsi à l'infini toute sortes de formes, vecteurs, tenseurs. La seule limite est votre imagination !

3. Attention : \otimes ne commute pas ! Par exemple, dans $\mathcal{M} = \mathbb{R}^2$, $\underline{\mathbf{dx}} \otimes \underline{\mathbf{dy}} - \underline{\mathbf{dy}} \otimes \underline{\mathbf{dx}} \neq \mathbf{0}$. En effet, deux vecteurs non colinéaires et non nuls \vec{u} et \vec{v} ont un déterminant non nul dans \mathbb{R}^2 , donc

$$(\underline{\mathbf{dx}} \otimes \underline{\mathbf{dy}} - \underline{\mathbf{dy}} \otimes \underline{\mathbf{dx}})(\vec{u}, \vec{v}) = u^x v^y - u^y v^x \neq 0 \quad (\text{VI.41})$$

En toute généralité, deux tenseurs a et b ne vérifient pas nécessairement $a \otimes b = b \otimes a$.

Désormais nous avons suffisamment d'outils pour déterminer une base de $T_s^r \mathcal{M}$.

Théorème VI.1.14. Munissons \mathcal{M} d'une carte (x^k) . Alors

$$\left(\vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{\mathbf{dx}}^{j_1} \otimes \dots \otimes \underline{\mathbf{dx}}^{j_s} \right)_{(i_1, \dots, i_r, j_1, \dots, j_s) \in [\![1, n]\!]^{r+s}} \quad (\text{VI.42})$$

est une base de $T_s^r \mathcal{M}$.

Preuve. On constate tout simplement qu'on a

$$\vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{\mathbf{dx}}^{j_1} \otimes \dots \otimes \underline{\mathbf{dx}}^{j_s} \left(\underline{\mathbf{dx}}^{p_1}, \dots, \underline{\mathbf{dx}}^{p_r}, \vec{\partial}_{q_1}, \dots, \vec{\partial}_{q_s} \right) = \prod_{k=1}^r \prod_{m=1}^s \delta_{i_k}^{p_k} \delta_{j_m}^{q_m} \quad (\text{VI.43})$$

De là, prouver que la famille proposée est libre et génératrice est évident.

Ainsi, dans une base naturelle, on pourra définir les coordonnées d'un tenseur comme ses composantes sur la base :

$$\mathbf{T} = T^{i_1 \dots i_r}_{j_1 \dots j_s} \vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{\mathbf{dx}}^{j_1} \otimes \dots \otimes \underline{\mathbf{dx}}^{j_s} \quad (\text{VI.44})$$

Il faut prendre garde à ne pas oublier qu'avec la convention de sommation des indices répétés en haut et en bas, cette équation sous entend une $(r+s)$ -uple somme qui s'écrirait, sans cette convention :

$$\mathbf{T} = \sum_{i_1=1}^n \dots \sum_{i_r=1}^n \sum_{j_1=1}^n \dots \sum_{j_s=1}^n T^{i_1 \dots i_r}_{j_1 \dots j_s} \vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{\mathbf{dx}}^{j_1} \otimes \dots \otimes \underline{\mathbf{dx}}^{j_s} \quad (\text{VI.45})$$

Les scalaires $T^{i_1 \dots i_r}_{j_1 \dots j_s}$ sont les composantes de \mathbf{T} dans la carte (x^k) . En particulier, on a

$$T^{i_1 \dots i_r}_{j_1 \dots j_s} = \mathbf{T} \left(\underline{\mathbf{dx}}^{p_1}, \dots, \underline{\mathbf{dx}}^{p_r}, \vec{\partial}_{q_1}, \dots, \vec{\partial}_{q_s} \right) \quad (\text{VI.46})$$

Sur les tenseurs de rang 1, on vérifie en effet que pour $\underline{\omega} = \omega_k \underline{\mathbf{dx}}^k$, on a bien

$$\underline{\omega} \left(\vec{\partial}_m \right) = \omega_k \underline{\mathbf{dx}}^k \left(\vec{\partial}_m \right) = \omega_k \delta_m^k = \omega_m \quad (\text{VI.47})$$

et pour $\vec{v} = v^k \vec{\partial}_k$, on a bien

$$\vec{v} \left(\underline{\mathbf{dx}}^m \right) = \vec{v} \left(\underline{\mathbf{dx}}^m \right) = \underline{\mathbf{dx}}^m \left(\vec{v} \right) = \underline{\mathbf{dx}}^m \left(v^k \vec{\partial}_k \right) = v^k \underline{\mathbf{dx}}^m \left(\vec{\partial}_k \right) = v^k \delta_k^m = v^m \quad (\text{VI.48})$$

Exercice Soient r formes linéaires $\underline{\omega}^a = \omega_k^a \underline{\mathbf{dx}}^k$, $a \in [\![1, r]\!]$, et s vecteurs $\vec{v}_b = v_b^k \vec{\partial}_k$, $b \in [\![1, s]\!]$, et soit un tenseur de type (r, s) avec les mêmes notations que précédemment : $\mathbf{T} = T^{i_1 \dots i_r}_{j_1 \dots j_s} \vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{\mathbf{dx}}^{j_1} \otimes \dots \otimes \underline{\mathbf{dx}}^{j_s}$. Calculer $\mathbf{T} (\underline{\omega}^1, \dots, \underline{\omega}^r, \vec{v}_1, \dots, \vec{v}_s)$ en fonction des coordonnées de chaque objet.

Solution : c'est un exercice de calcul brutal qui utilise la multilinéarité de \mathbf{T} par rapport à chacun de ses arguments. On pourrait tout calculer directement mais faisons d'abord avec la base de $T_s^r \mathcal{M}$. On a

$$\begin{aligned} & \vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{\mathbf{dx}}^{j_1} \otimes \dots \otimes \underline{\mathbf{dx}}^{j_s} \left(\omega_k^1 \underline{\mathbf{dx}}^k, \dots, \omega_m^s \underline{\mathbf{dx}}^m, v_1^p \vec{\partial}_p, \dots, v_s^q \vec{\partial}_q \right) \\ &= \vec{\partial}_{i_1} \left(\omega_k^1 \underline{\mathbf{dx}}^k \right) \dots \vec{\partial}_{i_r} \left(\omega_m^s \underline{\mathbf{dx}}^m \right) \underline{\mathbf{dx}}^{j_1} \left(v_1^p \vec{\partial}_p \right) \dots \underline{\mathbf{dx}}^{j_s} \left(v_s^q \vec{\partial}_q \right) \end{aligned} \quad (\text{VI.49})$$

$$= \omega_k^1 \dots \omega_m^s v_1^p \dots v_s^q \vec{\partial}_{i_1} \left(\underline{\mathbf{dx}}^k \right) \dots \vec{\partial}_{i_r} \left(\underline{\mathbf{dx}}^m \right) \dots \underline{\mathbf{dx}}^{j_1} \left(\vec{\partial}_p \right) \underline{\mathbf{dx}}^{j_s} \left(\vec{\partial}_q \right) \quad (\text{VI.50})$$

$$= \omega_k^1 \dots \omega_m^r v_1^p \dots v_s^q \delta_{i_1}^k \dots \delta_{i_r}^m \delta_{j_1}^p \dots \delta_{j_s}^q \quad (\text{VI.51})$$

$$= \omega_{i_1}^1 \dots \omega_{i_r}^r v_1^{j_1} \dots v_s^{j_s} \quad (\text{VI.52})$$

Connaissant l'action d'un élément de base, il suffit de faire la combinaison linéaire avec les coordonnées de \mathbf{T} , ce qui donne tout simplement

$$\mathbf{T}(\underline{\omega}^1, \dots, \underline{\omega}^r, \vec{v}_1, \dots, \vec{v}_s) = T^{i_1 \dots i_r}_{j_1 \dots j_s} \omega_{i_1}^1 \dots \omega_{i_r}^r v_1^{j_1} \dots v_s^{j_s} \quad (\text{VI.53})$$

Définition VI.1.15. On définit la contraction entre un tenseur de type (r, s) et un vecteur, ou une forme linéaire, le tenseur de type respectivement $(r, s - 1)$ ou $(r - 1, s)$ défini ainsi :

$$\vec{u} \dashv \mathbf{T} = \mathbf{T}(\cdot, \dots, \cdot, \vec{u}, \dots, \cdot) \quad (\text{VI.54})$$

où \vec{u} apparaît à la $r + 1$ -ème position, et

$$\underline{\omega} \dashv \mathbf{T} = \mathbf{T}(\underline{\omega}, \dots) \quad (\text{VI.55})$$

Les composantes de $\vec{u} \dashv \mathbf{T}$ sont $T^{i_1 \dots i_r}_{k j_2 \dots j_s} v^k$ et les composantes de $\underline{\omega} \dashv \mathbf{T}$ sont $T^{k i_2 \dots i_r}_{j_1 \dots j_s} \omega_k$. Quitte à réordonner les indices (ou les arguments) de \mathbf{T} , on peut très bien définir la contraction sur l'argument que l'on souhaite. En fait, on peut généraliser cette définition et contracter deux tenseurs de rangs quelconques sur les arguments que l'on souhaite, par exemple pour deux tenseurs \mathbf{T} et \mathbf{S} on peut très bien considérer le tenseur de composantes $T^{k m n \dots i_r}_{j_1 \dots j_{s-1} q} S^{a_1 q}_{b_1 b_{c-1} k m n}$. Ici encore la seule limite est notre imagination. On peut également contracter un tenseur sur ses propres indices, par exemple considérer le tenseur de composantes $T^{k i_2 \dots i_r}_{k j_2 \dots j_s}$.

Toccata : transformation contravariante et covariante des coordonnées d'un tenseur

On peut maintenant se demander comment varient les composantes d'un tenseur quand on change de carte. Pour cela il suffit de regarder comment se transforme les vecteurs de base de $T_s^r \mathcal{M}$. Notons

$$t = t^{i_1 \dots i_r}_{j_1 \dots j_s} \frac{\vec{\partial}}{\partial x^{i_1}} \otimes \dots \otimes \frac{\vec{\partial}}{\partial x^{i_r}} \otimes \underline{\mathbf{d}\mathbf{x}}^{j_1} \otimes \dots \otimes \underline{\mathbf{d}\mathbf{x}}^{j_s} \quad (\text{VI.56})$$

$$= T^{k_1 \dots k_r}_{m_1 \dots m_s} \frac{\vec{\partial}}{\partial X^{k_1}} \otimes \dots \otimes \frac{\vec{\partial}}{\partial X^{k_r}} \otimes \underline{\mathbf{d}\mathbf{X}}^{m_1} \otimes \dots \otimes \underline{\mathbf{d}\mathbf{X}}^{m_s} \quad (\text{VI.57})$$

$$(\text{VI.58})$$

On cherche à exprimer, par exemple, $t^{k_1 \dots k_r}_{m_1 \dots m_s}$ en fonction de $T^{i_1 \dots i_r}_{j_1 \dots j_s}$. Pour ce faire, il suffit de se rappeler que

$$\frac{\vec{\partial}}{\partial X^k} = \frac{\partial x^j}{\partial X^k} \frac{\vec{\partial}}{\partial x^j} \quad (\text{VI.59})$$

et

$$\underline{\mathbf{d}\mathbf{X}}^k = \frac{\partial X^k}{\partial x^j} \underline{\mathbf{d}\mathbf{x}}^j \quad (\text{VI.60})$$

et en injectant cela dans le deuxième membre de l'équation (VI.57) et en identifiant aux éléments de base, on trouve la loi de transformation des coordonnées des tenseurs de type (r, s) :

$$t^{k_1 \dots k_r}_{m_1 \dots m_s} = T^{i_1 \dots i_r}_{j_1 \dots j_s} \frac{\partial x^{k_1}}{\partial X^{i_1}} \dots \frac{\partial x^{k_r}}{\partial X^{i_r}} \frac{\partial X^{j_1}}{\partial x^{m_1}} \dots \frac{\partial X^{j_s}}{\partial x^{m_s}} \quad (\text{VI.61})$$

Ici on voit que les r premières composantes sont contravariantes comme les composantes des vecteurs, et que les s dernières sont covariantes, comme les composantes des formes linéaires. Les vieux ouvrages définissent les tenseurs de type (k, s) comme des tableaux de chiffres qui se transforment selon la loi énoncée ci-dessus. Ici, nous avons déduit la loi de transformation des composantes du tenseur en partant de sa définition intrinsèque.

VI.1.5 Le tenseur métrique

Une variété différentielle peut être munie d'une métrique intrinsèque qui définira le produit scalaire sur \mathcal{M} . Cela revient à munir \mathcal{M} d'une première forme fondamentale sans que celle-ci ne soit nécessairement induite par une métrique euclidienne d'un espace plus gros¹. Dans le cas des variétés munies de métriques, on parlera de variétés métrées, ou de variétés riemanniennes.

1. Cependant, Nash a démontré que toute variété différentiable de dimension n de classe C^3 est plongeable de façon isométrique dans un espace euclidien de dimension $n(n+1)(3n+11)/2$.

Propriétés du tenseur métrique

Définition VI.1.16. On dit que \mathcal{M} est munie d'une métrique \mathbf{g} , et l'on parle de variété riemannienne $(\mathcal{M}, \mathbf{g})$, si \mathbf{g} est un champ tensoriel de type $(0, 2)$ qui vérifie

- \mathbf{g} est symétrique : $\forall \vec{u}, \vec{v} \in T\mathcal{M}, \mathbf{g}(\vec{u}, \vec{v}) = \mathbf{g}(\vec{v}, \vec{u})$.
- \mathbf{g} est non dégénérée définie positive : $\forall \vec{u} \in T\mathcal{M}, \mathbf{g}(\vec{u}, \vec{u}) \geq 0$ et $\mathbf{g}(\vec{u}, \vec{u}) = 0 \Rightarrow \vec{u} = \vec{0}$.
- En relativité générale, on peut affaiblir l'hypothèse précédente. On travaille alors dans des variétés de dimension 4 avec des métriques qui ont une signature $(-, +, +, +)$, c'est-à-dire qu'en tout point, il existe une carte telles que les coordonnées de \mathbf{g} soient $\text{diag}(-1, +1, +1, +1)$. Nous préciserons à chaque fois si nous travaillons avec une métrique définie positive ou la métrique de la relativité générale.

Remarques.

1. La bilinéarité de \mathbf{g} découle immédiatement de sa nature tensorielle. Elle a donc toutes les bonnes propriétés pour être une première forme fondamentale.
2. Si \mathbf{g} est définie positive, alors $\vec{u} \mapsto \sqrt{\mathbf{g}(\vec{u}, \vec{u})}$ définit une norme sur $T\mathcal{M}$.
3. \mathbf{g} définit un produit scalaire si elle est définie positive, ou ce qu'on appelle un pseudo-produit scalaire en relativité générale. Désormais le produit scalaires de vecteurs se fera toujours au sens de \mathbf{g} , c'est-à-dire que l'on aura par définition

$$\forall \vec{u}, \vec{v} \in T\mathcal{M}, \vec{u} \cdot \vec{v} = \mathbf{g}(\vec{u}, \vec{v}) \quad (\text{VI.62})$$

4. On peut décomposer \mathbf{g} sur la base des formes coordonnées : $\mathbf{g} = g_{ij} \underline{\mathbf{dx}}^i \otimes \underline{\mathbf{dx}}^j$. Par définition, on a donc, dans toute carte locale, $g_{ij} = \mathbf{g}(\vec{\partial}_i, \vec{\partial}_j) = \vec{\partial}_i \cdot \vec{\partial}_j$. On retrouve ce qu'on avait déjà déduit de la première forme fondamentale pour les surfaces plongées dans l'espace 3D euclidien : en deux dimensions on a donc $E = g_{uu} = \vec{\partial}_u \cdot \vec{\partial}_u$, $F = g_{vv} = \vec{\partial}_v \cdot \vec{\partial}_v$, et $G = \vec{\partial}_v \cdot \vec{\partial}_u$.
5. Les coordonnées du tenseur métrique se transforment comme celles de tout bon tenseur de type $(0, 2)$. Si l'on a $\mathbf{g} = g_{ab} \underline{\mathbf{dx}}^a \otimes \underline{\mathbf{dx}}^b = G_{ij} \underline{\mathbf{dX}}^i \otimes \underline{\mathbf{dX}}^j$, alors

$$G_{ij} = g_{mn} \frac{\partial x^i}{\partial X^m} \frac{\partial x^j}{\partial X^n} \quad (\text{VI.63})$$

Exercice : calculer la métrique de \mathbb{R}^3 en coordonnées cartésiennes, puis en coordonnées sphériques.
Réponse : $g_{\text{cart}} = \text{diag}(1, 1, 1)$ et $g_{\text{spher}} = \text{diag}(1, r^2, r^2 \sin^2 \theta)$.

Définition VI.1.17. Quand \mathbf{g} est définie positive, on dit que (\vec{e}_i) est une base orthonormée de $T_x\mathcal{M}$ si $\mathbf{g}(\vec{e}_i, \vec{e}_j) = \delta_{ij} = 1$ si $i = j$, 0 sinon. Si la signature de \mathbf{g} est $(-, +, \dots)$, on aura $\mathbf{g}(\vec{e}_1, \vec{e}_1) = -1$, et $\forall j > 1, \mathbf{g}(\vec{e}_i, \vec{e}_j) = \delta_{ij}$.

Comme \mathbf{g} est diagonalisable définie positive, il est toujours possible de trouver une carte telle qu'en P , les vecteurs de la base naturelle sont une base orthonormée, autrement dit qu'on puisse exprimer les composantes du tenseur métrique comme celles de la métrique de Minkowski en coordonnées cartésiennes. Cela constitue en fait le principe d'équivalence qui postule que localement, on peut toujours se ramener à un système de référence dans lequel les lois de la physique sont celles du cadre théorique de la relativité restreinte.

Définition VI.1.18. Soit $\gamma : I \rightarrow \mathcal{M}$ un arc paramétré sur \mathcal{M} . La distance parcourue le long de l'arc entre deux paramètres $t_1, t_2 \in I$ vaut par définition

$$\ell_\gamma(t_1, t_2) = \int_{t_1}^{t_2} \sqrt{\pm \mathbf{g}(\vec{\gamma}'(t), \vec{\gamma}'(t))} dt \quad (\text{VI.64})$$

Il faudra mettre un plus sous la racine carrée pour toutes les trajectoires des variétés à métrique définie positive, et moins pour les particules matérielles en relativité générale dont la trajectoire vérifie toujours $\|\vec{\gamma}'(t)\|^2 < 0$.

Cette définition est naturelle au sens de la norme définie par le tenseur métrique. En effet, on a

$$\|\gamma(t + \varepsilon) - \gamma(t)\|^2 = \varepsilon^2 \mathbf{g}(\vec{\gamma}'(t), \vec{\gamma}'(t)) + o(\varepsilon^3) \quad (\text{VI.65})$$

de telle sorte que l'abscisse curviligne sera définie par $ds = \sqrt{\pm \mathbf{g}(\vec{\gamma}'(t), \vec{\gamma}'(t))} dt$. En termes de coordonnées, en notant $\vec{\gamma}'(t) = \vec{v} = v^i \vec{\partial}_i$, on aura $ds = \sqrt{\pm g_{ij} v^i v^j} dt$.

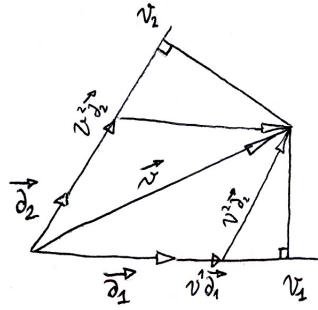


FIGURE VI.7 – Coordonnées covariantes et contravariantes dans la base naturelle quand la métrique est définie positive. Quand celle-ci n'est pas orthonormée, les coordonnées covariantes et contravariantes ne sont pas égales. Quand la métrique est de signature $(-, +, +, +)$, l'orthonormalité de la base n'assure même pas l'égalité des coordonnées covariantes et contravariantes en raison de la géométrie hyperbolique qui lie l'espace et le temps (voir tout bon livre de relativité restreinte).

Dualité métrique entre vecteurs et formes

Les coefficients g_{ij} forment une matrice symétrique définie positive (ou en relativité, diagonalisable avec des valeurs propres non nulles), donc inversible. On note g^{ij} les composantes de l'inverse de cette matrice. Ainsi, on a $g^{ab}g_{bc} = \delta_c^a$. Munis de cette notation, voyons si le tenseur métrique nous permet de passer de $T\mathcal{M}$ à $T^*\mathcal{M}$ de façon privilégiée. On sait que dans un espace muni d'un produit scalaire, à tout vecteur \vec{v} correspond une unique forme linéaire définie par $\underline{v} = \langle \vec{v}, \cdot \rangle$. Ici on peut faire exactement la même chose avec le tenseur métrique et définir $\underline{v} = \mathbf{g}(\vec{v}, \cdot)$. \underline{v} est le dual métrique de \vec{v} . On a en particulier

$$\underline{v}(\vec{\partial}_i) = \mathbf{g}(\vec{v}, \vec{\partial}_i) = g_{ab}v^a\delta_i^b = g_{ai}v^a \quad (\text{VI.66})$$

de telle sorte que si l'on décompose \underline{v} sur la base des formes coordonnées ainsi $\underline{v} = v_i \underline{dx}^i$, on a nécessairement

$$v_i = g_{ai}v^a \quad (\text{VI.67})$$

On dit alors que v_i sont les coordonnées covariantes de \vec{v} définies par dualité métrique. On peut interpréter l'équation précédente comme une équation matricielle. On peut inverser cette équation et écrire que

$$v^a = g^{ai}v_i \quad (\text{VI.68})$$

On peut évidemment tout faire dans l'autre sens. Soit $\underline{\omega} = \omega_i \underline{dx}^i$ une forme linéaire. Alors on peut lui faire correspondre un unique vecteur $\vec{\omega} = \omega^i \vec{\partial}_i$ qui est tel que $\underline{\omega} = \mathbf{g}(\vec{\omega}, \cdot)$. Par le même raisonnement on trouvera que $\omega_i = g_{ai}\omega^a$ et que $\omega^i = g^{ai}\omega_a$.

Remarquez que l'on a $v_i = \vec{v} \cdot \vec{\partial}_i$ et $v^i = \underline{dx}^i(\vec{v})$. On peut alors visualiser à quoi correspondent géométriquement les coordonnées covariantes et contravariantes d'un objet de $T\mathcal{M}$, si on le considère comme un vecteur, ou de $T^*\mathcal{M}$, si on le considère comme une forme. Nous représentons cela à la figure VI.7 quand la base naturelle n'est pas orthonormée.

Musique tensorielle : dualité métrique des tenseurs

De même on peut passer des coordonnées covariantes aux contravariantes et réciproquement pour tout type de tenseur.

Définition VI.1.19. Soit $\mathbf{T} \in T_s^r \mathcal{M}$. On appelle dual métrique de \mathbf{T} sur le k -ème argument contravariante le tenseur de $T_{s+1}^{r-1} \mathcal{M}$ noté \mathbf{T}^\flat dont le graphe vérifie

$$\mathbf{T}^\flat(\underline{\omega}_1, \dots, \vec{\omega}_k, \dots, \underline{\omega}_r, \vec{u}_1, \dots, \vec{u}_s) = \mathbf{T}(\underline{\omega}_1, \dots, \mathbf{g}(\vec{\omega}_k, \cdot), \dots, \underline{\omega}_r, \vec{u}_1, \dots, \vec{u}_s) \quad (\text{VI.69})$$

et de même on appelle dual métrique de \mathbf{T} sur le p -ème argument covariant le tenseur de $T_{s-1}^{r+1} \mathcal{M}$ le tenseur noté \mathbf{T}^\sharp dont le graphe vérifie

$$\mathbf{T}^\sharp(\underline{\omega}_1, \dots, \underline{\omega}_r, \vec{u}_1, \dots, \underline{u}_p, \dots, \vec{u}_s) = \mathbf{T}(\underline{\omega}_1, \dots, \dots, \underline{\omega}_r, \vec{u}_1, \dots, \vec{u}_p, \dots, \vec{u}_s) \quad (\text{VI.70})$$

où l'on a $\underline{u}_p = \mathbf{g}(\vec{u}_p, \cdot)$.

Exercice : calculer les coordonnées de \mathbf{T}^\flat et \mathbf{T}^\sharp en fonctions de celles de \mathbf{T} que l'on notera respectivement $T^{a_1 \dots a_{k-1}}{}_{a_k}{}^{a_{k+1} \dots a_r}{}_{b_1 \dots b_s}$, $T^{a_1 \dots a_r}{}_{b_1 \dots b_{p-1}}{}^{b_p}{}_{b_{p+1} \dots b_s}$ et $T^{a_1 \dots a_r}{}_{b_1 \dots b_s}$.

Solution : un calcul simple montre que

$$T^{a_1 \dots a_{k-1}}{}_{a_k}{}^{a_{k+1} \dots a_r}{}_{b_1 \dots b_s} = g_{a_k q_k} T^{a_1 \dots a_{k-1} q_k a_{k+1} \dots a_r}{}_{b_1 \dots b_s} \quad (\text{VI.71})$$

$$T^{a_1 \dots a_r}{}_{b_1 \dots b_{p-1}}{}^{b_p}{}_{b_{p+1} \dots b_s} = g^{b_p q_p} T^{a_1 \dots a_r}{}_{b_1 \dots b_{p-1} q_p b_{p+1} \dots b_s} \quad (\text{VI.72})$$

Pour des tenseurs de rang moins élevés, par exemple pour $\mathbf{T} = T_{ab} \mathbf{dx}^a \otimes \mathbf{dx}^b$, on aura typiquement $T_a^b = g^{ai} T_{ib}$, $T_a^b = g^{bi} T_{ai}$, $T^{ab} = g^{ai} g^{bj} T_{ij}$ etc. Notez que comme \mathbf{g} est symétrique, l'ordre des indices dans g^{ai} n'a pas d'importance. On aura également, dans l'autre sens $T_{ab} = g_{ai} T_b^i$, etc. (le reste est laissé à titre d'exercice).

Ces relations sont vraies en particulier pour le tenseur métrique lui-même, de quoi l'on peut déduire une propriété intéressante :

$$g_b^a = g_a^b = g^{ai} g_{ib} = \delta_b^a \quad (\text{VI.73})$$

dans tous les systèmes de coordonnées.

Certains ouvrages parlent d'isomorphismes musicaux pour passer des indices contravariants aux covariants ou le contraire ; on utilise la notation \flat pour faire descendre un indice et \sharp pour le faire monter, par dualité métrique. On aurait pu noter $\underline{\omega}^\sharp = \vec{\omega}$ et $\vec{v}^\flat = \underline{v}$ mais nous pensons que les notations soulignées et fléchées sont plus parlantes.

VI.2 Algèbre extérieure

Ce chapitre sera l'occasion d'approfondir un type particulier de tenseurs : les formes différentielles, autrement appelées les k -formes extérieures. Ce sont tout simplement les formes k -linéaires antisymétriques. Elles présentent un grand intérêt en physique car elles permettent de généraliser la notion de déterminant dans des variétés et permettent d'établir une théorie de l'intégration (et donc de la dérivation) dans les variétés et les sous-variétés. Cela sera infiniment utile pour faire des sommes continues en relativité générale. L'objet de ce chapitre est de présenter ces objets en tant que tel et de déjà apercevoir quelques illustrations physiques de ces objets. Le chapitre suivant traîtera spécifiquement de la dérivation et l'intégration sur les variétés.

VI.2.1 En guise de prélude : les 2-formes

Propriétés des 2-formes

Définition VI.2.1. Une 2-forme extérieure, ou tout simplement une 2-forme, est une forme bilinéaire (ou un tenseur de type $(0, 2)$) alternée si \mathbf{b} est une 2-forme, alors

$$\forall \vec{u} \in T_x \mathcal{M}, \mathbf{b}(\vec{u}, \vec{u}) = 0 \quad (\text{VI.74})$$

Proposition VI.2.2. Toute forme bilinéaire alternée est antisymétrique. Autrement dit, si \vec{b} est une 2-forme, on a

$$\forall \vec{u}, \vec{v} \in T_x \mathcal{M}, \mathbf{b}(\vec{u}, \vec{v}) = -\mathbf{b}(\vec{v}, \vec{u}) \quad (\text{VI.75})$$

Preuve. Soit \mathbf{b} une 2-forme. Soient $\vec{u}, \vec{v} \in T_x \mathcal{M}$.

$$\mathbf{b}(\vec{u} - \vec{v}, \vec{u} - \vec{v}) = 0 = \mathbf{b}(\vec{u}, \vec{u}) - \mathbf{b}(\vec{u}, \vec{v}) - \mathbf{b}(\vec{v}, \vec{u}) + \mathbf{b}(\vec{v}, \vec{v}) = -\mathbf{b}(\vec{u}, \vec{v}) - \mathbf{b}(\vec{v}, \vec{u}) \quad (\text{VI.76})$$

Définition VI.2.3. Par convention, les formes linéaires (ou tenseurs de type $(0, 1)$) sont des 1-formes, et les champs scalaires sont des 0-formes. On note $\mathcal{A}_0 \mathcal{M}$ les champs scalaires, $\mathcal{A}_1 \mathcal{M}$ les champs de formes linéaires, et $\mathcal{A}_2 \mathcal{M}$ les champs de 2-formes sur \mathcal{M} .

Définition VI.2.4. Soient $\underline{\omega}, \underline{\nu} \in \mathcal{A}_1 \mathcal{M}$. On définit le produit extérieur de $\underline{\omega}$ et $\underline{\nu}$ l'élément de $\mathcal{A}_2 \mathcal{M}$ défini ainsi

$$\underline{\omega} \wedge \underline{\nu} = \underline{\omega} \otimes \underline{\nu} - \underline{\nu} \otimes \underline{\omega} \quad (\text{VI.77})$$

Remarque : on a

$$\underline{\omega} \wedge \underline{\nu}(\vec{u}, \vec{v}) = \det \begin{pmatrix} \underline{\omega}(\vec{u}) & \underline{\nu}(\vec{u}) \\ \underline{\omega}(\vec{v}) & \underline{\nu}(\vec{v}) \end{pmatrix} \quad (\text{VI.78})$$

Théorème VI.2.5. Soit une variété \mathcal{M} de dimension n . Soit x^k une carte générique de \mathcal{M} . $(\underline{dx}^i \wedge \underline{dx}^j)_{1 \leq i < j \leq n}$ est une base de $A_2\mathcal{M}$.

Preuve. Montrons que la famille est génératrice. Soit $b \in A_2\mathcal{M}$. Soient $\vec{u} = u^i \vec{\partial}_i$ et $\vec{v} = v^i \vec{\partial}_i$ deux éléments de $T\mathcal{M}$. On a

$$b(\vec{u}, \vec{v}) = u^k v^\ell b(\vec{\partial}_k, \vec{\partial}_\ell) \quad (\text{VI.79})$$

mais comme $b(\vec{\partial}_k, \vec{\partial}_\ell) = -b(\vec{\partial}_\ell, \vec{\partial}_k)$ et $b(\vec{\partial}_k, \vec{\partial}_k) = 0$, on a

$$b(\vec{u}, \vec{v}) = \sum_{1 \leq k < \ell \leq n} (u^k v^\ell - u^\ell v^k) b(\vec{\partial}_k, \vec{\partial}_\ell) = \sum_{1 \leq k < \ell \leq n} b(\vec{\partial}_k, \vec{\partial}_\ell) \underline{dx}^k \wedge \underline{dx}^\ell \quad (\text{VI.80})$$

ce qui montre bien que la famille annoncée engendre $A_2\mathcal{M}$.

Montrons que la famille est libre. Soient $\alpha_{k\ell} \in \mathbb{R}^{n(n-1)/2}$ tel que $\sum_{1 \leq k < \ell \leq n} \alpha_{k\ell} \underline{dx}^k \wedge \underline{dx}^\ell = 0$. En appliquant cette forme à $(\vec{\partial}_i, \vec{\partial}_j)$ avec $1 \leq i < j \leq n$, on trouve $\alpha_{ij} = 0$.

Corrigé VI.2.6. $A_2\mathcal{M}$ est un espace vectoriel de dimension $n(n-1)/2$.

Donnons une interprétation géométrique des 2-formes.

- Si $\mathcal{M} = \mathbb{R}^2$, $\underline{dx} \wedge \underline{dy} = \det(\cdot, \cdot)$. C'est donc le volume du parallélogramme formé par les deux vecteurs en argument.
- Pour une variété quelconque, dans l'espace tangent en un point M , $\underline{dx}^i \wedge \underline{dx}^j(\vec{u}, \vec{v})$ est l'aire algébrique du parallélogramme formé par la projection des vecteurs \vec{u} et \vec{v} sur le plan affine d'origine M et engendré par $\text{Vect}(\vec{\partial}_i, \vec{\partial}_j)$.

Le champ électromagnétique comme 2-forme

Nous allons donner un exemple de champ de 2-formes particulièrement utile en physique : la forme de Faraday, ou autrement dit, le tenseur électro-magnétique.

Commençons par observer que ce qu'on appelle vulgairement le "vecteur champ électrique" et le "vecteur champ magnétique" ne sont pas des vecteurs au sens des champs tensoriels dans les variétés. C'est particulièrement évident pour le champ électromagnétique. En effet, considérons une particule au repos subissant un champ magnétique non nul. Elle subit une force de Lorentz nulle :

$$\vec{F} = q \vec{0} \times \vec{B} = \vec{0} \quad (\text{VI.81})$$

À présent plaçons nous dans un référentiel en translation rectiligne uniforme de vitesse \vec{U} par rapport au mobile au repos. Alors du point de vue de cet observateur, étant donné l'invariance des lois de la physique, il faut que la force soit également nulle malgré la vitesse, c'est-à-dire que du point de vue de cet observateur, on a

$$\vec{F} = \vec{0} = q(\vec{E}' - \vec{U} \times \vec{B}') \quad (\text{VI.82})$$

Il apparaît ici nécessaire d'avoir $\vec{E}' \neq \vec{0}$ pour que le mobile continue d'être au repos. Pourtant, la transformation de Lorentz comme de Galilée sont des C^1 -difféomorphismes de \mathbb{R}^4 dans \mathbb{R}^4 . Si \vec{E} était un vecteur au sens défini dans le chapitre sur les tenseurs, alors ses composantes devraient se transformer selon la loi linéaire des transformations contravariantes des coordonnées d'un vecteur. Aucune transformation linéaire ne peut transformer le vecteur nul en un vecteur non nul. \vec{E} n'est donc pas un vecteur. De la même manière, considérons une spire circulaire traversée par un courant électrique dans le plan $x0y$, centrée en $(0, 0)$. Le champ magnétique en $(0, 0, 0)$ sera dirigé vers $+\vec{\partial}_z$. À présent effectuons une symétrie centrale du système dans le plan $x0y$. Bien que l'axe z soit invariant, le sens du courant a été inversé, par conséquent, le champ magnétique pointe dorénavant vers $-\vec{\partial}_z$. Si le champ magnétique était un vecteur, sa composante z aurait été invariante par une telle transformation. Le champ magnétique n'est donc pas un vecteur au sens des champs tensoriels.

Les "vecteurs" champ électrique et champ magnétique ne sont donc pas les meilleurs objets. Nous voudrions pouvoir travailler avec des tenseurs. Il existe un objet plus intrinsèque qui englobe tout le champ électromagnétique et c'est avec cet objet que nous allons travailler.

À ce stade du cours nous n'avons pas les moyens de montrer qu'il s'agit de la seule forme possible du champ électromagnétique, nous nous contenterons de le donner "dogmatiquement". Nous montreront tout à la fin de ce chapitre pourquoi le champ électromagnétique a cette forme². Le but de cette anticipation est pédagogique : il vous permettra de manipuler une 2-forme et de vous habituer à ce formalisme avant de passer aux k -formes.

Munissons l'espace de Minkowski \mathcal{E} d'une carte cartésienne $(x^\mu) = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$. Ici, les lettres grecques noteront les indices allant de 0 à 3 et les indices latins indiqueront les lettres allant de 1 à 3. La 1-forme travail du champ électromagnétique s'écrit

$$\underline{\mathbf{E}} = E_i \underline{\mathbf{d}x}^i \quad (\text{VI.83})$$

Et la 2-forme flux magnétique s'écrit

$$\phi = B_x \underline{\mathbf{d}y} \wedge \underline{\mathbf{d}z} + B_y \underline{\mathbf{d}z} \wedge \underline{\mathbf{d}x} + B_z \underline{\mathbf{d}x} \wedge \underline{\mathbf{d}y} = B_i \underline{\mathbf{d}S}^i \quad (\text{VI.84})$$

où $\underline{\mathbf{d}S}^x = \underline{\mathbf{d}y} \wedge \underline{\mathbf{d}z}$, $\underline{\mathbf{d}S}^y = \underline{\mathbf{d}z} \wedge \underline{\mathbf{d}x}$, et $\underline{\mathbf{d}S}^z = \underline{\mathbf{d}x} \wedge \underline{\mathbf{d}y}$ sont les formes d'aire. La forme de Faraday s'écrit

$$\underline{\mathbf{F}} = \underline{\mathbf{E}} \wedge c \underline{\mathbf{dt}} + \phi = E_i \underline{\mathbf{d}x}^i \wedge \underline{\mathbf{d}x}^0 + B_i \underline{\mathbf{d}S}^i \quad (\text{VI.85})$$

On aimeraient établir les lois de transformation du champ électromagnétique lorsque nous changeons d'observateur inertiel. À ce titre considérons le changement de carte associé à une transformation de Lorentz le long de l'axe x : posons $x^\mu = \Lambda^\mu_\nu X^\nu$ où l'on a posé

$$\Lambda^\mu_\nu = \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{VI.86})$$

la matrice de Lorentz³. Voyons comment se transforment les formes élémentaires. Nous avons,

$$c \underline{\mathbf{dt}} = \cosh\varphi \underline{\mathbf{d}X}^0 + \sinh\varphi \underline{\mathbf{d}X}, \quad \underline{\mathbf{d}x} = \sinh\varphi c \underline{\mathbf{dT}} + \cosh\varphi \underline{\mathbf{d}X}, \quad \underline{\mathbf{d}y} = \underline{\mathbf{d}Y}, \quad \underline{\mathbf{d}z} = \underline{\mathbf{d}Z} \quad (\text{VI.87})$$

De quoi nous déduisons

$$E_x \underline{\mathbf{d}x} \wedge c \underline{\mathbf{dt}} = E_x (\sinh\varphi c \underline{\mathbf{dT}} + \cosh\varphi \underline{\mathbf{d}X}) \wedge (\cosh\varphi \underline{\mathbf{d}X}^0 + \sinh\varphi \underline{\mathbf{d}X}) = E_x \underline{\mathbf{d}X} \wedge c \underline{\mathbf{dT}} \quad (\text{VI.88})$$

où nous avons utilisé le fait que $\underline{\mathbf{d}X}^i \wedge \underline{\mathbf{d}X}^i = \mathbf{0}$ pour tout i , que $\underline{\mathbf{d}X} \wedge \underline{\mathbf{d}T} = -\underline{\mathbf{d}T} \wedge \underline{\mathbf{d}X}$, et que $\cosh^2\varphi - \sinh^2\varphi = 1$. Par le même type de calcul on trouve

$$E_y \underline{\mathbf{d}y} \wedge c \underline{\mathbf{dt}} = \cosh\varphi E_y \underline{\mathbf{d}Y} \wedge c \underline{\mathbf{dT}} - \sinh\varphi E_y \underline{\mathbf{d}X} \wedge \underline{\mathbf{d}Y} \quad (\text{VI.89})$$

$$E_z \underline{\mathbf{d}z} \wedge c \underline{\mathbf{dt}} = \cosh\varphi E_z \underline{\mathbf{d}Z} \wedge c \underline{\mathbf{dT}} + \sinh\varphi E_z \underline{\mathbf{d}Z} \wedge \underline{\mathbf{d}X} \quad (\text{VI.90})$$

$$B_x \underline{\mathbf{d}y} \wedge \underline{\mathbf{d}z} = B_x \underline{\mathbf{d}Y} \wedge \underline{\mathbf{d}Z} \quad (\text{VI.91})$$

$$B_y \underline{\mathbf{d}z} \wedge \underline{\mathbf{d}x} = \sinh\varphi B_y \underline{\mathbf{d}Z} \wedge c \underline{\mathbf{dT}} + \cosh\varphi B_y \underline{\mathbf{d}Z} \wedge \underline{\mathbf{d}X} \quad (\text{VI.92})$$

$$B_z \underline{\mathbf{d}x} \wedge \underline{\mathbf{d}y} = -\sinh\varphi B_z \underline{\mathbf{d}Y} \wedge c \underline{\mathbf{dT}} + \cosh\varphi B_z \underline{\mathbf{d}X} \wedge \underline{\mathbf{d}Y} \quad (\text{VI.93})$$

En rassemblant les termes sur les composantes des vecteurs de base de $\mathcal{A}_2\mathcal{E}$, on trouve que

$$\begin{aligned} \underline{\mathbf{F}} = & E_x \underline{\mathbf{d}X} \wedge \underline{\mathbf{d}Y} + (\cosh\varphi E_y - \sinh\varphi B_z) \underline{\mathbf{d}Y} \wedge c \underline{\mathbf{dT}} + (\cosh\varphi E_z + \sinh\varphi B_y) \underline{\mathbf{d}Z} \wedge c \underline{\mathbf{dT}} \\ & + B_x \underline{\mathbf{d}Y} \wedge \underline{\mathbf{d}Z} + (\cosh\varphi B_y + \sinh\varphi E_z) \underline{\mathbf{d}Z} \wedge \underline{\mathbf{d}X} + (\cosh\varphi B_z - \sinh\varphi E_y) \underline{\mathbf{d}X} \wedge \underline{\mathbf{d}Y} \end{aligned} \quad (\text{VI.94})$$

2. Nous avons en effet besoin du dual de Hodge, qui intervient assez tard dans cette leçon.

3. Pour exprimer cette matrice en fonction de V la vitesse relative des référentiels et c la vitesse de la lumière, il suffit de se rappeler que dans le référentiel mobile, l'origine se déplace à $x(t) = Vt$, d'où $\frac{x}{ct} = \frac{\sinh\varphi}{\cosh\varphi} = \operatorname{th}\varphi$. De là on tire $\frac{\cosh^2\varphi - \sinh^2\varphi}{\cosh^2\varphi} = 1 - \operatorname{th}\varphi = 1 - \frac{V^2}{c^2}$ et donc $\cosh\varphi = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$, et $\sinh\varphi = \operatorname{th}\varphi \cosh\varphi = \frac{V/c}{\sqrt{1 - \frac{V^2}{c^2}}}$.

On sait que dans la carte (X^μ) , l'observateur va ressentir le champ électromagnétique \vec{E} et \vec{B} tels que

$$\begin{aligned} \mathbf{F} = & E_X \underline{\mathbf{d}\mathbf{X}} \wedge \underline{\mathbf{c}\mathbf{d}\mathbf{T}} + E_Y \underline{\mathbf{d}\mathbf{Y}} \wedge \underline{\mathbf{c}\mathbf{d}\mathbf{T}} + E_Z \underline{\mathbf{d}\mathbf{Z}} \wedge \underline{\mathbf{c}\mathbf{d}\mathbf{T}} \\ & + B_X \underline{\mathbf{d}\mathbf{Y}} \wedge \underline{\mathbf{d}\mathbf{Z}} + B_Y \underline{\mathbf{d}\mathbf{Z}} \wedge \underline{\mathbf{d}\mathbf{X}} + B_Z \underline{\mathbf{d}\mathbf{X}} \wedge \underline{\mathbf{d}\mathbf{Y}} \end{aligned} \quad (\text{VI.95})$$

par unicité de la décomposition sur les éléments de base, on en déduit la loi de transformation du champ électromagnétique

$$E_X = E_x, \quad E_Y = \cosh\varphi E_y - \sinh\varphi B_z, \quad E_Z = \cosh\varphi E_z + \sinh\varphi B_y \quad (\text{VI.96})$$

$$B_X = B_x, \quad B_Y = \cosh\varphi B_y + \sinh\varphi E_z, \quad B_Z = \cosh\varphi B_z - \sinh\varphi E_y \quad (\text{VI.97})$$

On voit ici que le champ électrique et le champ magnétique n'ont pas des composantes qui se transforment comme celles d'un vecteur. Les champs électriques et magnétiques sont liés : lors d'un changement d'observateur, une partie du champ magnétique se retrouve dans le champ électrique et réciproquement. On ne peut donc plus les dissocier, c'est pourquoi dans la physique moderne⁴ on parle d'électromagnétisme. Par ailleurs, le tenseur \mathbf{F} a des composantes qui se transforment comme un "vrai" tenseur. On aurait trouvé le même résultat en appliquant les formules du chapitre 6. Ici, comme la transformation est linéaire, la matrice jacobienne est la matrice de Lorentz elle-même. Le tenseur électromagnétisme est donc le seul objet intrinsèque au sens des variétés, c'est en étudiant cet objet qu'on a pu déduire les lois de transformation des composantes des champs électrique et magnétique.

VI.2.2 Formes multilinéaires alternées, ou p -formes

Définition VI.2.7. Une forme p -linéaire antisymétrique, ou une p -forme différentielle – plus simplement une p -forme –, est un tenseur de type $(0, p)$, alterné si ω est une p -forme, alors on a :

$$\forall \vec{u} \in \mathcal{T}\mathcal{M}, \omega(\dots, \vec{u}, \dots, \vec{u}, \dots) = \mathbf{0} \quad (\text{VI.98})$$

ou, ce qui revient au même, ω est antisymétrique :

$$\forall \vec{u}, \vec{v} \in \mathcal{T}\mathcal{M}, \omega(\dots, \vec{u}, \dots, \vec{v}, \dots) = -\omega(\dots, \vec{v}, \dots, \vec{u}, \dots) \quad (\text{VI.99})$$

On note $\mathcal{A}_p T_x \mathcal{M}$ l'ensemble des p -formes sur $T_x \mathcal{M}$, et $\mathcal{A}_p \mathcal{M}$ l'ensemble des champs de p -formes sur $T \mathcal{M}$.

Remarques.

1. Soit (\vec{e}_i) une base de $T_x \mathcal{M}$ (cela peut être la base naturelle dans une carte donnée, mais pas nécessairement) et soit ω une forme p -linéaire sur l'espace tangent en x ($\omega \in T_x \binom{0}{p} \mathcal{M}$). Soient $\omega_{i_1 \dots i_p}$ les coordonnées de ω dans la base $\underline{e}^{i_1} \otimes \dots \otimes \underline{e}^{i_p}$ où $\underline{e}^i = g(\vec{e}_i, \cdot)$. On a $\omega = \omega_{i_1 \dots i_p} \underline{e}^{i_1} \otimes \dots \otimes \underline{e}^{i_p}$. Autrement dit, $\omega_{i_1 \dots i_p} = \omega(\vec{e}_{i_1}, \dots, \vec{e}_{i_p})$. Alors on a la propriété suivante : ω est un élément de $\mathcal{A}_p T_x \mathcal{M}$ si et seulement si ses coordonnées sont antisymétriques par rapport à leurs indices, c'est-à-dire si et seulement si $\omega_{\dots i \dots j \dots} = -\omega_{\dots j \dots i \dots}$. Les anciens ouvrages se contentent de définir les p -formes comme des tableaux de chiffres qui se transforment comme les tenseurs par changement de carte et qui sont antisymétriques par rapport à leurs indices. Ils parlent tout simplement de "tenseurs antisymétriques".
2. Si \mathcal{M} est de dimension n , alors pour tout $p > n$, $\mathcal{A}_p \mathcal{M}$ ne contient que des formes nulles. En effet, si les n premiers arguments sont une famille libre, alors le $n+1$ -ème argument est nécessairement une combinaison linéaire des n premiers, de telle sorte que la famille est liée. Comme les p -formes sont linéaires et alternées, elles donneront nécessairement un résultat nul. Ainsi, on se limitera à l'étude des p -formes pour $p \leq n$.

Théorème VI.2.8. $\dim \mathcal{A}_p \mathcal{M} = C_n^p = \binom{n}{p} = \frac{n!}{p!(n-p)!}$

Preuve. Pour déterminer la dimension de $\mathcal{A}_p \mathcal{M}$, il suffit de savoir combien de nombres indépendants figurent dans les $\omega_{i_1 \dots i_p}$, composantes de $\omega \in \mathcal{A}_p \mathcal{M}$, à partir de la contrainte d'antisymétrie. Nous savons que $\omega_{i_1 \dots i_p}$ est nul si deux indices sont égaux. Le nombre de composantes non nulles est donc $n(n-1)\dots(n-p+1)$ si $p \leq n$, 0 si $p > n$. De plus, comme ω est antisymétrique, le nombre de composantes égales à $\omega_{i_1 \dots i_p}$ au signe près se déduit par le nombre de permutations de (i_1, \dots, i_p) , égal à $p!$. Le nombre de composantes indépendantes est donc égal à $\frac{n(n-1)\dots(n-p+1)}{p!}$.

4. Plus précisément, depuis le XIX^e siècle. Auparavant, les notions d'électricité et de magnétisme étaient séparées.

Ce théorème appelle quelques remarques.

1. Par convention, $\binom{n}{p} = 0$ si $p > n$, le théorème se suffit donc à lui-même pour tout $p \in \mathbb{N}$.
2. On note que $\dim \mathcal{A}_0 \mathcal{M} = \dim \mathcal{A}_n \mathcal{M} = 1$. Cela montre que l'ensemble des formes n -linéaires antisymétriques dans un espace de dimension n est un espace vectoriel de dimension 1. ce résultat très important permet de restreindre l'étude des formes n -linéaires dans les espaces vectoriels de dimension n à l'étude du déterminant et permet d'établir toutes les propriétés intéressantes de celui-ci. Par exemple, que le déterminant d'une application linéaire n'est qu'un multiplicateur de volume : $\det(f(\vec{u}_1), \dots, f(\vec{u}_n)) = \det f \det(\vec{u}_1, \dots, \vec{u}_n)$.
3. On remarque aussi que $\dim \mathcal{A}_p \mathcal{M} = \dim \mathcal{A}_{n-p} \mathcal{M}$. Il doit donc exister une dualité qui permet de passer de façon bijective de $\mathcal{A}_p \mathcal{M}$ à $\mathcal{A}_{n-p} \mathcal{M}$. Nous verrons cela plus loin dans le cours : il s'agit de la dualité de Hodge.

À présent voyons comment construire des p -formes différentielles à partir de q -formes où $q < p$. Pour cela nous avons besoin de poser du vocabulaire sur les permutations.

Définition VI.2.9. On note \mathcal{G}_n l'ensemble des permutations de $\llbracket 1, n \rrbracket$ (c'est-à-dire l'ensemble des bijections de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, n \rrbracket$). Soit $\sigma \in \mathcal{G}_n$. Toute permutation peut s'écrire comme succession d'échange de deux éléments en laissant les autres invariants (ce qu'on appelle des transpositions).

- Si le nombre de transposition est pair, on dit que σ est une permutation paire. Sinon on dit que σ est une permutation impaire.
- On note $k(\sigma)$ la signature de σ . Si σ est une permutation paire, $k(\sigma) = 1$, sinon $k(\sigma) = -1$.

En fait, on peut également définir les p -formes différentielles à partir des permutations :

Proposition VI.2.10.

$$\omega \in \mathcal{A}_p \mathcal{M} \Leftrightarrow \left\{ \begin{array}{l} \omega \in T_p^0 \mathcal{M} \\ \forall (\vec{v}_i) \in (T \mathcal{M})^p, \forall \sigma \in \mathcal{G}_p, \omega(\vec{v}_1, \dots, \vec{v}_p) = k(\sigma) \omega(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(n)}) \end{array} \right. \quad (\text{VI.100})$$

Armé de cet outillage, nous pouvons généraliser le produit extérieur aux p -formes.

Définition VI.2.11. Le produit d'une p -forme et d'une q -forme, noté \wedge est une $p+q$ -forme :

$$\begin{aligned} \wedge : \mathcal{A}_p \mathcal{M} \times \mathcal{A}_q \mathcal{M} &\rightarrow \mathcal{A}_{p+q} \mathcal{M} \\ (\omega, \mu) &\mapsto \omega \wedge \mu \end{aligned} \quad (\text{VI.101})$$

qui se définit par son action sur $p+q$ vecteurs. Soient $(\vec{v}_i) \in (T \mathcal{M})^{p+q}$, on a

$$\omega \wedge \mu(\vec{v}_1, \dots, \vec{v}_{p+q}) = \frac{1}{p!q!} \sum_{\sigma \in \mathcal{G}_{p+q}} k(\sigma) \omega(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(p)}) \mu(\vec{v}_{\sigma(p+1)}, \dots, \vec{v}_{\sigma(p+q)}) \quad (\text{VI.102})$$

Proposition VI.2.12.

- $\forall (\omega, \mu) \in \mathcal{A}_p \mathcal{M} \times \mathcal{A}_q \mathcal{M}, \omega \wedge \mu = (-1)^{pq} \mu \wedge \omega$ (anticommutativité).
- $\forall (\omega, \mu, \nu) \in \mathcal{A}_p \mathcal{M} \times \mathcal{A}_q \mathcal{M} \times \mathcal{A}_r \mathcal{M}, \omega \wedge (\mu \wedge \nu) = (\omega \wedge \mu) \wedge \nu = \omega \wedge \mu \wedge \nu$ (associativité).

Voyons quelques exemples. Pour les 1-formes, on a

$$\forall \underline{\omega}, \underline{\mu} \in \mathcal{A}_1 \mathcal{M}, \underline{\omega} \wedge \underline{\mu} = \underline{\omega} \otimes \underline{\mu} - \underline{\mu} \otimes \underline{\omega} \quad (\text{VI.103})$$

soit

$$\forall \vec{u}, \vec{v} \in T \mathcal{M}, \underline{\omega} \wedge \underline{\mu}(\vec{u}, \vec{v}) = \underline{\omega}(\vec{u}) \underline{\mu}(\vec{v}) - \underline{\omega}(\vec{v}) \underline{\mu}(\vec{u}) \quad (\text{VI.104})$$

Plus généralement, si $(\underline{\omega}_i) \in (\mathcal{A}_1 \mathcal{M})^p$, on a

$$\forall (\vec{v}_i) \in (T \mathcal{M})^p, \underline{\omega}_1 \wedge \dots \wedge \underline{\omega}_n(\vec{v}_1, \dots, \vec{v}_n) = \begin{vmatrix} \underline{\omega}_1(\vec{v}_1) & \dots & \underline{\omega}_p(\vec{v}_1) \\ \vdots & \ddots & \vdots \\ \underline{\omega}_1(\vec{v}_p) & \dots & \underline{\omega}_p(\vec{v}_p) \end{vmatrix} \quad (\text{VI.105})$$

Théorème VI.2.13. $(\underline{dx}^{i_1} \wedge \dots \wedge \underline{dx}^{i_p})$ avec $1 \leq i_1 < i_2 < \dots < i_p \leq n$, est une base de $\mathcal{A}_p \mathcal{M}$.

Preuve. Comme on a $\underline{dx}^{i_1} \wedge \dots \wedge \underline{dx}^{i_p}(\vec{\partial}_{i_1}, \dots, \vec{\partial}_{i_p}) = 1$, la famille proposée est libre. Par ailleurs, cette famille contient autant d'éléments qu'il y a de choix pour sélectionner p éléments parmi n , c'est-à-dire $\binom{n}{p}$, la dimension de $\mathcal{A}_p\mathcal{M}$. C'est donc une base.

Quelques remarques.

1. Le produit tensoriel de deux formes différentielles alternées n'est pas une forme alternée !
2. En dimension 3, les permutations circulaires sont paires. Attention ! elles sont impaires en dimension 4, donc en relativité.

VI.2.3 Le tenseur de Levi-Civita

Espace des n -formes comme espace vectoriel de dimension 1

Nous allons nous attarder sur $\mathcal{A}_n\mathcal{M}$ où $n = \dim\mathcal{M}$. On sait que $\dim\mathcal{A}_n\mathcal{M} = 1$, donc tous les éléments de $\mathcal{A}_n\mathcal{M}$ sont proportionnels. Si \mathcal{M} est munie d'une métrique \mathbf{g} , on peut définir une base orthonormale de $T_x\mathcal{M}$, que l'on notera $(\vec{e}_i) \in (T_x\mathcal{M})^n$ avec

$$\mathbf{g}(\vec{e}_i, \vec{e}_j) = \delta_{ij} \text{ si } \mathbf{g} \text{ est définie positive} \quad (\text{VI.106})$$

$$\mathbf{g}(\vec{e}_0, \vec{e}_0) = -1, \quad \forall (i, j) \in \llbracket 0, 3 \rrbracket \times \llbracket 1, 3 \rrbracket, \mathbf{g}(\vec{e}_i, \vec{e}_j) = \delta_{ij} \text{ en relativité.} \quad (\text{VI.107})$$

Pour tout élément non nul de $\mathcal{A}_n T_x \mathcal{M}$, on a nécessairement $\omega(\vec{e}_1, \dots, \vec{e}_n) \neq 0$. Comme $\dim \mathcal{A}_n T_x \mathcal{M} = 1$, il n'existe que deux n -formes distinctes qui vérifient respectivement $\lambda(\vec{e}_1, \dots, \vec{e}_n) = 1$ et $\nu(\vec{e}_1, \dots, \vec{e}_n) = -1$. Définir une orientation de l'espace $T_x\mathcal{M}$ revient à choisir l'une de ces deux formes comme référence. On dit que \mathcal{M} est orientable si on peut la munir d'un atlas et y choisir une orientation qui reste la même sur chaque carte et qui ne change pas sur les zones de recollement. Nous avons déjà vu des exemples avec les surfaces régulières. Désormais nous ne travaillerons qu'avec des variétés orientables.

Définition VI.2.14. Choisissons une base orthonormée $\vec{e}_1, \dots, \vec{e}_n$ en $T_x\mathcal{M}$. On appelle tenseur (ou forme) de Levi-Civita, associé à la métrique \mathbf{g} , et on note ε le tenseur de $\mathcal{A}_n T_x \mathcal{M}$ qui vérifie $\varepsilon(\vec{e}_1, \dots, \vec{e}_n) = 1$.

Pour l'instant cette définition est complètement arbitraire et dépend de la base orthonormée choisie. Nous allons voir qu'il n'en est rien grâce à la propriété suivante.

Proposition VI.2.15. Pour toute base orthonormée $(\vec{u}_1, \dots, \vec{u}_n)$ de $T_x\mathcal{M}$, $|\varepsilon((\vec{u}_1, \dots, \vec{u}_n)) = 1|$.

Preuve. Puisque $\underline{dx}^1 \wedge \underline{dx}^2 \wedge \dots \wedge \underline{dx}^n$ est une base de $\mathcal{A}_n T_x \mathcal{M}$, on a $\varepsilon = \omega_{12\dots n} \underline{dx}^1 \wedge \underline{dx}^2 \wedge \dots \wedge \underline{dx}^n$. S on se place dans la base orthonormée de départ $(\vec{e}_1, \dots, \vec{e}_n)$, on aura tout simplement

$$\varepsilon = \underline{e}^1 \wedge \dots \wedge \underline{e}^n = k(\sigma) \underline{e}^{\sigma(1)} \wedge \dots \wedge \underline{e}^{\sigma(n)} \quad (\text{VI.108})$$

où $\sigma \in \mathcal{G}_n$. Posons

$$[i_1, \dots, i_n] = \begin{cases} k(\sigma) & \text{si } (i_1, \dots, i_n) = \sigma(1, 2, \dots, n) \text{ avec } \sigma \in \mathcal{G}_n \\ 0 & \text{sinon} \end{cases} \quad (\text{VI.109})$$

On a $\varepsilon = [\sigma(1), \dots, \sigma(n)] \underline{e}^{\sigma(1)} \wedge \dots \wedge \underline{e}^{\sigma(n)}$. Ainsi, dans une base orthonormée, on a

$$\boxed{\varepsilon = \sum_{i_1, \dots, i_n} [i_1, \dots, i_n] \underline{e}^{i_1} \otimes \dots \otimes \underline{e}^{i_n} \quad ((\vec{e}_i) \text{ base orthonormée})} \quad (\text{VI.110})$$

Voyons ce qui se passe quand on passe dans une base quelconque. Soit P^μ_ν la matrice de passage de (\vec{e}_i) , la base orthonormée de départ, à (\vec{u}_i) , autre base orthonormée on a $\vec{u}_i = P^j_i \vec{e}_j$. On a

$$\varepsilon(\vec{u}_1, \dots, \vec{u}_n) = \underline{e}^1 \wedge \dots \wedge \underline{e}^n(\vec{u}_1, \dots, \vec{u}_n) = \sum_{\sigma \in \mathcal{G}_n} k(\sigma) \prod_{j=1}^n P_j^{\sigma(j)} = \det P \quad (\text{VI.111})$$

Or on sait que le déterminant de la matrice de passage entre deux base orthonormées est égal à ± 1 , d'où le résultat.

Cela nous permet de poser la définition suivante :

Définition VI.2.16. On dit qu'une base $(\vec{u}_1, \dots, \vec{u}_n)$ de $T_x\mathcal{M}$ est directe si $\varepsilon(\vec{u}_1, \dots, \vec{u}_n) > 0$, indirecte si $\varepsilon(\vec{u}_1, \dots, \vec{u}_n) < 0$. De même, on dit qu'une base orthonormée est directe si $\varepsilon(\vec{u}_1, \dots, \vec{u}_n) = 1$, indirecte si $\varepsilon(\vec{u}_1, \dots, \vec{u}_n) = -1$.

Ainsi, la définition du tenseur de Levi-Civita ne dépend plus de la base choisie de départ mais uniquement du choix de l'orientation.

Théorème VI.2.17. Dans une base quelconque, on a $\varepsilon = \varepsilon_{i_1 \dots i_n} \underline{\mathrm{d}x}^{i_1} \otimes \dots \otimes \underline{\mathrm{d}x}^{i_n}$ avec

$$\boxed{\varepsilon_{i_1 \dots i_n} = \pm \sqrt{|g|}[i_1, \dots, i_n]} \quad (\text{VI.112})$$

où $g = \det g$ est le déterminant de la matrice formé par les coordonnées du tenseur métrique dans la base considérée. On prendra le signe + si la base est directe, - si la base est indirecte.

Preuve. Soit (\vec{e}_i) une base orthonormée directe et (\vec{u}_i) une base quelconque de $T_x\mathcal{M}$, et P la matrice de passage entre les deux bases. On sait que dans la base orthonormée, $g_{ij} = \delta_{ij}$ et dans la base quelconque, les composantes du tenseur métrique peuvent s'écrire sous forme matricielle⁵ $g_{ij} = {}^t P P$ où ${}^t P$ est la matrice transposée de P . De telle sorte qu'on a $g = \det g = \det P^2$. D'abord constatons que

$$\varepsilon = \varepsilon_{12 \dots n} \underline{\mathrm{u}}^1 \wedge \dots \wedge \underline{\mathrm{u}}^n = \varepsilon_{12 \dots n} \sum_{(i_k)} [i_1, \dots, i_n] \underline{\mathrm{u}}^{i_1} \otimes \dots \otimes \underline{\mathrm{u}}^{i_n} \quad (\text{VI.113})$$

En particulier, on a

$$\varepsilon_{12 \dots n} = \underline{\mathrm{e}}^1 \wedge \dots \wedge \underline{\mathrm{e}}^n (\vec{u}_1, \dots, \vec{u}_n) = \sum_{\sigma \in \mathcal{G}_n} k(\sigma) P_1^{\sigma(1)} \dots P_n^{\sigma(n)} = \det P = \pm \sqrt{g} \quad (\text{VI.114})$$

d'où $\varepsilon_{12 \dots n} = \pm \sqrt{g}$ selon le caractère direct ou indirect de la base, ce qui achève la preuve.

Dans le cas où la signature de \mathbf{g} est $(-, +, +, +)$, la preuve est similaire, au détail près qu'il faut ajouter des signes moins où il faut dans la définition de \mathbf{g} . En effet, il suffit de constater que dans ce cas, on aurait matriciellement $g_{ij} = {}^t P \eta P$ avec η la matrice de Minkowski, et donc $\det g_{ij} = \det {}^t P \det \eta \det P = -\det P^2$.

Exercices - exemples

Nous mettons sous forme d'exercice des démonstrations de résultats importants pour la suite du cours. Les résultats importants sont encadrés.

Produit vectoriel de $n - 1$ vecteurs

Soit un espace vectoriel de dimension n , orienté, et muni d'une métrique (par exemple $T_x\mathcal{M}$). Soient $n - 1$ vecteurs $\vec{v}_1, \dots, \vec{v}_{n-1}$. On pose

$$\underline{\omega} = \varepsilon(\vec{v}_1, \dots, \vec{v}_{n-1}, \cdot) \quad (\text{VI.115})$$

Dans la base naturelle, déterminer les coordonnées de $\vec{\omega}$, dual métrique de $\underline{\omega}$. Considérer le cas particulier de \mathbb{R}^3 avec $\mathbf{g} = \mathbb{I}$ la matrice identité.

Solution. On a $\underline{\omega} = \omega_i \underline{\mathrm{d}x}^i$ et $\vec{\omega} = \omega^i \vec{\partial}_i$ avec $\omega^i = g^{ij} \omega_j$. D'autre part,

$$\begin{aligned} \omega_j &= \underline{\omega}(\vec{\partial}_j) \\ &= \varepsilon(\vec{v}_1, \dots, \vec{v}_{n-1}, \vec{\partial}_j) \\ &= \pm \sqrt{g}[i_1, \dots, i_{n-1}, i_n] v_1^{i_1} \dots v_{n-1}^{i_{n-1}} \delta_j^{i_n} \\ &= \pm \sqrt{g}[i_1, \dots, i_{n-1}, j] v_1^{i_1} \dots v_{n-1}^{i_{n-1}} \end{aligned} \quad (\text{VI.116})$$

5. Nous identifions de façon impropre la matrice et les composantes du tenseur métrique dans la base donnée par la notation " g_{ij} " afin d'éviter toute confusion avec la notation $g = \det g_{ij}$.

et donc

$$\omega^i = g^{ij} \pm \sqrt{g}[i_1, \dots, i_{n-1}, j] v_1^{i_1} \dots v_{n-1}^{i_{n-1}} \quad (\text{VI.117})$$

En dimension 3, cela donne

$$\omega^j = [i_1, i_2, j] v_1^{i_1} v_2^{i_2} = \begin{pmatrix} v_1^y v_2^z - v_1^z v_2^y \\ v_1^z v_2^y - v_1^y v_2^z \\ v_1^x v_2^y - v_1^y v_2^x \end{pmatrix} \quad (\text{VI.118})$$

C'est le produit vectoriel. Dans les espaces courbes en dimension 3, ou dans un système de coordonnées quelconques dans \mathbb{R}^3 , on a donc :

$$\varepsilon(\vec{v}_1, \vec{v}_2, \cdot) = \mathbf{g}(\vec{v}_1 \times \vec{v}_2, \cdot) \quad (\text{VI.119})$$

ce qui justifie le cas particulier de la "formule du produit mixte" souvent utilisée en physique.

Gymnastique avec le tenseur de Levi-Civita

Nous nous proposons de démontrer quelques propriétés intéressantes qui seront utiles pour établir des propriétés intéressantes de la dualité de Hodge et surtout pour construire le champ électromagnétique.

1. Montrer que $\varepsilon^{i_1 \dots i_n} = \pm \frac{1}{\sqrt{g}}[i_1, \dots, i_n]$, où $\varepsilon^{i_1 \dots i_n}$ sont les composantes du dual métrique du tenseur de Levi-Civita.
2. Montrer que $\varepsilon^{i_1 \dots i_n} \varepsilon_{j_1 \dots j_n} = \pm \delta_{j_1 \dots j_n}^{i_1 \dots i_n}$ où $\delta_{j_1 \dots j_n}^{i_1 \dots i_n}$ est le delta de Kronecker généralisé : il vaut $k(\sigma)$ si $(i_1, \dots, i_n) = \sigma(j_1, \dots, j_n)$ où $\sigma \in \mathcal{G}_n$, 0 sinon. Distinguer les cas où l'on doit choisir + ou -.
3. Montrer que pour tout $p \in \llbracket 1, n \rrbracket$, $\varepsilon^{i_1 \dots i_p a_1 \dots a_{n-p}} \varepsilon_{i_1 \dots i_p b_1 \dots b_{n-p}} = \pm p! \delta_{b_1 \dots b_{n-p}}^{a_1 \dots a_{n-p}}$. Indication : faire une récurrence finie en initialisant avec le résultat de la question précédente.

Solution.

1. On a

$$\varepsilon^{i \dots n} = \varepsilon_{i_1 \dots i_n} g^{1i_1} \dots g^{ni_n} \quad (\text{VI.120})$$

$$= \sqrt{g}[i_1, \dots, i_n] g^{1i_1} \dots g^{ni_n} \quad (\text{VI.121})$$

$$= \sqrt{g} \sum_{\sigma \in \mathcal{G}_n} k(\sigma) g^{1\sigma(1)} \dots g^{n\sigma(n)} \quad (\text{VI.122})$$

$$= \sqrt{g} \det(g^{ij}) \quad (\text{VI.123})$$

$$= \begin{cases} 1/\sqrt{g} & \text{si } \mathbf{g} \text{ est définie positive,} \\ -1/\sqrt{g} & \text{en relativité générale,} \end{cases} \quad (\text{VI.124})$$

car $\det(g^{ij}) = 1/\det(g_{ij})$. Par ailleurs on sait que $\tilde{\varepsilon}$, le dual métrique de ε , est un élément de $\mathcal{A}_2 T^* \mathcal{M}$, donc a les mêmes propriétés que ε sauf que ses arguments sont des formes linéaires. On sait donc qu'il vit dans un espace vectoriel de dimension 1 et que ses composantes, $\varepsilon^{i_1 \dots i_n}$ sont nécessairement proportionnelles à $[i_1, \dots, i_n]$. La constante de proportionnalité est $\varepsilon^{12 \dots n}$, d'où

$$\varepsilon^{i_1 \dots i_n} = \pm \frac{1}{\sqrt{g}}[i_1, \dots, i_n] \quad (\text{VI.125})$$

où il faudra choisir le signe + si la base est directe et que \mathbf{g} est définie positive, ou que la base est indirecte et qu'on est en relativité générale, et le signe - si la base est indirecte et que \mathbf{g} est définie positive, ou que la base est directe et qu'on est en relativité générale.

2. On sait que $\tilde{\varepsilon}$ et ε vivent chacun dans des espaces vectoriels de dimension 1. Leur produit tensoriel, qui a pour composantes $\varepsilon^{i_1 \dots i_n} \varepsilon_{j_1 \dots j_n}$, vit donc également dans un espace de dimension 1 : l'ensemble des tenseurs de type (n, n) , antisymétriques sur les n premiers arguments et antisymétrique sur les n derniers. Le tenseur de type (n, n) et de composantes $\delta_{j_1 \dots j_n}^{i_1 \dots i_n}$ satisfait à cette contrainte. Il est donc proportionnel à $\tilde{\varepsilon} \otimes \varepsilon$. La constante

de proportionnalité se calcule comme suit. Si l'on prend une base orthonormée directe $(\vec{e}_1, \dots, \vec{e}_n)$ et sa base duale métrique $(\underline{e}^1, \dots, \underline{e}^n)$, on voit immédiatement que

$$\tilde{\epsilon} \otimes \epsilon(\underline{e}^1, \dots, \underline{e}^n, \vec{e}_1, \dots, \vec{e}_n) = \pm 1 \quad (\text{VI.126})$$

avec $+$ si \mathbf{g} est définie positive et $-$ en relativité. En effet, pour une base orthonormée, on a $e_0 = \eta_{0i} e^i = -e^0 = -1$, donc $\tilde{\epsilon}(\underline{e}^1, \dots, \underline{e}^n) = -1$. D'autre part, considérant le tenseur $\delta = \delta_{j_1 \dots j_n}^{i_1 \dots i_n} \overrightarrow{\partial}_{i_1} \otimes \dots \otimes \overrightarrow{\partial}_{i_n} \otimes \underline{dx}^{j_1} \otimes \dots \otimes \underline{dx}^{j_n}$, on a

$$\delta(\underline{e}^1, \dots, \underline{e}^n, \vec{e}_1, \dots, \vec{e}_n) = \delta_{j_1 \dots j_n}^{i_1 \dots i_n} \delta_{i_1}^1 \dots \delta_{i_n}^n \delta_{j_1}^{j_1} \dots \delta_{j_n}^{j_n} = \delta_{12 \dots n}^{12 \dots n} = 1 \quad (\text{VI.127})$$

d'où le résultat.

3. On montre le résultat par récurrence. Au rang $p = 0$, la propriété est vraie d'après la question précédente. Supposons que la propriété annoncée soit vraie au rang p avec $p \in \llbracket 1, n-1 \rrbracket$. Montrons qu'elle est vraie au rang $p+1$. Nous avons, d'après l'hypothèse de récurrence :

$$\varepsilon^{i_1 \dots i_p k a_1 \dots a_{n-p-1}} \varepsilon_{i_1 \dots i_p k b_1 \dots b_{n-p-1}} = \pm p! \delta_{k b_1 \dots b_{n-p-1}}^{k a_1 \dots a_{n-p-1}} \quad (\text{VI.128})$$

L'indice k va se répéter n fois, mais $\delta_{k b_1 \dots b_{n-p-1}}^{k a_1 \dots a_{n-p-1}}$ n'est non nul que si $k \notin \{a_1, \dots, a_{n-p-1}\}$ ce qui ne peut arriver que $p+1$ fois. Il faudra compter $(p+1)$ fois si (a_1, \dots, a_{n-p-1}) est une permutation de (b_1, \dots, b_{n-p-1}) , soit $\delta_{k b_1 \dots b_{n-p-1}}^{k a_1 \dots a_{n-p-1}} = (p+1) \delta_{b_1 \dots b_{n-p-1}}^{a_1 \dots a_{n-p-1}}$. Notons que l'égalité est aussi vérifiée si (a_1, \dots, a_{n-p-1}) n'est pas une permutation de (b_1, \dots, b_{n-p-1}) , puisque les deux membres de l'égalité valent 0. Comme on a $p!(p+1) = (p+1)!$, cela permet de conclure la récurrence. L'égalité annoncée est donc vraie pour tout p entier compris entre 1 et n . Le choix des signes $+$ ou $-$ est similaire à celui de la question précédente.

Dualité de Hodge

Nous avons déjà vu que $\dim \mathcal{A}_p \mathcal{M} = \dim \mathcal{A}_{n-p} \mathcal{M}$. Ces deux ensembles sont donc en bijection d'après des théorèmes classiques d'algèbre linéaire. Une utilisation particulière du tenseur de Levi-Civita va nous permettre de passer de l'un à l'autre.

Définition VI.2.18. Soit $p \in \llbracket 1, n \rrbracket$. On appelle étoile de Hodge l'application

$$\star : \mathcal{A}_p \mathcal{M} \rightarrow \mathcal{A}_{n-p} \mathcal{M}, \omega \mapsto \star \omega \quad (\text{VI.129})$$

que l'on définit explicitement par les transformation des composantes :

$$\star \omega_{i_1 \dots i_{n-p}} = \frac{1}{p!} \varepsilon_{a_1 \dots a_p i_1 \dots i_{n-p}} g^{a_1 b_1} \dots g^{a_p b_p} \omega_{b_1 \dots b_p} \quad (\text{VI.130})$$

Explicitement, si $n = 3$:

$$p = 0 : \star \omega_{ijk} = \omega \varepsilon_{ijk} \quad (\text{VI.131})$$

$$p = 1 : \star \omega_{ij} = \varepsilon_{aij} g^{ab} \omega_b \quad (\text{VI.132})$$

$$p = 2 : \star \omega_i = \frac{1}{2} \varepsilon_{abi} g^{ap} g^{bq} \omega_{pq} \quad (\text{VI.133})$$

$$p = 3 : \star \omega = \frac{1}{6} \varepsilon_{abc} g^{ap} g^{bq} g^{cr} \omega_{pqr} \quad (\text{VI.134})$$

Proposition VI.2.19. $\forall \omega \in \mathcal{A}_p \mathcal{M}, \star \star \omega = \pm (-1)^{np} \omega$, + si \mathbf{g} est définie positive, $-$ en relativité.

Preuve. D'abord, on sait que $\star \star \omega \in \mathcal{A}_p \mathcal{M}$. De là, on calcule

$$\begin{aligned} \star(\star \omega)_{i_1 \dots i_p} &= \frac{1}{(n-p)!} \varepsilon_{a_1 \dots a_{n-p} i_1 \dots i_p} g^{a_1 b_1} \dots g^{a_{n-p} b_{n-p}} \frac{1}{p!} \varepsilon_{u_1 \dots u_p b_1 \dots b_{n-p}} g^{u_1 v_1} \dots g^{u_p v_p} \omega_{v_1 \dots v_p} \\ &\quad (\text{VI.135}) \end{aligned}$$

$$= \frac{1}{p!(n-p)!} \varepsilon_{a_1 \dots a_{n-p} i_1 \dots i_p} \varepsilon^{v_1 \dots v_p a_1 \dots a_{n-p}} \omega_{v_1 \dots v_p} \quad (\text{VI.136})$$

D'une part on substitue $\varepsilon^{v_1 \dots v_p a_1 \dots a_{n-p}} = (-1)^{np} \varepsilon^{a_1 \dots a_{n-p} v_1 \dots v_p}$, puis d'après l'exercice VI.2.3, on a $\varepsilon_{a_1 \dots a_{n-p} i_1 \dots i_p} \varepsilon^{a_1 \dots a_{n-p} v_1 \dots v_p} = \pm(n-p)! \delta_{i_1 \dots i_p}^{v_1 \dots v_p}$. De telle sorte que

$$\star \star \omega_{i_1 \dots i_p} = (-1)^{np} \frac{1}{p!} \delta_{i_1 \dots i_p}^{v_1 \dots v_p} \omega_{v_1 \dots v_p} \quad (\text{VI.137})$$

Comme ω est un tenseur antisymétrique, pour toute permutation σ on a $\delta_{j_1 \dots j_n}^{\sigma(j_1) \dots \sigma(j_n)} \omega_{\sigma(j_1) \dots \sigma(j_n)} = k(\sigma)^2 \omega_{j_1 \dots j_n} = \omega_{j_1 \dots j_n}$. Dans la sommation des indices répétés de l'équation VI.137, les seuls termes non nuls qui nous intéressent sont ceux pour lesquels $(i_1 \dots i_n)$ est une permutation de $(j_1 \dots j_n)$. Il y en a exactement $p!$. D'où le résultat.

Remarquons que cette propriété établit bien que \star est une bijection, puisque l'on a $\star^{-1} = \pm(-1)\star$.

Quelques applications dans l'espace-temps relativiste

Désormais et jusqu'à la fin du chapitre, nous nous plaçons dans le cadre de la relativité (restreinte ou générale) : $\dim \mathcal{M} = 4$ et \mathbf{g} a une signature $-,+,+,+.$ Nous aurons besoin de deux propriétés très intéressantes avant d'appliquer tout ce formalisme à l'électromagnétisme.

Proposition VI.2.20. $\forall \underline{a}, \underline{b} \in \mathcal{A}_1 \mathcal{M}, \star(\underline{a} \wedge \underline{b}) = \varepsilon(\overrightarrow{a}, \overrightarrow{b}, \cdot, \cdot)$ où \overrightarrow{a} et \overrightarrow{b} sont les vecteurs duals métriques de \underline{a} et \underline{b} .

Preuve. Nous allons simplement montrer l'égalité des composantes.

$$(\star(\underline{a} \wedge \underline{b}))_{ij} = \frac{1}{2} \varepsilon_{rsij} g^{rp} g^{sq} (a_p b_q - a_q b_p) \quad (\text{VI.138})$$

$$= \frac{1}{2} \varepsilon_{rsij} (a^r b^s - a^s b^r) \quad (\text{VI.139})$$

$$= \varepsilon_{rsij} a^r b^s \quad (\text{VI.140})$$

Ce théorème va jouer un rôle central dans la décomposition du tenseur électromagnétique en sa composante électrique puis sa composante magnétique.

Théorème VI.2.21 (Décomposition orthogonale des 2-formes). Soit $\omega \in \mathcal{A}_2 \mathcal{M}$. Soit $\overrightarrow{u} \in \mathcal{T} \mathcal{M}$ tel que $g(\overrightarrow{u}, \overrightarrow{u}) = -1$. Alors ω peut se décomposer de façon unique selon deux formes linéaires \underline{q} et \underline{b} toutes deux orthogonales à \overrightarrow{u} , c'est-à-dire qu'il existe deux uniques formes linéaires \underline{q} et \underline{b} vérifiant $\underline{q}(\overrightarrow{u}) = 0$ et $\underline{b}(\overrightarrow{u}) = 0$ telles que $\omega = \underline{u} \wedge \underline{q} + \star(\underline{u} \wedge \underline{b})$ où \underline{u} est le dual métrique de \overrightarrow{u} . Autrement dit

$$\forall \omega \in \mathcal{A}_2 \mathcal{M}, \forall \overrightarrow{u} \in \mathcal{T} \mathcal{M} / \|\overrightarrow{u}\|^2 = -1, \exists! (\underline{q}, \underline{b}) \in (\mathcal{A}_1 \mathcal{M})^2 / \underline{q}(\overrightarrow{u}) = 0, \underline{b}(\overrightarrow{u}) = 0, \omega = \underline{u} \wedge \underline{q} + \star(\underline{u} \wedge \underline{b}) \quad (\text{VI.141})$$

Preuve. On va démontrer ce théorème sous forme d'exercice.

1. Soit $\omega \in \mathcal{A}_2 \mathcal{M}$. On pose $\underline{q} = \omega(\overrightarrow{u}, \cdot)$ et $\mathbf{B} = \omega - \underline{u} \wedge \underline{q}$, où $\underline{u} = \mathbf{g}(\overrightarrow{u}, \cdot)$. Montrer que $\mathbf{B}(\cdot, \overrightarrow{u}) = \underline{0}$.
2. Soient $(\overrightarrow{e}_1, \overrightarrow{e}_2, \overrightarrow{e}_3)$ tels que $(\overrightarrow{u}, \overrightarrow{e}_1, \overrightarrow{e}_2, \overrightarrow{e}_3)$ forme une base orthonormée directe. On pose $\overrightarrow{b} = b^i \overrightarrow{e}_i$ où la sommation ne s'effectue que pour $i \in \{1, 2, 3\}$, avec $b^i = \frac{1}{2}[i, j, k] \mathbf{B}(\overrightarrow{e}_i, \overrightarrow{e}_j)$. Montrer que $\forall \overrightarrow{v}, \overrightarrow{w} \in \text{Vect}(\overrightarrow{e}_1, \overrightarrow{e}_2, \overrightarrow{e}_3) = (\mathbb{R} \overrightarrow{u})^\perp$, on a

$$\mathbf{B}(\overrightarrow{v}, \overrightarrow{w}) = \varepsilon_{\overrightarrow{u}}(\overrightarrow{b}, \overrightarrow{v}, \overrightarrow{w}) = \varepsilon(\overrightarrow{u}, \overrightarrow{b}, \overrightarrow{v}, \overrightarrow{w}) \quad (\text{VI.142})$$

3. En déduire le théorème.

Solution.

1. $\mathbf{B}(\cdot, \overrightarrow{u}) = \omega(\cdot, \overrightarrow{u}) - \omega(\overrightarrow{u}, \overrightarrow{u}) \underline{u} + \underline{u}(\overrightarrow{u}) \omega(\cdot, \overrightarrow{u})$. Le deuxième terme est nul car ω est alternée. En se rappelant que $\underline{u}(\overrightarrow{u}) = \|\overrightarrow{u}\|^2 = -1$, on trouve bien 0.
2. Soient $\overrightarrow{v} = v^i \overrightarrow{e}_i$, $\overrightarrow{w} = w^i \overrightarrow{e}_i$, deux vecteurs de $(\mathbb{R} \overrightarrow{u})^\perp$. On a $\mathbf{B}(\overrightarrow{v}, \overrightarrow{w}) = v^i w^j \mathbf{B}(\overrightarrow{e}_i, \overrightarrow{e}_j)$. Comme \mathbf{B} est antisymétrique, il ne faut sommer que quand $i \neq j$. En utilisant le fait que

$$b^1 = \mathbf{B}(\overrightarrow{e}_2, \overrightarrow{e}_3), \quad b^2 = \mathbf{B}(\overrightarrow{e}_3, \overrightarrow{e}_1), \quad b^3 = \mathbf{B}(\overrightarrow{e}_1, \overrightarrow{e}_2) \quad (\text{VI.143})$$

le développement de cette somme donne

$$\mathbf{B}(\vec{v}, \vec{w}) = v^1 v^2 b^3 - v^2 w^1 b^3 + v^3 w^1 b^2 - v^1 w^3 b^2 + v^2 w^3 b^1 - v^3 w^2 b^1 = \begin{vmatrix} b^1 & v^1 & w^1 \\ b^2 & v^2 & w^2 \\ b^3 & v^3 & w^3 \end{vmatrix} = \varepsilon_{\vec{u}}(\vec{b}, \vec{v}, \vec{w}) \quad (\text{VI.144})$$

où $\varepsilon_{\vec{u}} = \vec{u} \dashv \varepsilon = \varepsilon(\vec{u}, \cdot, \cdot, \cdot)$ est le tenseur de Levi-Civita restreint à $(\mathbb{R}\vec{u})^\perp$.

3. Le résultat précédent peut se généraliser à deux vecteurs de $T_x\mathcal{M}$. En effet, si on ajoute une composante le long de \vec{u} à \vec{v} et \vec{w} , le résultat sera indifférent :

$$\forall \vec{v}, \vec{w} \in T_x\mathcal{M}, \mathbf{B}(\vec{v}, \vec{w}) = \varepsilon(\vec{u}, \vec{b}, \vec{v}, \vec{w}) \quad (\text{VI.145})$$

Sachant que $\mathbf{B} = \omega - \underline{u} \wedge \underline{q}$, on en déduit l'existence d'une forme linéaire \vec{q} et d'un vecteur \vec{b} tels que $\underline{q}(\vec{u}) = \vec{b} \cdot \underline{u} = 0$ et $\omega = \underline{u} \wedge \underline{q} + \varepsilon(\vec{u}, \vec{b}, \cdot, \cdot) = \underline{u} \wedge \underline{q} + \star(\underline{u} \wedge \vec{b})$ d'après la propriété VI.2.20. Nous venons donc de démontrer l'existence de cette décomposition puisque nous avons exibé des formes linéaires qui conviennent. Montrons l'unicité. Pour \underline{q} c'est facile : s'il existe une autre forme \underline{q}' et un autre vecteur \vec{b}' qui conviennent également, alors on a, pour tout vecteur \vec{v}

$$0 = \underline{u} \wedge \underline{q}(\vec{v}, \vec{u}) + \varepsilon(\vec{u}, \vec{b} \vec{v}, \vec{u}) - \underline{u} \wedge \underline{q}'(\vec{v}, \vec{u}) + \varepsilon(\vec{u}, \vec{b}' \vec{v}, \vec{u}) \quad (\text{VI.146})$$

Toutes les formes de Levi-Civita vont donner 0. Si on décompose les produits extérieurs, on va trouver

$$0 = \underline{u}(\vec{v})\underline{q}(\vec{u}) - \underline{u}(\vec{u})\underline{q}(\vec{v}) - \underline{u}(\vec{v})\underline{q}(\vec{u}) + \underline{u}(\vec{u})\underline{q}(\vec{v}) = \underline{q}(\vec{v}) - \underline{q}'(\vec{v}) \quad (\text{VI.147})$$

soit $\forall \vec{v} \in \mathcal{T}\mathcal{M}, \underline{q}(\vec{v}) = \underline{q}'(\vec{v})$, d'où l'unicité de \underline{q} . Pour l'unicité de \vec{b} , considérons deux vecteurs \vec{v} et \vec{w} dans $(\mathbb{R}\vec{u})^\perp$. S'il existe un autre \vec{b}' qui convient, alors on a également $\varepsilon_{\vec{u}}(\vec{b} - \vec{b}', \vec{v}, \vec{w}) = 0 = (\vec{b} - \vec{b}') \cdot (\vec{v} \times \vec{w})$ ce qui n'est possible que si $\vec{b} = \vec{b}'$.

VI.2.4 Le champ électromagnétique

Nous allons montrer que la forme du champ électromagnétique exposée section VI.2.1 est la seule possible. Pour toute cette section nous nous plaçons dans l'espace-temps de la relativité générale modélisé par une variété riemannienne (\mathcal{M}, g) où $\dim \mathcal{M} = 4$ et g est de signature $(-, +, +, +)$.

VI.2.5 Prolégomène : particule dans un champ de force pure

Considérons une particule matérielle. On sait qu'en relativité générale cela signifie que cette particule suit une ligne d'univers que l'on peut paramétrier ainsi : $\mathcal{L} : I \rightarrow \mathcal{M}, \lambda \mapsto \mathcal{P}(\lambda)$, avec $\forall \lambda \in I, g(\dot{\gamma}'(\lambda), \dot{\gamma}'(\lambda)) < 0$. On reparamétrise cette trajectoire selon son abscisse curviligne qui vérifie $ds = \sqrt{-g(\dot{\gamma}'(\lambda), \dot{\gamma}'(\lambda))} d\lambda$. En notant $\vec{u} = \vec{d}/ds$, on trouve $g(\vec{u}, \vec{u}) = \|\vec{u}\|^2 = -1$. C'est ce qu'on appelle la 4-vitesse, ou la vitesse normalisée, du point matériel. Si la particule a une masse au repos m , on définit $\vec{p} = m\vec{u}$ sa quadri-impulsion. Le principe fondamental de la dynamique en relativité suppose qu'il existe un champ de force \vec{f} tel que l'on a

$$\frac{d\vec{p}}{d\tau} = \vec{f} \quad (\text{VI.148})$$

où $d\tau = cds$ est ce qu'on appelle le temps propre. Si l'on développe cette équation, on obtient

$$\vec{f} = c \frac{dm}{ds} \vec{u} + mc \frac{d\vec{u}}{d\tau} = c \frac{dm}{ds} \vec{u} + mc \vec{a} \quad (\text{VI.149})$$

Or on sait que pour les trajectoires unitaires, $\vec{u} \cdot \vec{a} = 0$. La variation de la masse au repos de la particule se calcule donc de cette façon : $g(\vec{f}, \vec{u}) = cdm/ds = dm/d\tau$.

Définition VI.2.22. Les 4-forces pures sont les forces qui conservent la masse des particules, ce sont des champs de vecteurs \vec{f} (ou, par dualité métrique, des champs de formes linéaires \underline{f}) qui vérifient $\vec{f} \cdot \vec{u} = 0$ (ou $\underline{f}(\vec{u}) = 0$).

À moins que nous ne traitions de physique quantique relativiste avec collisions de particules élémentaires (ce qui n'est pas le cas ici), nous ne traiterons que des forces pures.

Construction du champ électromagnétique

Il est raisonnable de penser que le champ électromagnétique induit une 4-force pure sur les particules. D'autre part, si l'on peut imaginer des interactions scalaires qui ne dépendent que de la position des particules dans l'espace, on peut aussi imaginer des interactions vectorielles, c'est-à-dire qui dépendent de la direction intrinsèque du mouvement des particules, c'est-à-dire de leur 4-vitesse \vec{u} . C'est ce que nous supposons pour le champ électromagnétique⁶. La dernière hypothèse forte que nous ajoutons est que l'action sur les particules est proportionnelle à la 4-vitesse \vec{u} . Si chaque particule a un coefficient de couplage électromagnétique intrinsèque q que l'on peut appeler "charge électromagnétique", alors il doit exister une forme bilinéaire \mathbf{F} telle que la 4-force électromagnétique s'écrit ainsi

$$\underline{\mathbf{f}} = q\mathbf{F}(\cdot, \vec{u}) \quad (\text{VI.150})$$

C'est ce qu'on appelle la "force de Lorentz" en physique. Si l'on suppose que cette force est pure, alors pour tout vecteur unitaire, on a $\mathbf{F}(\vec{u}, \vec{u}) = 0$. Cela indique que \mathbf{F} est une forme bilinéaire alternée, c'est donc un élément de $\mathcal{A}_2\mathcal{M}$. Mais, nous savons que toute 2-forme peut se décomposer de façon unique orthogonalement à un vecteur unitaire \vec{u} donné. Ainsi, pour tout observateur de quadritvitesse \vec{u} , il existe deux formes linéaires, uniques, $\underline{\mathbf{E}}$ et $\underline{\mathbf{B}}$, telles que

$$\mathbf{F} = \underline{\mathbf{u}} \wedge \underline{\mathbf{E}} + \star(\underline{\mathbf{u}} \wedge c\underline{\mathbf{B}}) \quad (\text{VI.151})$$

$\underline{\mathbf{E}}$ et $\underline{\mathbf{B}}$ sont les formes électriques et magnétiques. Par dualité métrique on peut évidemment leur faire correspondre les vecteurs électrique et magnétique respectivement. On constate que $\underline{\mathbf{E}}$ et $\underline{\mathbf{B}}$ ne sont pas des champs de formes linéaires définis intrinsèquement, car ils dépendent de l'observateur et sa quadritvitesse. Si l'on choisit un autre observateur qui croise le chemin du premier mais avec une autre quadritvitesse, alors il faudra changer la décomposition de \mathbf{F} et l'expression des champs électrique et magnétique.

VI.2.6 Exercices

1. Montrer que $\underline{\mathbf{E}} = \mathbf{F}(\cdot, \vec{u})$ et $c\underline{\mathbf{B}} = \star\mathbf{F}(\vec{u}, \cdot)$.
2. Exprimer les coordonnées de $\vec{\mathbf{E}}$ et $\vec{\mathbf{B}}$ en fonction de celles de \mathbf{F} .
3. Exprimer les composantes de \mathbf{F} en fonction de celles de $\underline{\mathbf{E}}$, $\vec{\mathbf{B}}$, dans une base quelconque, puis dans le repère orthonormé dont le premier vecteur est \vec{u} .

Solution.

1. C'est un calcul direct.
2. C'est aussi du calcul direct : $E^i = (\mathbf{F}(\cdot, \vec{u}))^i = F^{ib}u_b$. $cB^i = (\star\mathbf{F}(\vec{u}, \cdot))^i = -\frac{1}{2}\varepsilon^{ijk\ell}F_{jk}u_\ell$
3. $F_{ab} = u_a E_\beta - E_a u_b + c\varepsilon_{cdab}u^c E^d$. Dans une base orthonormée décomposé le long de \vec{u} , on a $\vec{u} = \vec{e}_0$. En utilisant la matrice de Minkowski on vérifie immédiatement que $u^a = (1, 0, 0, 0) = \delta_0^a$ et $u_a = (-1, 0, 0, 0) = -\delta_a^0$ (les indices vont de 0 à 3), ainsi que $E_a = (0, E_x, E_y, E_z)$ et $B^a = (0, B^x, B^y, B^z)$, avec $\varepsilon_{abcd} = [a, b, c, d]$. D'où

$$F_{ab} = -\delta_0^a E_b + E_a \delta_0^b + c[0, k, a, b]B^k = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & cB^x & -cB^y \\ E_y & -cB^z & 0 & cB^x \\ E_z & cB^y & -cB^x & 0 \end{pmatrix} \quad (\text{VI.152})$$

Quelques compléments :

$$F^{ab} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & cB^x & -cB^y \\ -E_y & -cB^z & 0 & cB^x \\ -E_z & cB^y & -cB^x & 0 \end{pmatrix} \quad (\text{VI.153})$$

6. Le cas de la gravitation est plus compliqué de ce point de vue-là. On peut consulter l'ouvrage de Feynman "leçons sur la gravitation" où il traite la gravitation comme un champ quantique et retrouve les lois de la relativité générale à partir de principes variationnels.

$$\star \mathbf{F} = c\mathbf{B} \wedge \mathbf{u} + \star(\mathbf{u} \wedge \mathbf{E}) \quad (\text{VI.154})$$

$$\star F_{ab} = \begin{pmatrix} 0 & cB_x & cB_y & cB_z \\ -cB_x & 0 & E^x & -E^y \\ -cB_y & -E^z & 0 & E^x \\ -cB_z & E^y & -E^x & 0 \end{pmatrix} \quad (\text{VI.155})$$

D'amples compléments sur le champ électromagnétiques peuvent être trouvés dans le livre d'Éric Gourgoulhon, "relativité restreinte". Nous continuerons de travailler sur le champ électromagnétiques dans le chapitre suivant sur la dérivation dans les variétés.

VI.3 Dérivation

Nous allons maintenant voir comment ces quantités (scalaire, vecteur, tenseur et forme différentiel) se dérive sur une variété. Nous avons vu au chapitre VI comment dériver des fonctions de la position sur une surface régulière plongé dans \mathbb{R}^3 . Pour cela nous avons vu deux approches, la première consisté à dériver la quantité en question le long d'une courbe paramétrée incluse dans la surface. La seconde plus brutal, où nous ramenions la dépendance de la coordonnée non pas à la position sur la surface, mais au "coordonnée" dans un paramétrage décrivant la position des points au voisinage du point où l'on effectuera cette dérivée. Nous avions ensuite vue l'indépendance de cette définition d'avec le choix du paramétrage utilisé.

On a ensuite vu en sous-section VI.1.2 qu'un champs de vecteur de l'espace-tangent pouvait être assimilé à une forme de dérivation sur les champs scalaires $\mathcal{F}(P)$ d'une variété. Cela correspondait à la première approche effectué sur les surfaces. Nous allons voir comment, à partir des définitions, adoper à nouveau la seconde approche.

VI.3.1 Connexion covariante de Koszul, de Riemann, de Levi-Civita

La connexion covariante est la notion qui se substitue a celles de différentiel (pour les fonctions définie à l'arrivée et au départ dans les espaces vectoriels) dans le cas des tenseurs sur les variétés. Nous allons définir abstrairement, cette notion de dérivation, appliquée au variété. Puis nous verrons que cette définitions abstraites (Connexion covariante de Koszul) conduit à un ensemble de connexion localement défini auxquels sont associée des notions de courbure et de torsion. La connexion de Riemann sera un sous-ensemble des connexions de Koszul qui préservent le produit scalaire, ce qui est équivalent à dire que la métrique est parrallèle vis-à-vis de cette connexion. Enfin nous introduirons la connexion de Levi-Civita, celles qu'utilise tout les jours les physiciens dans le cadre de la relativité général. Et qui correspond à une connexions de Riemann dont la torsion sont nulles, nous démontrerons qu'une tel connexion existe est qu'elle est unique.

Différentiel des scalaires sur une variété

Pour calculer la différentiel d'un champs scalaire quelconque $f : \mathcal{M} \rightarrow \mathbb{R}$ sur une variété, il suffit simplement d'évaluer celles-ci dans un atlas. En effet reprenons les notations de la section VI.1, alors au voisinage d'un point $P = \phi_i^{-1}(p) \in \mathcal{V}_i \subset \mathcal{M}$, on peut se ramener à $f \circ \phi_i^{-1} : \mathcal{U}_i \rightarrow \mathbb{R}$ dont nous pouvons calculer la différentielle. Cette différentielle s'applique naturellement sur des accroisements vivant dans \mathcal{U}_i or nous avons envie d'utiliser la différentiel directement sur les vecteurs tangents de la variété, il ne reste donc qu'à composer la différentielle avec l'isomorphisme $h^k \partial_k \in \mathcal{T}\mathcal{M} \mapsto (h^1, \dots, h^n) \in \mathbb{R}^n$ transportant naturellement les vecteurs de $\mathcal{T}\mathcal{M}$ vers des accroissement de \mathcal{U}_i . On posera donc,

Définition VI.3.1 (Différentiel, dérivée partielle et gradient d'un champs scalaire). *Pour tout champ scalaire $f : \mathcal{M} \mapsto \mathbb{R}$ lisse sur la variété \mathcal{M} d'atlas $(\phi_i, \mathcal{U}_i, \mathcal{V}_i)_{i \in I}$, (on notera $\psi_i = \phi_i^{-1}$ les applications inverses). Pour tout point $P \in \mathcal{M}$, il existe $i \in I$ et $p \in \mathcal{U}_i$ tel que $P = \psi_i(p) \in \mathcal{V}_i \subset \mathcal{M}$, alors la différentiel de f en P , le long d'un vecteur $v \in \mathcal{T}\mathcal{M}$ est,*

$$df(P)v \hat{=} d(f \circ \psi_i)(p)(d\phi(p))v$$

où on a utilisé la notation abusive,

$$\begin{aligned} (d\phi(p))^{-1} &: \mathcal{T}\mathcal{M} &\rightarrow &\mathbb{R}^n \\ h^k \partial_k &\mapsto &(h^1, \dots, h^n) \end{aligned}$$

On définira alors les dérivées partielles par,

$$\partial_i f(P) \hat{=} df(P) \partial_i$$

ce qui permet d'écrire en utilisant les formes coordonnées \mathbf{dx}^k ,

$$df(P) = \partial_i f(P) \mathbf{dx}^i.$$

Le gradient $\nabla f(P)$ en P est alors défini comme le dual métrique de l'application différentiel,

$$df(P) \cdot \hat{=} g(\nabla f(P), \cdot)$$

et donc pour tout vecteur $\mathbf{v} \in \mathcal{T}\mathcal{M}$, on a,

$$df(P)\mathbf{v} \hat{=} g(\nabla f(P), \mathbf{v})$$

Pour être tout à fait rigoureux il faudrait montrer que la définition qu'on a pris ne dépend pas de l'atlas choisi. Une démonstration analogue a été effectué dans le chapitre sur les surfaces nous laisserons les étudiants les plus motivé vérifié que cette définition est cohérente.

Avec une tel définition, on montre facilement que pour toute courbe paramétrée $\gamma : I \longrightarrow \mathcal{M}$ incluse dans la variété on a,

$$\frac{df \circ \gamma}{dt} = df(P) \frac{d\gamma}{dt} = \partial_i f(P) \frac{d\gamma^i}{dt} = g\left(\nabla f(P), \frac{d\gamma}{dt}\right).$$

Coordonné orthogonal On appelle système de coordonnée orthogonal sur une variété, un atlas pour lequel les composantes $g_{i,j}$ du tenseur métrique sont diagonal. On notera alors $[g_{ij}] = \text{diag}(\pm h_1^2, \dots, \pm h_n^2)$, les quantités $(h_i)_{i \in \llbracket 1, n \rrbracket}$ sont alors appelées éléments de lignes. Alors la métrique inverse se calcule aisément et $[g^{ij}] = \text{diag}(\pm 1/h_1^2, \dots, \pm 1/h_n^2)$. Il est utiles dans ce cas la d'introduire une base orthonormal construite à partir de la base naturel $\partial_i = h_i e_i$. On notera alors $\mathbf{v} = v^k \partial_k = v^{\hat{k}} e_k$, ce qui implique que les coordonnée dans la base orthonormal vérifie $v^k h_k = v^{\hat{k}}$ (où il n'y a pas de sommation). Dans ce cas il est aisément de calculer les composantes du gradient, en effet.

$$\begin{aligned} \nabla f(P)^k &= g^{kl} df(P)_l \\ &= g^{kl} df(P) \partial_l \\ &= g^{kl} \partial_l f(P) \quad \Rightarrow \quad \boxed{\nabla f(P) = \frac{\partial_k f(P)}{h_k^2} \partial_k = \frac{\partial_k f(P)}{h_k} e_k} \\ &= \frac{\partial_k f(P)}{h_k^2} \end{aligned}$$

Exercice 39. —

On travaillera avec $\mathbb{R}^3 - O_y$ munit de la carte des coordonnées

$$\begin{aligned} \psi &: \mathbb{R} \times \mathbb{R}_{+,*} \times \mathbb{R} \rightarrow \mathbb{R}^3 - O_y \\ (u, v, z) &\mapsto \left(uv, \frac{u^2 - v^2}{2}, z\right) \end{aligned}$$

1. Vérifier qu'il s'agit bien d'un difféomorphisme.
2. Calculer la base naturelle et les coefficients de la métrique en déduire que le système de coordonnées est orthogonal et en déduire la base orthonormal associé.
3. Que vaut le gradient dans un tel système de coordonnée.

Connexion covariantes de Koszul

Pour les tenseurs, nous allons présenter la notion de connexion covariantes de Koszul. Pour l'heure encore très peu présente en physique, elle généralise la notion de connexion de Levi-civita qui sera réellement l'objet de nos développements. En effet les équations de la relativité général sont écrites à l'aide de la connexion de Levi-civita qui est une connexion de Koszul particulière. Les connexions de Koszul sont associés à une quantités géométriques qui s'annule avec l'usage d'une connexion de Levi-civita : la torsion. De nouvelle théorie de la gravitation incluent cette notion dans leurs développements. De plus tel qu'elle est construite la notion de connexion covariantes de Koszul, s'affranchit de l'usage d'une base particulière. C'est pour ces raisons que nous avons jugé bon d'en parler tout de même. Ces parties sont hors programme, le lecteur peut une fois lut la définition de la connexion de Koszul reprendre sa lecture en sous-sous-section VI.3.1

Commençons par la définition.

Définition VI.3.2 (Connexion covariante de Koszul). *Une connexion de Koszul ∇ est un opérateur linéaire qui à un tenseur $T \in \mathcal{T}_k^l \mathcal{M}$ associe un tenseur $\nabla T \in \mathcal{T}_{k+1}^l \mathcal{M}$,*

$$\begin{array}{ccc} \nabla & : & \mathcal{T}_k^l \mathcal{M} \\ & & \mapsto \end{array} \begin{array}{c} \nabla T \in \mathcal{T}_{k+1}^l \mathcal{M} \\ \nabla T \end{array}$$

vérifiant la propriété de Leibniz vis-à-vis de la multiplication par un scalaire, c'est-à-dire pour que tout vecteur v tenseurs T et U ,

$$\nabla_v fT = (v(f)) \cdot T + f \cdot \nabla_v T$$

où l'on a noté $\nabla_v T$ la contraction $v \dashv \nabla T$ avec un champ de vecteur sur l'indice "crée" par l'opérateur. Il vérifie également une règle de Leibniz pour la multiplication tensoriel,

$$\nabla_v (T \otimes U) = (\nabla_v T) \otimes U + T \otimes (\nabla_v U)$$

Tenseur parrallèle On définit également la notion de tenseur parallèle vis-à-vis d'une connexion.

Définition VI.3.3 (Tenseur parallèle). *Soit $T \in \mathcal{T}_k^l \mathcal{M}$, on dis que T est parallèle vis-à-vis d'une connexion de Koszul ∇ si et seulement si le champ $\nabla T = 0$ sur toute la variété.*

Christoffel d'une base dans une connexion de Koszul Que donne les propriétés de la connexion de Koszul sur la décomposition d'un vecteur dans une base ? Posons $(e_k)_{k \in \llbracket 0, n \rrbracket}$ une base de l'espace tangent que nous supposerons lisse, alors tout vecteur $v \in \mathcal{T} \mathcal{M}$ se décompose dans cette base $v = v^k e_k$. En utilisant linéarité et Leibniz, on a,

$$\nabla v = (dv^\alpha + \omega^\alpha_\beta \sigma^\beta) e_\alpha$$

où on a posé,

$$\omega^\alpha_\beta e_\alpha = \nabla e_\beta \in \mathcal{T}_1^1 \mathcal{M} \quad \Rightarrow \quad \nabla e_\beta = \Gamma^\alpha_{\gamma\beta} e_\alpha \otimes dx^\gamma$$

ce qui donne en décomposant sur les formes coordonnées de la différentiel sur la base $(e_k)_{k \in \llbracket 0, n \rrbracket}$,

$$\nabla v = (\partial_\gamma v^\alpha + \omega^\alpha_\beta \sigma^\beta) dx^\gamma \otimes e_\alpha$$

les composantes $\Gamma^\alpha_{\gamma\beta} \in \mathbb{R}$ sont ce qu'on appelle *les symboles de Christoffel de la base $(e_k)_{k \in \llbracket 0, n \rrbracket}$ dans la connexion de Koszul ∇* . Dans une connexion de Koszul ces composantes n'ont aucune raison d'être symétriques sur leurs indices du bas. Nous verrons plus bas que cette propriété est lié à la torsion de la connexion considéré.

Notons e_*^α les formes coordonnées associés à la base $(e_k)_{k \in \llbracket 0, n \rrbracket}$. Alors l'application de la connexion de Koszul au formes coordonnés ∇dx^α est un tenseur appartenant à $\mathcal{T}_2^0 \mathcal{M}$ et donc se décompose donc naturellement dans la base $(e_*^\alpha \otimes e_*^\beta)_{\alpha, \beta=1 \dots n}$, on note,

$$\nabla dx^\alpha = -\gamma^\alpha_{\gamma\delta} e_*^\gamma \otimes e_*^\delta.$$

Les quantités $\gamma^\alpha_{\gamma\delta}$ s'appelle *les symboles de Christoffel de la base dual $(e_*^\alpha)_{k \in \llbracket 0, n \rrbracket}$ dans la connexion de Koszul ∇* . Ce qui nous permet de déterminer l'expression de la connexion des formes linéaires $\omega = \omega_\alpha e_*^\alpha$,

$$\nabla \omega = (\partial_\beta \omega_\alpha - \omega_\delta \gamma^\delta_{\alpha\beta}) e_*^\alpha \otimes e_*^\beta$$

À l'aide des quantités construites et de la règle de Leibniz sur les produits tensoriels, nous pouvons construire l'expression de la connexion de Koszul d'un tenseur de variance quelconque.

Soit $\mathbf{T} \in \mathcal{T}\left(\begin{smallmatrix} l \\ k \end{smallmatrix}\right) \mathcal{M}$ ce tenseur se décompose dans $(\mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \mathbf{e}_*^{j_l})$ et donc,

$$\begin{aligned} \nabla \mathbf{T} &= \nabla [T^{i_1, \dots, i_k}{}_{j_1, \dots, j_l} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \mathbf{e}_*^{j_l}] \\ &= \partial_\alpha T^{i_1, \dots, i_k}{}_{j_1, \dots, j_l} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^\alpha \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \mathbf{e}_*^{j_l} \\ &\quad + \sum_{m=1}^k T^{i_1, \dots, i_k}{}_{j_1, \dots, j_l} \mathbf{e}_{i_1} \otimes \dots \otimes \Gamma_{i_m \alpha}^\delta e_\delta \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^\alpha \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \mathbf{e}_*^{j_l} \\ &\quad - \sum_{p=1}^l T^{i_1, \dots, i_k}{}_{j_1, \dots, j_l} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^\alpha \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \gamma_{\alpha \delta}^{j_p} \otimes \dots \otimes \mathbf{e}_*^{j_l} \\ &= \left[\partial_\alpha T^{i_1, \dots, i_k}{}_{j_1, \dots, j_l} + \sum_{m=1}^k T^{i_1, \dots, \delta, \dots, i_k}{}_{j_1, \dots, j_l} \Gamma_{\delta \alpha}^{i_m} - \sum_{p=1}^l T^{i_1, \dots, i_k}{}_{j_1, \dots, \delta, \dots, j_l} \gamma_{\alpha j_p}^\delta \right] \\ &\quad \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^\alpha \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \mathbf{e}_*^{j_l} \end{aligned}$$

où en changeant les indices muet afin de faire apparaître les même élément de la base conduit à,

$$\nabla \mathbf{T} = \left[\partial_\alpha T^{i_1, \dots, i_k}{}_{j_1, \dots, j_l} + \sum_{m=1}^k T^{i_1, \dots, \delta, \dots, i_k}{}_{j_1, \dots, j_l} \Gamma_{\delta \alpha}^{i_m} - \sum_{p=1}^l T^{i_1, \dots, i_k}{}_{j_1, \dots, \delta, \dots, j_l} \gamma_{\alpha j_p}^\delta \right] \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_k} \otimes \mathbf{e}_*^\alpha \otimes \mathbf{e}_*^{j_1} \otimes \dots \otimes \mathbf{e}_*^{j_l} \quad (\text{VI.156})$$

Crochet de commutation On a vu plus haut que les vecteurs de l'espace tangent pouvait se définir par leurs actions sur les champs scalaires. Pour tout champ de vecteur \mathbf{v}, \mathbf{w} , on note $\mathbf{x} = [\mathbf{v}, \mathbf{w}]$ le vecteur correspondant au commutateur des actions des champs sur un champs scalaire :

$$\mathbf{x}(f) = [\mathbf{v}, \mathbf{w}](f) = \mathbf{v}(\mathbf{w}(f)) - \mathbf{w}(\mathbf{v}(f))$$

ce qui, dans la base naturelle lié à un atlas et puisque les composantes des vecteurs sont des fonction lisse des coordonnées, donne,

$$\mathbf{x} = [\mathbf{v}, \mathbf{w}] = (v^\alpha \partial_\alpha w^\beta - w^\alpha \partial_\alpha v^\beta) \partial_\beta$$

Torsion dans une connexion de Koszul

Définition VI.3.4 (Torsion). *On appelle torsion le tenseur de $\mathbf{T}^\nabla \in \mathcal{T}\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right) \mathcal{M}$ construit par la différence entre la commutation construit à l'aide de la connexion de Koszul et le crochet de commutation, c'est-à-dire le tenseur vérifiant pour tout champ de vecteur \mathbf{v}, \mathbf{w} ,*

$$\mathbf{T}^\nabla(\cdot, \mathbf{v}, \mathbf{w}) = \nabla_{\mathbf{v}} \mathbf{w} - \nabla_{\mathbf{w}} \mathbf{v} - [\mathbf{v}, \mathbf{w}]$$

Les composantes $\Gamma_{\gamma \beta}^\alpha$ ne soit pas forcement symétrique sur les indices du bas. On a,

$$\nabla_{\mathbf{e}_\alpha} \mathbf{e}_\beta - \nabla_{\mathbf{e}_\beta} \mathbf{e}_\alpha = [\Gamma_{\alpha \beta}^\delta - \Gamma_{\beta \alpha}^\delta] \mathbf{e}_\delta$$

dans le cas où l'on travaille avec la base naturelle lié à l'atlas, on a,

$$[\partial_\alpha, \partial_\beta] = \mathbf{0}$$

et donc

$$\mathbf{T}^\nabla(\cdot, \partial_\alpha, \partial_\beta) = [\Gamma_{\alpha \beta}^\delta - \Gamma_{\beta \alpha}^\delta] \partial_\delta.$$

Ainsi,

Proposition VI.3.5. La torsion \mathbf{T}^∇ d'une connexion de Koszul ∇ est nul si et seulement si les symboles de Christoffel $(\Gamma_{\alpha \beta}^\gamma)_{\alpha, \beta, \gamma=1, \dots, n}$ d'une base naturel dans la connexion de Koszul en question sont symétriques par rapport à leurs indices du bas.

Connexion trivial

Cette partie assez technique, va permettre de construire la connexion trivial à l'aide d'un plongement ce qui permettra d'utiliser la différentiel pour construire cette connexion. L'existence de tel plongement est assuré par le théorème de Whitney. Nous verrons ensuite quelques propriétés de cette connexion. La connexion trivial est la connexion qui se substituera à la notion de différentielle. Comme pour la différentielle, la connexion trivial permet de faire des "développement limité à l'ordre 1", c'est pour cela qu'elle sera central. Cette partie est relativement importante. En effet, nous montrerons que la connexion de Levi-Civita (celle qui est utilisé en relativité général) correspond effectivement à cette connexion trivial.

Construction de la connexion trivial à l'aide d'un plongement On considère ici \mathcal{M} une variété de dimension n qui est un plongement dans un espace de dimension $N \geq n$, c'est-à-dire qu'il existe un ouvert $\mathcal{U} \subset \mathbb{R}^n$ et $\phi : \mathcal{U} \rightarrow \mathcal{M} \subset \mathbb{R}^N$ tel que ϕ soit un plongement⁷.

On cherche à construire une connexion de Koszul particulière qu'on appellera connexion trivial (c'est quand on utilise ce mots que cela ne l'est pas...). Dans une tel variété, la base naturelle $(\partial_\alpha = \partial_\alpha \phi)_{\alpha=1\dots n}$ associé à ce plongement et les formes coordonnées associé $(dx^\alpha)_{\alpha=1\dots n}$ peuvent être décomposé dans la base canonique et la base dual canonique de \mathbb{R}^N . Notons $(e_k)_{k=1\dots N}$ la base canonique et $(e_\star^k)_{k=1\dots N}$ la base dual associé, alors $\forall x \in \mathcal{U}, \mathcal{T}_{\phi(x)}\mathcal{M} = \text{vect}\{(\partial_\alpha)_{\alpha=1\dots n}\} \subset \text{vect}\{(e_\star^k)_{k=1\dots N}\}$ et $\mathcal{T}_{\phi(x)}^*\mathcal{M} = \text{vect}\{(dx^\alpha)_{\alpha=1\dots n}\} \subset \text{vect}\{(e_\star^k)_{k=1\dots N}\}$.

On a donc $\forall P \in \mathcal{M}$, l'existence d'un supplémentaire orthogonal,

$$\mathbb{R}^N = \mathcal{T}_P\mathcal{M} \overset{\perp}{\oplus} \mathcal{T}_P\mathcal{M}^\perp$$

Alors on étant le domaine de définition des formes coordonnées $(dx^\alpha)_{\alpha=1\dots n}$ à l'espace \mathbb{R}^N ,

$$\begin{aligned} dx^\alpha &: \mathcal{T}_P\mathcal{M} \overset{\perp}{\oplus} \mathcal{T}_P\mathcal{M}^\perp \rightarrow \mathbb{R} \\ \mathbf{x} = \mathbf{x}_T + \mathbf{x}_N &\mapsto dx^\alpha(\mathbf{x}_T) \end{aligned}$$

Alors les applications,

$$\begin{aligned} \mathcal{P}_0^1(P) &: \mathbb{R}^N \rightarrow \mathcal{T}_P\mathcal{M} & \mathcal{P}_1^0(P) &: \mathbb{R}^{N,*} \rightarrow \mathcal{T}_P\mathcal{M}^* \\ \mathbf{V} &\mapsto dx^\alpha(\mathbf{V}) \partial_\alpha & \boldsymbol{\omega} &\mapsto \boldsymbol{\omega}(\partial_\alpha) dx^\alpha \end{aligned}$$

définissent la projection orthogonal sur l'espace tangent et sur l'espace dual à l'espace tangents des vecteurs et des formes. Ainsi quelque soit la variance k, l considéré et quelque soit le point $x \in \mathcal{U}$ on a,

$$\mathcal{T}_{\phi(x)} \left(\begin{smallmatrix} l \\ k \end{smallmatrix} \right) \mathcal{M} \subset \overbrace{\mathbb{R}^N \otimes \dots \otimes \mathbb{R}^N}^{l \text{ fois}} \otimes \overbrace{\mathbb{R}^{N,*} \otimes \dots \otimes \mathbb{R}^{N,*}}^{k \text{ fois}}$$

ce qui nous permet de disposer de la multiplication par un scalaire et de l'addition sur $\mathbb{R}^N \otimes \dots \otimes \mathbb{R}^N \otimes \mathbb{R}^{N,*} \otimes \dots \otimes \mathbb{R}^{N,*}$ qui ne dépend pas du point et sommer des tenseurs évalué en des points différents. Enfin on définie un projecteur sur $\mathcal{T}_P \left(\begin{smallmatrix} l \\ k \end{smallmatrix} \right) \mathcal{M}$:

$$\begin{aligned} \mathcal{P}_k^l(P) &: \overbrace{\mathbb{R}^N \otimes \dots \otimes \mathbb{R}^N}^{l \text{ fois}} \otimes \overbrace{\mathbb{R}^{N,*} \otimes \dots \otimes \mathbb{R}^{N,*}}^{k \text{ fois}} \rightarrow \mathcal{T}_P \left(\begin{smallmatrix} l \\ k \end{smallmatrix} \right) \mathcal{M} \\ \mathbf{T} = T_{j_1, \dots, j_k}^{i_1, \dots, i_l} e_{i_1} \otimes \dots \otimes e_{i_l} \otimes e_\star^{j_1} \otimes \dots \otimes e_\star^{j_k} &\mapsto T_{j_1, \dots, j_k}^{i_1, \dots, i_l} \mathcal{P}_0^1(P)(e_{i_1}) \otimes \dots \otimes \mathcal{P}_0^1(P)(e_{i_l}) \otimes \\ &\quad \mathcal{P}_1^0(P)(e^{j_1,*}) \otimes \dots \otimes \mathcal{P}_1^0(P)(e_\star^{j_k}) \end{aligned}$$

Et donc pour tout champ de tenseur \mathbf{T} , l'application $\mathbf{T} \circ \phi : \mathcal{U} \rightarrow \overbrace{\mathbb{R}^N \otimes \dots \otimes \mathbb{R}^N}^{l \text{ fois}} \otimes \overbrace{\mathbb{R}^{N,*} \otimes \dots \otimes \mathbb{R}^{N,*}}^{k \text{ fois}}$ est donc définie à l'arrivé et au départ dans des sous-espaces de type \mathbb{R}^M , il y a donc un sens à parler de sa différentielle. Néanmoins, il est nécessaire de la composer avec \mathcal{P}_k^l car à priori pour $\mathbf{x} \in \mathcal{U}$ et $\mathbf{h} \in \mathbb{R}^n$ $d(\mathbf{T} \circ \phi)(x)\mathbf{h} \in \mathbb{R}^N \otimes \dots \otimes \mathbb{R}^N \otimes \mathbb{R}^{N,*} \otimes \dots \otimes \mathbb{R}^{N,*}$. De plus, pour construire une connexion de Koszul $\nabla_v \mathbf{T}$, il faut que la direction v appartienne à l'espace tangent. Si $v \in \mathcal{T}\mathcal{M}$ alors l'accroissement correspondant dans \mathcal{U} est $(d\phi(\phi^{-1}(P)))^{-1}v$ (où $d\phi$ est définie en Eq.VI.3.1). On peut donc construire la connexion trivial par l'opérateur défini pour tout $v \in \mathcal{T}\mathcal{M}$ par,

$$\begin{aligned} \nabla_v &: \mathcal{T}_k^l \mathcal{M} \rightarrow \mathcal{T}_k^l \mathcal{M} \subset \mathbb{R}^{N(k+l)} \\ \mathbf{T} &\mapsto P \in \mathcal{M} \mapsto \mathcal{P}_k^l(P) \left[d(\mathbf{T} \circ \phi)(\phi^{-1}(P)) (d\phi(\phi^{-1}(P)))^{-1} v(P) \right] \end{aligned}$$

7. L'application ϕ doit est lisse, injective et sa différentiel de rang n

Il s'agit tout simplement de la projection de la différentiel induit par le plongement.

Le fait que la connexion trivial soit une connexion de Koszul découle des propriétés de la différentiel. On a donc pour tout $x \in \mathcal{U}$ et pour tout $\mathbf{h} \in \mathbb{R}^n$,

$$\nabla_{d\phi(x)\mathbf{h}} \mathbf{T}(\phi(x)) = \mathcal{P}_k^l d(\mathbf{T} \circ \phi)(x) \mathbf{h}$$

Remarque VI.3.6. Cette construction est assez général en effet le théorème de Whitney assure l'existence d'un plongement pour toute les variétés. Le théorème de Nash va plus loin et garantit le caractère isométrique et C^3 d'un plongement qui lui aussi existe. Une fois le plongement définis, de manière similaire à ce qui est fais au dessus on peut à l'aide de l'atlas sur la variété construire des applications semblable à ϕ qui constitueront l'atlas correspondant à la variété plongé. A partir de là il n'est pas difficile de construire la connexion trivial, en montrant que la définitions n'est pas ambigu sur les zones de recollement. Et enfin,achever la construction en la transporter en retour dans la variété. Dans le cas d'un plongement isométrique, on peut également montrer que cette connexion est une connexion de Riemann. Dans la suite on admettra donc qu'une telle connexion peut être construite sur toutes variétés.

Dans le cas d'un champs de vecteur \mathbf{v} sur une surfaces plongé dans \mathbb{R}^3 , la partie orthogonal à l'espace tangents de l'accélération du champs $\nabla_{\mathbf{v}} \mathbf{v}$ pouvait être calculé à l'aide de la seconde forme fondamental (Voir Eq.V.1).

Propriétés de la connexion trivial. Par souci de simplicité et sans perte de généralité, on continue à se placer dans les cas où la variété \mathcal{M} de dimension n qui est un plongement dans \mathbb{R}^N avec $N \geq n$. On notera ∇ la connexion trivial définie au dessus. La connexion trivial peut s'exprimer à l'aide de courbe paramètre plongé dans la variété,

Proposition VI.3.7. On sait que pour tout point P et tout vecteur $\mathbf{v} \in T_P \mathcal{M}$, il existe $\Gamma :]-\epsilon, \epsilon[\rightarrow \mathcal{M}$ tel que $\Gamma(0) = P$ et $\dot{\Gamma} = \mathbf{v}$, alors pour tout tenseur $\mathbf{T} \in \mathcal{T}(^l_k)$

$$\nabla_{\mathbf{v}} \mathbf{T}(P) = \mathcal{P}_k^l \frac{d}{dt} \mathbf{T} \circ \Gamma \Big|_{t=0}$$

Démonstration. La courbe paramétrée peut se transporter dans l'espace de départ \mathcal{U} à l'aide du plongement $\gamma = \phi^{-1} \circ \Gamma :]-\epsilon, \epsilon[\rightarrow \mathcal{U}$ et donc,

$$\begin{aligned} \mathcal{P}_k^l \frac{d\mathbf{T} \circ \Gamma}{dt} &= \mathcal{P}_k^l \frac{d\mathbf{T} \circ \phi \circ \gamma}{dt} \\ &= \mathcal{P}_k^l d(\mathbf{T} \circ \phi)(\gamma(t)) \dot{\gamma}(t) \\ &= \nabla_{d\phi(\gamma(t))\gamma'(t)} \mathbf{T}(\phi \circ \gamma(t)) \\ &= \nabla_{\Gamma'(t)} \mathbf{T}(\Gamma(t)) \end{aligned}$$

où on a utilisé le fait que $\dot{\Gamma}(t) = d\phi(\gamma(t))\dot{\gamma}(t)$. On conclut en évaluant en 0. ■

Proposition VI.3.8. Soit pour tout champ de vecteur $\mathbf{v}, \mathbf{w} \in T_P \mathcal{M}$, de 1-forme différentielle ω et tenseur $\mathbf{T} \in \mathcal{T}(^l_k)$ la connexion trivial vérifie,

$$\nabla_{\mathbf{w}} [\mathbf{T}(\dots, \mathbf{v}, \dots)] = \nabla_{\mathbf{w}} \mathbf{T}(\dots, \mathbf{v}, \dots) + \mathbf{T}(\dots, \nabla_{\mathbf{v}} \mathbf{v}, \dots)$$

et

$$\nabla_{\mathbf{w}} [\mathbf{T}(\dots, \omega, \dots)] = \nabla_{\mathbf{w}} \mathbf{T}(\dots, \omega, \dots) + \mathbf{T}(\dots, \nabla_{\mathbf{v}} \omega, \dots)$$

Démonstration. Il existe $\epsilon > 0$ et une courbe paramétrée $\Gamma :]-\epsilon, \epsilon[\rightarrow \mathcal{M}$ tel que $\Gamma(0) = P$ et $\dot{\Gamma} = \mathbf{v}$. Ce qui conduit à,

$$\begin{aligned} \nabla_{\mathbf{w}} [\mathbf{T}(\dots, \mathbf{v}, \dots)] &= \lim_{t \rightarrow 0} \frac{1}{t} [\mathbf{T}(\Gamma(t))(\dots, \mathbf{v}(\Gamma(t)), \dots) - \mathbf{T}(P)(\dots, \mathbf{v}(P), \dots)] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} [\mathbf{T}(\Gamma(t))(\dots, \mathbf{v}(\Gamma(t)), \dots) - \mathbf{T}(P)(\dots, \mathbf{v}(\Gamma(t)), \dots)] \\ &\quad + \lim_{t \rightarrow 0} \frac{1}{t} [\mathbf{T}(P)(\dots, \mathbf{v}(\Gamma(t)), \dots) - \mathbf{T}(P)(\dots, \mathbf{v}(P), \dots)] \\ &= \left[\lim_{t \rightarrow 0} \frac{1}{t} (\mathbf{T}(\Gamma(t)) - \mathbf{T}(P))(\dots, \mathbf{v}(\Gamma(t)), \dots) + \mathbf{T}(P)(\dots, \lim_{t \rightarrow 0} \frac{1}{t} (\mathbf{v}(\Gamma(t)) - \mathbf{v}(P)), \dots) \right] \\ &= \nabla_{\mathbf{w}} \mathbf{T}(\dots, \mathbf{v}, \dots) + \mathbf{T}(\dots, \nabla_{\mathbf{v}} \mathbf{v}, \dots) \end{aligned}$$

On démontre la chose de la même manière pour la contraction avec une 1-forme différentielle. ■

Proposition VI.3.9. La métrique induite par le plongement est parallèle vis-à-vis de la connexion trivial.

Démonstration. Rappelons que la métrique induite par le plongement est tout simplement,

$$g = \langle \partial_\mu | \partial_\nu \rangle dx^\mu \otimes dx^\nu$$

où $\langle . | . \rangle$ est le produit scalaire canonique de \mathbb{R}^N . Pour tout champ de vecteur $u, v, w \in \mathcal{T}\mathcal{M}$, on considère pour tout point P la courbe Γ qui est en $P = \Gamma(0)$ et $\dot{\Gamma}(0)$. De tel sort qu'on ai,

$$\begin{aligned} \nabla_w g(u, v) &= \nabla_w \langle u | v \rangle \\ &= \frac{d}{dt} (\langle u \circ \Gamma(t) | v \circ \Gamma(t) \rangle) \Big|_{t=0} \\ &= \left\langle \frac{d}{dt} (u \circ \Gamma(t)) \Big|_{t=0} \middle| v \right\rangle + \left\langle u \middle| \frac{d}{dt} (v \circ \Gamma(t)) \Big|_{t=0} \right\rangle \\ &= \left\langle \mathcal{P}_0^1 \frac{d}{dt} (u \circ \Gamma(t)) \Big|_{t=0} \middle| v \right\rangle + \left\langle u \middle| \mathcal{P}_0^1 \frac{d}{dt} (v \circ \Gamma(t)) \Big|_{t=0} \right\rangle \\ &= \langle \nabla_w u | v \rangle + \langle u | \nabla_w v \rangle \\ &= g(\nabla_w u, v) + g(u, \nabla_w v) \end{aligned}$$

où le passage à la quatrième ligne provient du fait que les vecteurs $u, v \in \mathcal{T}\mathcal{M}$ et donc $\forall x \in \mathbb{R}^N$ on a $\langle x | u \rangle = \langle \mathcal{P}_0^1 x | u \rangle$. Pour finir, on utilise la règle de Leibniz à propos d'une contraction avec un vecteur de la proposition précédentes pour aboutir, à

$$\begin{aligned} \nabla_w g(u, v) &= \nabla_w g(u, v) + g(\nabla_w u, v) + g(u, \nabla_w v) = g(\nabla_w u, v) + g(u, \nabla_w v) \\ \Rightarrow \quad \nabla_w g(u, v) &= 0 \Rightarrow \quad \nabla g = 0 \end{aligned}$$

La dernière implication provenant du fait que l'égalité précédentes est vrai pour tout champs de vecteur $u, v, w \in \mathcal{T}\mathcal{M}$. ■

Proposition VI.3.10. Les symboles de Christoffel de la base naturelle $(\partial_\alpha)_{\alpha \in \llbracket 0, n \rrbracket}$ et de sa base dual associé $(dx^\alpha)_{\alpha \in \llbracket 0, n \rrbracket}$ dans la connexion trivial sont égaux,

$$\gamma_{\gamma\beta}^\alpha = \Gamma_{\gamma\beta}^\alpha$$

Démonstration. C'est une simple conséquence de la proposition VI.3.8 appliquée à $T = dx^\alpha$, $v = \partial_\beta$ et $w = \partial_\gamma$, en effet,

$$\begin{aligned} \nabla_{\partial_\gamma} (dx^\alpha(\partial_\beta)) &= (\nabla_{\partial_\gamma} dx^\alpha)(\partial_\beta) + dx^\alpha(\nabla_{\partial_\gamma} \partial_\beta) \\ \partial_\gamma \delta_\beta^\alpha &= -\gamma_{\gamma\delta}^\alpha dx^\delta(\partial_\beta) + dx^\alpha(\Gamma_{\gamma\beta}^\delta \partial_\delta) \\ 0 &= -\gamma_{\gamma\beta}^\alpha + \Gamma_{\gamma\beta}^\alpha \end{aligned}$$

■

Proposition VI.3.11. La torsion associé à la connexion trivial est nul.

Démonstration. On appelle $(e_\alpha)_{\alpha=1\dots n}$ la base canonique de \mathbb{R}^n et $(\partial_\alpha = \partial_\alpha \varphi)_{\alpha=1\dots n}$ la base naturel associé au plongement et on considère pour tout $x \in \mathcal{U}$ pour tout $\epsilon \in \mathbb{R}$ suffisamment petit et pour tout $\alpha, \beta \in \llbracket 0, n \rrbracket$ les points,

$$\begin{aligned} M &= \phi(x) & M_\alpha &= \phi(x + \epsilon e_\alpha) \\ M_\beta &= \phi(x + \epsilon e_\beta) & M_{\alpha,\beta} &= \phi(x + \epsilon e_\alpha + \epsilon e_\beta) \end{aligned}$$

Alors les propriétés de la différentielle permettent, en développant d'abord dans la direction e_β puis dans la direction e_α , d'aboutir à,

$$\begin{aligned} M_{\alpha\beta} &= M_\alpha + \epsilon \partial_\beta(M_\alpha) + \frac{\epsilon^2}{2} \partial_\beta^2 \phi(M) + o(\epsilon^3) \\ &= M + \epsilon [\partial_\alpha(M) + \partial_\beta(M_\alpha)] + \frac{\epsilon^2}{2} [\partial_\alpha^2 \phi(M) + \partial_\beta^2 \phi(M)] + o(\epsilon^2) \end{aligned} .$$

Par symétrie, en développant d'abord dans la direction e_α puis dans la direction e_β ,

$$M_{\alpha\beta} = M + \epsilon [\partial_\beta(M) + \partial_\alpha(M_\beta)] + \frac{\epsilon^2}{2} [\partial_\alpha^2\phi(M) + \partial_\beta^2\phi(M)] + o(\epsilon^2)$$

Or les propriétés de la différentielles assure que,

$$\begin{cases} \partial_\beta(M_\alpha) = \partial_\beta\phi(x + \epsilon e_\alpha) = \partial_\beta\phi(x) + \epsilon d\partial_\beta\phi(x)e_\alpha + o(\epsilon) \\ = \partial_\beta\phi(x) + \epsilon \nabla_{d\phi(x)}e_\alpha \partial_\beta\phi(x) + \epsilon n_{\beta,\alpha} + o(\epsilon) \\ = \partial_\beta(M) + \epsilon \nabla_{\partial_\alpha}\partial_\beta(M) + \epsilon n_{\beta,\alpha} + o(\epsilon) \\ \partial_\alpha(M_\beta) = \partial_\alpha(M) + \epsilon \nabla_{\partial_\beta}\partial_\alpha(M) + \epsilon n_{\alpha,\beta} + o(\epsilon) \end{cases}$$

où $n_{\alpha,\beta}$ et $n_{\beta,\alpha}$ sont deux vecteurs orthogonaux à l'espace tangent. Ainsi, on a,

$$\begin{aligned} M_{\alpha\beta} &= M + \epsilon [\partial_\beta(M) + \partial_\alpha(M)] + \epsilon^2 \nabla_{\partial_\beta}\partial_\alpha(M) + \epsilon^2 n_{\alpha,\beta} + \frac{\epsilon^2}{2} [\partial_\alpha^2\phi(M) + \partial_\beta^2\phi(M)] + o(\epsilon^2) \\ &= M + \epsilon [\partial_\beta(M) + \partial_\alpha(M)] + \epsilon^2 \nabla_{\partial_\alpha}\partial_\beta(M) + \epsilon^2 n_{\beta,\alpha} + \frac{\epsilon^2}{2} [\partial_\alpha^2\phi(M) + \partial_\beta^2\phi(M)] + o(\epsilon^2) \end{aligned}$$

Et donc, puisque $\nabla_{\partial_\beta}\partial_\alpha$ et $\nabla_{\partial_\alpha}\partial_\beta$ sont des vecteurs de l'espace-tangents, on obtient finalement.

$$\nabla_{\partial_\beta}\partial_\alpha = \nabla_{\partial_\alpha}\partial_\beta$$

La propositions VI.3.5 permet de conclure. ■

Remarque VI.3.12. En examinant bien la démonstration, l'on voit que c'est la possibilité d'utiliser la connexion trivial dans développement limité d'ordre 1 qui implique l'annulation de la torsion associé à la connexion trivial.

Connexion de Riemann et de Levi-Civita

Riemann En terme géométrique, les connexions de Koszul physiquement intéressantes sont celles pour lequel, la variation d'une "longueur élémentaire" reposant sur un déplacement élémentaire n'est déterminé que par la variation de ce vecteur élémentaire. Puisque il y a équivalence entre forme quadratique et forme bilinéaire associé ce réquisit nous amènes à définir la notion de connexion de Riemann.

Définition VI.3.13 (Connection de Riemann). *Dans une variété (\mathcal{M}, g) , une connexion de Koszul ∇ est une connexion de Riemann si est seulement si pour tout champs de vecteur $u, v, w \in T\mathcal{M}$ on ait,*

$$\nabla_w(g(u, v)) = g(\nabla_w u, v) + g(u, \nabla_w v)$$

On demandera de plus que les symboles de Christoffel $\Gamma_{\alpha\beta}^\gamma$ de la base naturelle $(\partial_\alpha)_{\alpha \in [0, n]}$ et des formes coordonnées $\gamma_{\alpha\beta}^\gamma$ correspondante $(dx^\alpha)_{\alpha \in [0, n]}$ dans la connexion ∇ soit égaux. Ce qui implique, combiné à la relation précédentes le fait que la métrique soit une tenseur parallèle vis-à-vis de ∇

$$\nabla g = 0$$

Remarque VI.3.14. Sachant qu'une connexion de Koszul vérifie l'Eq.VI.156 et que les symboles de Christoffel de la base naturelle et des formes coordonnées correspondante dans la connexion ∇ sont égaux, on en déduit qu'une connexion de Riemann est entièrement déterminé par par la donnés des Christoffels de la base naturelle $\Gamma_{\beta\gamma}^\alpha$.

Soit $u, v, w \in T\mathcal{M}$, on dis que u et v est transporté parallèlement par w vis-à-vis de la connexion de Koszul ∇ si et seulement si $\nabla_w u = 0$ et $\nabla_w v = 0$. Dans le cas d'une connexion de Riemann, cette propriété implique que $\nabla_w(g(u, v)) = 0$ et donc que le produit scalaire de ces vecteurs est conservés.

Les connexions de Riemann, bien que satisfaisante concernant la dérivation d'un produit scalaire, s'avère difficile à appréhender à cause de la torsion. En effet, toute connexion qui permet de faire des développements limité (comme dans la démonstration de la proposition Eq.VI.3.11), aura une torsion nul.

Connexion de Levi-Civita On a vu précédemment que les symboles de Christoffel d'une connexion de Koszul n'étais pas forcement symétrique par rapport au indices du bas, cela est également vrai des connexions de Riemann. Les connexions qui possèdent cette symétries sont appellées connexion de Levi-Civita.

Définition VI.3.15 (Connection de Levi-Civita). *Dans une variété $(\mathcal{M}, \mathbf{g})$, une connexion de Riemann ∇ est une connexion de Levi-Civita si et seulement si la torsion qui y est associée est nul, c'est-à-dire pour tout champs de vecteur $\mathbf{v}, \mathbf{w} \in \mathcal{T}\mathcal{M}$ on a,*

$$\nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{w}}\mathbf{v} = [\mathbf{v}, \mathbf{w}].$$

Dans la suite, dans une base naturel définie, pour tout tenseur $\mathbf{T} \in \mathcal{T}\binom{l}{k}\mathcal{M}$ on notera,

$$\nabla_{\partial_\alpha}\mathbf{T} = \nabla_\alpha\mathbf{T}$$

et en terme de composantes,

$$\nabla_\alpha\mathbf{T} = \nabla_\alpha T^{i_1, i_1, \dots, i_k}_{j_1, j_1, \dots, j_l} \partial_{i_1} \otimes \dots \otimes \partial_{i_k} \otimes \mathbf{d}x^{j_1} \dots \otimes \mathbf{d}x^{j_l}$$

L'annulation de la torsion implique en utilisant la proposition VI.3.5 la symétrie sur les indices du base des quantités de Christoffel $\Gamma^\alpha_{\beta\gamma} = \Gamma^\alpha_{\gamma\beta}$.

Théorème fondamental de la géométrie Riemannienne Le théorème fondamental de la géométrie Riemannienne affirme l'existence et l'unicité de cette connexion.

Théorème VI.3.16 (Fondamental de la géométrie Riemannienne). *Pour toute variété Riemannienne $(\mathcal{M}, \mathbf{g})$, sur $\mathcal{T}\mathcal{M}$ il existe une unique connexion de Levi-Civita ∇ . On l'appellera *dérivation covariante*.*

Démonstration. Unicité : Supposons que la connexion de Levi-Civita existe. Alors son caractère Riemannien implique que pour tout $\alpha, \beta, \gamma \in \llbracket 1, n \rrbracket$.

$$\begin{cases} \partial_\gamma g_{\alpha\beta} = g_{\mu\beta} \Gamma^\mu_{\gamma\alpha} + g_{\mu\alpha} \Gamma^\mu_{\gamma\beta} \\ \partial_\beta g_{\gamma\alpha} = g_{\mu\alpha} \Gamma^\mu_{\beta\gamma} + g_{\mu\gamma} \Gamma^\mu_{\beta\alpha} \\ \partial_\alpha g_{\beta\gamma} = g_{\mu\gamma} \Gamma^\mu_{\alpha\beta} + g_{\mu\beta} \Gamma^\mu_{\alpha\gamma} \end{cases}$$

Ce qui en additionnant les deux premières avant de retrancher la dernière, puis en utilisant la proposition VI.3.5 et l'annulation de la torsion d'une connexion de Levi-Civita aboutit à,

$$\partial_\gamma g_{\alpha\beta} + \partial_\beta g_{\gamma\alpha} - \partial_\alpha g_{\beta\gamma} = g_{\mu\beta} (\Gamma^\mu_{\gamma\alpha} - \Gamma^\mu_{\alpha\gamma}) + g_{\mu\gamma} (\Gamma^\mu_{\beta\alpha} - \Gamma^\mu_{\alpha\beta}) + g_{\mu\beta} (\Gamma^\mu_{\beta\gamma} + \Gamma^\mu_{\gamma\beta}) = 2g_{\mu\beta} \Gamma^\mu_{\beta\gamma}$$

Soit encore, en utilisant la métrique inverse $g^{\alpha\delta}$,

$$\Gamma^\delta_{\beta\gamma} = \frac{1}{2} g^{\delta\alpha} (\partial_\gamma g_{\alpha\beta} + \partial_\beta g_{\gamma\alpha} - \partial_\alpha g_{\beta\gamma}) \quad (\text{VI.157})$$

Ainsi tout les quantités de Christoffel sont entièrement déterminé par l'expression et les dérivées partielles des coefficients de la métrique : si la connexion de Levi-Civita existe, elle est unique et son expression est déterminé par l'expression de Christoffel ci-dessus.

Existence : On définit ∇ par l'expression Eq.VI.156, où les Christoffel de la base naturelle et des formes coordonnées associés sont égaux et sont définis à partir des dérivées partielles de la métrique par l'Eq.VI.157. C'est-à-dire qu'on posera,

$$\begin{aligned} \nabla &: \mathcal{T}_k^l\mathcal{M} \rightarrow \mathcal{T}_{k+1}^l\mathcal{M} \\ \mathbf{T} &\mapsto \nabla_\alpha T^{i_1, i_1, \dots, i_k}_{j_1, j_1, \dots, j_l} \partial_{i_1} \otimes \dots \otimes \partial_{i_k} \otimes \mathbf{d}x^\alpha \otimes \mathbf{d}x^{j_1} \otimes \dots \otimes \mathbf{d}x^{j_l} \end{aligned}$$

où on définira les composantes par,

$$\nabla_\alpha T^{i_1, i_1, \dots, i_k}_{j_1, j_1, \dots, j_l} = \partial_\alpha T^{i_1, \dots, i_k}_{j_1, \dots, j_l} + \sum_{m=1}^k T^{i_1, \dots, \delta, \dots, i_k}_{j_1, \dots, j_l} \Gamma_{\delta\alpha}^{i_m} - \sum_{p=1}^l T^{i_1, \dots, i_k}_{j_1, \dots, \delta, \dots, j_l} \Gamma_{\alpha j_p}^\delta$$

Il suffit alors de vérifier que l'opérateur ainsi construite est bien une connexion de Levi-Civita. Il faut tout d'abords que ce soit une connexion de Koszul. L'expression de ∇ , assure que ∇ soit un opérateur linéaire envoyant un élément de $\mathcal{T}\binom{l}{k}\mathcal{M}$ vers un élément de $\mathcal{T}\binom{l}{k+1}\mathcal{M}$. Les relations de Leibniz sont un peu technique à vérifié mais ne pose pas de difficulté majeur (on laissera l'étudiants motivé s'en charger). Enfin, l'expression des Christoffel assure $\mathbf{g} = 0$ et l'annulation de la torsion. Ainsi posé ∇ est donc une connexion de Levi-Civita. ■

Propriété de la dérivation covariante.

Corrolaire VI.3.17. La connexion trivial construite à partir d'un plongement isométrique est égale à la dérivation covariante. Ainsi la dérivée covariante possède toutes les propriétés de la connexion trivial. Rappelons les propriétés ne découlant pas immédiatement de la définitions d'une connexions de Levi-Civita :

- Pour tout $P \in \mathcal{M}$, pour tout vecteur $\mathbf{v} \in \mathcal{T}_P \mathcal{M}$ et pour tout tenseur $\mathbf{T} \in \mathcal{T}_k^l \mathcal{M}$

$$\nabla_{\mathbf{v}} \mathbf{T}(P) = \mathcal{P}_k^l \left. \frac{d\mathbf{T} \circ \gamma}{dt} \right|_{t=0}$$

où \mathcal{P}_k^l est le projecteur orthogonal sur $\mathcal{T}_k^l \mathcal{M}$ et $\gamma :]-\epsilon, \epsilon[\rightarrow \mathcal{M}$ est une courbe paramétrée de la variété vérifiant $\gamma(0) = P$ et $\dot{\gamma}(0) = \mathbf{v}$.

- Et on a pour tout champs de vecteur $\mathbf{v}, \mathbf{w} \in \mathcal{T}_P \mathcal{M}$, de forme ω et de tenseur $\mathbf{T} \in \mathcal{T}(\binom{l}{k}) \mathcal{M}$ la dérivée covariante vérifie,

$$\nabla_{\mathbf{w}} [\mathbf{T}(\dots, \mathbf{v}, \dots)] = \nabla_{\mathbf{w}} \mathbf{T}(\dots, \mathbf{v}, \dots) + \mathbf{T}(\dots, \nabla_{\mathbf{v}} \mathbf{v}, \dots)$$

et

$$\nabla_{\mathbf{w}} [\mathbf{T}(\dots, \omega, \dots)] = \nabla_{\mathbf{w}} \mathbf{T}(\dots, \omega, \dots) + \mathbf{T}(\dots, \nabla_{\mathbf{v}} \omega, \dots)$$

Démonstration. C'est une simple conséquence des propriétés de la connexion trivial, les propositions VI.3.9 et VI.3.10 assure son caractère de connexion Riemannienne et la proposition VI.3.11 assure qu'il s'agit d'une connexion de Levi-Civita. Or celles-ci étant unique elle est égale à la dérivation covariante.

L'expression en fonction de la dérivée avec la composition par un chemin est simplement la réécriture de la proposition VI.3.7. Enfin la dérivée covariante d'un tenseur agissant sur un vecteur ou une forme n'est que la réécriture de VI.3.8. ■

À titre pratique notons ici les expressions en terme de composantes de la dérivée covariantes,

Propriété VI.3.18. Soit pour tout champs de vecteur $\mathbf{v} \in \mathcal{T}$, pour toute forme $\omega \in \mathcal{T}^*$ et pour tout tenseur $\mathbf{T} \in \mathcal{T}(\binom{l}{k}) \mathcal{M}$ on a

$$\begin{cases} \nabla_{\alpha} v^{\beta} = \partial_{\beta} v^{\alpha} + v^{\gamma} \Gamma_{\beta\gamma}^{\alpha} & \nabla_{\alpha} \omega_{\beta} = \partial_{\alpha} \omega_{\beta} - \omega_{\gamma} \Gamma_{\beta\alpha}^{\gamma} \\ \nabla_{\alpha} T^{i_1, i_1, \dots, i_k}_{j_1, j_1, \dots, j_l} = \partial_{\alpha} T^{i_1, \dots, i_k}_{j_1, \dots, j_l} + \sum_{m=1}^k T^{i_1, \dots, \delta, \dots, i_k}_{j_1, \dots, j_l} \Gamma_{\delta\alpha}^{i_m} - \sum_{p=1}^l T^{i_1, \dots, i_k}_{j_1, \dots, \delta, \dots, j_l} \Gamma_{\alpha j_p}^{\delta} \end{cases}$$

Une opération souvent utilisé dans le calcul tensoriel est l'auto-contraction. Cette opération permet par exemple de définir et de calculer la trace d'un tenseur de $\mathcal{T}(\binom{1}{1}) \mathcal{M}$ (qui peut être considéré comme un endomorphisme).

Définition VI.3.19 (Auto-contraction d'un tenseur). Soit un tenseur $\mathbf{T} \in \mathcal{T}(\binom{l}{k}) \mathcal{M}$, pour tout $p \in [1, l]$ et tout $q \in [1, k]$, on appelle tenseur auto-contracté sur l'indice contrariant p et l'indice covariant q , le tenseur $C_q^p \mathbf{T}$ appartenant à $\mathcal{T}(\binom{l-1}{k-1}) \mathcal{M}$ et défini par,

$$\begin{aligned} C_q^p \mathbf{T} \left(\underbrace{\dots, \dots}_{l-1 \text{ forme}}, \underbrace{\dots, \dots}_{k-1 \text{ vecteur}} \right) &= \mathbf{T} \left(\dots, \underbrace{\mathbf{d}x^{\alpha}}_{\text{place } q}, \dots, \underbrace{\partial_{\alpha}}_{\text{place } p}, \dots \right) \\ &= T^{i_1, i_1, \dots, i_{q-1}, \alpha, i_{q+1}, \dots, i_k}_{j_1, j_1, \dots, j_{p-1}, \alpha, j_{p+1}, \dots, j_l} \partial_{i_1} \otimes \dots \otimes \partial_{i_{q-1}} \otimes \partial_{i_{q+1}} \otimes \dots \otimes \partial_{i_k} \otimes \mathbf{d}x^{j_1} \dots \otimes \mathbf{d}x^{j_{p-1}} \otimes \mathbf{d}x^{j_{p+1}} \otimes \dots \otimes \mathbf{d}x^{j_l} \end{aligned}$$

Cette opération commute avec la dérivation covariante.

Proposition VI.3.20. L'auto-contraction d'un tenseur commute avec la dérivation covariante, autrement dit pour tout champs de vecteur \mathbf{v} , pour tout tenseur $\mathbf{T} \in \mathcal{T}(\binom{l}{k}) \mathcal{M}$ et pour toute auto-contraction possible $(C_q^p)_{(p,q) \in [1,l] \times [1,k]}$, on a,

$$\nabla_{\mathbf{v}} C_q^p \mathbf{T} = C_q^p \nabla_{\mathbf{v}} \mathbf{T}$$

Démonstration. En effet, il s'agit juste d'utiliser la règle de Leibniz de la proposition VI.3.8 à propos de la contraction d'un tenseur.

$$\begin{aligned}
\nabla_{\mathbf{v}} C_q^p \mathbf{T} &= \nabla_{\mathbf{v}} [T(\dots, \dots, \mathbf{d}x^\alpha, \dots, \dots, \partial_\alpha, \dots)] \\
&= (\nabla_{\mathbf{v}} T)(\dots, \dots, \mathbf{d}x^\alpha, \dots, \dots, \partial_\alpha, \dots) + T(\dots, \dots, \nabla_{\mathbf{v}} \mathbf{d}x^\alpha, \dots, \dots, \partial_\alpha, \dots) + T(\dots, \dots, \mathbf{d}x^\alpha, \dots, \dots, \nabla_{\mathbf{v}} \partial_\alpha, \dots) \\
&= C_q^p \nabla_{\mathbf{v}} \mathbf{T} - v^\beta \Gamma_{\beta\gamma}^\alpha T(\dots, \dots, \mathbf{d}x^\gamma, \dots, \dots, \partial_\alpha, \dots) + v^\beta \Gamma_{\beta\alpha}^\gamma T(\dots, \dots, \mathbf{d}x^\alpha, \dots, \dots, \partial_\gamma, \dots) \\
&= C_q^p \nabla_{\mathbf{v}} \mathbf{T}
\end{aligned}$$

où pour le passage à la dernière ligne on a juste changé les indices muets. ■

Divergence... Une opération souvent effectuée car elle concerne l'expression mathématique de la conservation de quantités physiques est la divergence. Afin de calculer cette quantité on aura à calculer l'auto-contraction des symboles de Christoffels.

Proposition VI.3.21. On a,

$$\Gamma_{\alpha\beta}^\alpha = \partial_\beta \ln \sqrt{\pm g}$$

où g est le déterminant de la matrice $[g_{\alpha\beta}]$

Démonstration.

$$\begin{aligned}
\Gamma_{\alpha\beta}^\alpha &= \frac{1}{2} g^{\alpha\delta} (\partial_\alpha g_{\delta\beta} + \partial_\beta g_{\delta\alpha} - \partial_\delta g_{\alpha\beta}) \\
&= \frac{1}{2} g^{\alpha\delta} \partial_\alpha g_{\delta\beta} + \frac{1}{2} g^{\alpha\delta} \partial_\beta g_{\delta\alpha} - \frac{1}{2} g^{\alpha\delta} \partial_\delta g_{\alpha\beta} \\
&= \frac{1}{2} g^{\alpha\delta} \partial_\alpha g_{\delta\beta} + \frac{1}{2} g^{\alpha\delta} \partial_\beta g_{\delta\alpha} - \frac{1}{2} g^{\delta\alpha} \partial_\alpha g_{\delta\beta} \\
&= \frac{1}{2} g^{\alpha\delta} \partial_\beta g_{\delta\alpha}
\end{aligned}$$

Notons $\mathbf{G} = [g_{\alpha\beta}]$ la matrice carré de taille n contenant les éléments de lignes de la métrique dans la base naturel. Alors $\mathbf{G}^{-1} = [g^{\alpha\beta}]$. Alors,

$$\begin{aligned}
\partial_\mu \det \mathbf{G} &= d \det(\mathbf{G}) \partial_\mu \mathbf{G} \\
&= \text{Tr}({}^t \text{Com}(\mathbf{G}) \partial_\mu \mathbf{G}) \\
&= \det \mathbf{G} \text{Tr}(\mathbf{G}^{-1} \partial_\mu \mathbf{G}).
\end{aligned}$$

En effet $d \det(\mathbf{G}) = \text{Tr}({}^t \text{Com} \mathbf{G})$. En terme indiciel cette relation s'écrit,

$$\begin{aligned}
\partial_\mu g &= (\det \mathbf{G}) (\mathbf{G}^{-1} \partial_\mu \mathbf{G})^\alpha_\alpha \Rightarrow \partial_\mu \ln \sqrt{\pm g} = g^{\alpha\beta} \partial_\mu g_{\beta\alpha} \\
&= g g^{\alpha\beta} \partial_\mu g_{\beta\alpha}
\end{aligned}$$

D'où

$$\Gamma_{\alpha\beta}^\alpha = \frac{1}{2} g^{\alpha\beta} \partial_\mu g_{\beta\alpha} = \partial_\beta \ln \sqrt{\pm g}.$$

Le moins intervient dans les variétés pseudo-riemannienne ou la métrique n'est pas définie positive et que son déterminant est négatif. ■

Définition VI.3.22 (Divergence d'un vecteur). Pour tout vecteur $\mathbf{v} \in \mathcal{T}\mathcal{M}$ l'auto-contraction de $\nabla_{\mathbf{v}}$ est appelé divergence de \mathbf{v} . On note,

$$\text{div } \vec{\mathbf{v}} \hat{=} \nabla \cdot \mathbf{v} \hat{=} C_1^1 \nabla \mathbf{v}$$

Le résultat de la proposition VI.3.21, va nous permettre de calculer la divergence d'un vecteur.

Théorème VI.3.23. Alors on a,

$$\operatorname{div} \vec{v} \hat{=} \frac{1}{\sqrt{\pm g}} \partial_\mu (\sqrt{\pm g} v^\mu)$$

Dans un système de coordonnée orthonormal on obtient donc,

$$\operatorname{div} \vec{v} = \frac{1}{\prod_{k=1}^n h_k} \partial_\mu \left(\frac{\prod_{k=1}^n h_k}{h_\mu} v^\mu \right)$$

où $v^\hat{k} = v^k/k_k$ correspond aux coordonnées de v dans la base orthonormal associé à la base naturel orthogonal et h_k sont les élément de la diagonal de la métrique.

Démonstration. On a,

$$\begin{aligned} C_1^1 \nabla v g &= \nabla_\mu v^\mu \\ &= \partial_\mu v^\mu + v^\nu \Gamma_{\mu\nu}^\mu \\ &= \partial_\mu v^\mu + v^\nu \partial_\nu \ln \sqrt{\pm g} \\ &= \frac{1}{\sqrt{\pm g}} \partial_\mu (\sqrt{\pm g} v^\mu) \end{aligned}$$

■

Exercice 40. — Divergence en coordonnée parabolique.

On rappelle que dans $\mathbb{R}^3 - O_y$ munit de la carte les coordonnées paraboliques sont définie par

$$\begin{aligned} \psi : \mathbb{R} \times \mathbb{R}_{+,*} \times \mathbb{R} &\rightarrow \mathbb{R}^3 - O_y \\ (u, v, z) &\mapsto \left(uv, \frac{u^2 - v^2}{2}, z \right) \end{aligned}$$

Montrer alors que le divergence vaut,

$$\operatorname{div} \vec{V} = \frac{1}{u^2 + v^2} \left[\partial_u \left(\sqrt{u^2 + v^2} V^u \right) + \partial_v \left(\sqrt{u^2 + v^2} V^v \right) \right] + \partial_z V^z$$

Le résultat dans les coordonnée orthogonal est immédiat.

En relativité général on calcule également la divergence des tenseurs.

Théorème (et définition) VI.3.24 (Divergence d'un tenseur, cas d'un tenseur antisymétrique deux fois contravariant). Pour tout tenseur $T \in \mathcal{T} \binom{l}{k} \mathcal{M}$ où $l \geq 1$, on définit la divergence de T par,

$$\operatorname{div} T \hat{=} \nabla \cdot T = C_1^k \nabla T = \nabla_\alpha T^{i_1 \dots i_{k-1} \alpha}_{j_1 \dots j_l} \partial_{i_1} \otimes \dots \otimes \partial_{i_{k-1}} \otimes dx^\alpha \otimes dx^{j_1} \otimes \dots \otimes dx^{j_l}$$

Dans le cas où $T \in \mathcal{T} \binom{2}{0} \mathcal{M}$ à des composantes anti-symétriques $T^{\alpha\beta} = -T^{\beta\alpha}$ dans la base naturel, on a,

$$\operatorname{div} T = \frac{1}{\sqrt{\pm g}} \partial_\mu (\sqrt{\pm g} T^{\alpha\mu}) \partial_\alpha .$$

Démonstration. Soit encore,

$$\begin{aligned} \operatorname{div} T &= \nabla_\mu T^{\alpha\mu} \partial_\alpha \\ &= \left[\partial_\mu T^{\alpha\mu} + \overbrace{\Gamma_{\nu\mu}^\alpha T^{\nu\mu}}^{=0} + \Gamma_{\nu\mu}^\mu T^{\alpha\nu} \right] \partial_\alpha \\ &= [\partial_\mu T^{\alpha\mu} + T^{\alpha\mu} \partial_\mu \ln \sqrt{\pm g}] \partial_\alpha \end{aligned}$$

où l'annulation du terme au milieu de la seconde ligne vient de l'anti-symétrie de $T^{\nu\mu}$ et de la symétrie de $\Gamma_{\nu\mu}^\alpha$ sur les indices du bas.

■

Le tenseur ε de Levi-Civita est très utiles dans le traitement des l'électromagnétisme en relativité général mais aussi à propos du calculs intégrale dans les variétés, car il permet de calculer des volumes. Ce tenseur vérifie la propriété,

Théorème VI.3.25 (Dérivation covariante du tenseur de Levi-Civita). **Le Tenseur de Levi-Civita est parallèle vis-à-vis de la dérivation covariante**

Démonstration. Soit i_1, \dots, i_n une permutation de $1, \dots, n$. Calculons,

$$\begin{aligned}\nabla_\mu \varepsilon_{i_1, \dots, i_n} &= \partial_\mu \varepsilon_{i_1, \dots, i_n} - \sum_{q=1}^n \Gamma_{\mu i_q}^\sigma \varepsilon_{i_1, \dots, \underbrace{\sigma}_q, \dots, i_n} \\ &= \partial_\mu \sqrt{\pm g}[i_1, \dots, i_n] - \sum_{q=1}^n \Gamma_{\mu i_q}^\sigma \varepsilon_{i_1, \dots, \underbrace{\sigma}_q, \dots, i_n}\end{aligned}$$

Or puisque pour tout $\sigma \neq i_q$ on a $\varepsilon_{i_1, \dots, \sigma, \dots, i_n} = 0$,

$$\begin{aligned}\nabla_\mu \varepsilon_{i_1, \dots, i_n} &= (\partial_\mu \ln \sqrt{\pm g}) \varepsilon_{i_1, \dots, i_n} - \sum_{q=1}^n \Gamma_{\mu i_q}^{i_q} \varepsilon_{i_1, \dots, i_q, \dots, i_n} \\ &= \left[(\partial_\mu \ln \sqrt{\pm g}) - \sum_{q=1}^n \Gamma_{\mu i_q}^{i_q} \right] \varepsilon_{i_1, \dots, i_n}\end{aligned}$$

Or puisque i_1, \dots, i_n est une permutation de $1, \dots, n$ on a $\sum_{q=1}^n \Gamma_{\mu i_q}^{i_q} = \Gamma_{\mu\nu}^\nu = \partial_\mu \ln \sqrt{\pm g}$ et donc,

$$\nabla_\mu \varepsilon_{i_1, \dots, i_n} = [\partial_\mu \ln \sqrt{\pm g} - \partial_\mu \ln \sqrt{\pm g}] \varepsilon_{i_1, \dots, i_n} = 0$$

■

VI.3.2 Formes différentielles et dérivation extérieur

Dans la sous-section VI.2 on a introduit la notion de k -formes différentiel, il s'agit de tenseur alterné du type $\mathcal{T} \binom{0}{k} \mathcal{M}$. La dérivation covariante introduite en sous-sous-section VI.3.1 permet de passer d'un tenseur de type $\mathcal{T} \binom{l}{k} \mathcal{M}$ à une tenseur du type $\mathcal{T} \binom{l}{k+1} \mathcal{M}$. Ainsi sur une k -forme, il permet d'obtenir un tenseur de variance $\mathcal{T} \binom{0}{k+1} \mathcal{M}$. Afin d'obtenir une $k+1$ forme il suffit d'en calculer la partie entièrement anti-symétrisé, la succession des ces opérations : dérivation covariante et anti-symétrisation va définir la dérivation extérieur elle permet de passer d'une k -forme différentielle à une $k+1$ -forme.

Dérivation extérieur

Commençons par la définition,

Définition VI.3.26 (Derivée extérieur). Pour tout $p \in \mathbb{N}$, la dérivation extérieur est l'opérateur linéaire qui à tout champs de p -forme $\mathbf{A} \in \mathcal{A}_p \mathcal{M}$, associe une $p+1$ -forme $d\mathbf{A} \in \mathcal{A}_{p+1} \mathcal{M}$ valant $\forall (\mathbf{v}_1, \dots, \mathbf{v}_{p+1}) \mathcal{T} \mathcal{M}^{p+1}$,

$$d\mathbf{A}(\mathbf{v}_1, \dots, \mathbf{v}_{p+1}) = \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \nabla_{\mathbf{v}_{\sigma(1)}} \mathbf{A}(\mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(p+1)})$$

ce qui donne en terme de composante,

$$dA_{i_1 \dots i_{p+1}} = \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \nabla_{\sigma(i_1)} A_{\sigma(i_2), \dots, \sigma(i_{p+1})}$$

En fais la dérivation extérieur peut s'affranchir de la dérivation covariante....

Théorème VI.3.27. La dérivation extérieur peut également s'écrire,

$$dA_{i_1 \dots i_{p+1}} = \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \partial_{i_{\sigma(1)}} A_{i_{\sigma(2)}, \dots, i_{\sigma(p+1)}}$$

Démonstration. $\forall (\mathbf{v}_1, \dots, \mathbf{v}_{p+1}) \mathcal{T}\mathcal{M}^{p+1}$ et $\forall \sigma \in \mathcal{G}_n$ on pose,

$$c_\sigma = \nabla_{\mathbf{v}_{\sigma(1)}} \mathbf{A}(\mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(p+1)})$$

ce qui, d'après l'expression de la dérivée covariante d'un tenseur de ce type conduit à,

$$c_\sigma = v_{\sigma(1)}^\rho (\partial_\rho A_{\beta_1 \dots \beta_p}) v_{\sigma(2)}^{\beta_1} \dots v_{\sigma(p+1)}^{\beta_p} - \sum_{q=1}^p \Gamma_{\rho \beta_q}^\mu v_{\sigma(1)}^\rho A_{\beta_1 \dots \underbrace{\mu}_{q} \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots v_{\sigma(p+1)}^{\beta_p} = d_\sigma + \sum_{q=1}^p c_\sigma^q$$

On va travailler sur les termes c_σ^q , commençons par mettre en avant $v_\sigma(q+1)^{\beta_q}$,

$$\begin{aligned} c_\sigma^q &= \Gamma_{\rho \beta_q}^\mu v_{\sigma(1)}^\rho v_\sigma(q+1)^{\beta_q} A_{\beta_1 \dots \underbrace{\mu}_{q} \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots \overbrace{1}^q, \dots v_{\sigma(p+1)}^{\beta_p} \\ \sum_{\sigma \in \mathcal{G}_n} (-1)^{k(\sigma)} c_\sigma^q &= \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \Gamma_{\rho \beta_q}^\mu v_{\sigma(1)}^\rho v_{\sigma(q+1)}^{\beta_q} A_{\beta_1 \dots \mu \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots 1, \dots v_{\sigma(p+1)}^{\beta_p} \end{aligned}$$

ce qui donne l'idée de poser pour tout $q = 1 \dots p$,

$$\tau_q \in \mathcal{G}_{p+1} \text{ vérifiant } \begin{cases} q+1 & \text{si } k=1 \\ 1 & \text{si } k=q+1 \\ k & \text{sinon} \end{cases}$$

c'est la permutation des éléments $1 \leftrightarrow q+1$ et donc $(-1)^{k(\sigma \circ \tau_q)} = -(-1)^{k(\sigma)}$. L'application $\sigma \in \mathcal{G}_{p+1} \mapsto \sigma \circ \tau_q$ est un isomorphisme de $(\mathcal{G}_{p+1}, \circ)$ et donc,

$$\begin{aligned} \sum_{\sigma \in \mathcal{G}_n} (-1)^{k(\sigma)} c_\sigma^q &= \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \Gamma_{\rho \beta_q}^\mu v_{\sigma(1)}^\rho v_{\sigma(q+1)}^{\beta_q} A_{\beta_1 \dots \mu \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots \overbrace{1}^q, \dots v_{\sigma(p+1)}^{\beta_p} \\ &= \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma \circ \tau_q)} \Gamma_{\rho \beta_q}^\mu v_{\sigma \circ \tau_q(1)}^\rho v_\sigma \circ \tau_q(q+1)^{\beta_q} A_{\beta_1 \dots \mu \dots \beta_p} v_{\sigma \circ \tau_q(2)}^{\beta_1} \dots 1, \dots v_{\sigma \circ \tau_q(p+1)}^{\beta_p} \\ &= - \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \Gamma_{\rho \beta_q}^\mu v_{\sigma(q+1)}^\rho v_{\sigma(1)}^{\beta_q} A_{\beta_1 \dots \mu \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots 1, \dots v_{\sigma(p+1)}^{\beta_p} \end{aligned}$$

En renommant les indices muet $\rho \leftrightarrow \beta_q$, puis en utilisant la symétrie des Christoffels, on obtient,

$$\left. \begin{aligned} \sum_{\sigma \in \mathcal{G}_n} (-1)^{k(\sigma)} c_\sigma^q &= - \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \Gamma_{\beta_q \rho}^\mu v_{\sigma(q+1)}^{\beta_q} v_{\sigma(1)}^\rho A_{\beta_1 \dots \mu \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots 1, \dots v_{\sigma(p+1)}^{\beta_p} \\ &= - \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \Gamma_{\rho \beta_q}^\mu v_{\sigma(q+1)}^{\beta_q} v_{\sigma(1)}^\rho A_{\beta_1 \dots \mu \dots \beta_p} v_{\sigma(2)}^{\beta_1} \dots 1, \dots v_{\sigma(p+1)}^{\beta_p} \\ &= - \sum_{\sigma \in \mathcal{G}_n} (-1)^{k(\sigma)} c_\sigma^q \end{aligned} \right\} \Rightarrow \sum_{\sigma \in \mathcal{G}_n} (-1)^{k(\sigma)} c_\sigma^q = 0.$$

D'où

$$\mathbf{d}\mathbf{A}(\mathbf{v}_1, \dots, \mathbf{v}_{p+1}) = \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_n} (-1)^{k(\sigma)} v_{\sigma(1)}^\rho (\partial_\rho A_{\beta_1 \dots \beta_p}) v_{\sigma(2)}^{\beta_1} \dots v_{\sigma(p+1)}^{\beta_p}$$

ce qui donne en posant $\mathbf{v}_i = \partial_{j_i}$ pour tout $i = 1 \dots p+1$,

$$\mathbf{d}A_{i_1 \dots i_{p+1}} = \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \partial_{(i_{\sigma 1})} A_{i_{\sigma(2)}, \dots, i_{\sigma(p+1)}}$$

■

Remarque VI.3.28. Cette propriété est remarquable, en effet la dérivation extérieur s'affranchit des Christoffels qui exprime la variation des vecteurs de la base naturelle. Ainsi travailler dans une variété avec des coordonnées curviligne reviens, vis-à-vis de la dérivation extérieur, à travailler avec des espaces euclidiens munie de coordonnées cartésienne !

Exemple VI.3.29. Regardons ce que cela donne pour les composantes d'une dérivation extérieur d'une k -forme différentielle pour des petites valeurs de k . Quelques lignes de calculs conduit à,

- Pour les 0-forme (scalaire), $\mathbf{d}f_\alpha = \partial_\alpha f = \mathbf{d}dd_\alpha$
- Pour les 1-forme (scalaire), $\mathbf{d}\omega_{\alpha\beta} = \partial_\alpha \omega_\beta - \partial_\beta \omega_\alpha$
- Pour les 2-forme (scalaire), $\mathbf{d}\Omega_{\alpha,\beta,\gamma} = \partial_\alpha \Omega_{\beta\gamma} + \partial_\beta \Omega_{\gamma\alpha} + \partial_\gamma \Omega_{\alpha\beta}$

Propriété VI.3.30. Pour tout champs scalaire f et $\forall p \in \mathbb{N}, \forall (i_1, \dots, i_p) \in \llbracket 1, n \rrbracket^p$ on a,

$$\alpha = f \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \quad \Rightarrow \quad \mathbf{d}\alpha = \mathbf{d}f \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p}$$

Démonstration. En utilisant le théorème VI.3.27,

$$\begin{aligned} \mathbf{d}\alpha_{j_1 \dots j_{p+1}} &= \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \partial_{j_{\sigma(1)}} \alpha_{j_{\sigma(2)} \dots j_{\sigma(p+1)}} \\ &= \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \partial_{j_{\sigma(1)}} \left(f \delta_{j_{\sigma(2)}}^{i_1} \dots \delta_{j_{\sigma(p+1)}}^{i_p} \right) \\ &= \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} (\partial_{j_{\sigma(1)}} f) \delta_{j_{\sigma(2)}}^{i_1} \dots \delta_{j_{\sigma(p+1)}}^{i_p} \\ &= \frac{1}{p!} \sum_{\sigma \in \mathcal{G}_{p+1}} (-1)^{k(\sigma)} \mathbf{d}f_{j_{\sigma(1)}} \delta_{j_{\sigma(2)}}^{i_1} \dots \delta_{j_{\sigma(p+1)}}^{i_p} \\ &= \mathbf{d}f \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \end{aligned}$$

■

Propriété VI.3.31. Pour tout $p, q \in \mathbb{N}, \forall \mathbf{A}, \mathbf{B} \in \mathcal{A}_p \mathcal{M} \times \mathcal{A}_q \mathcal{M}$ on a,

$$\mathbf{d}(\mathbf{A} \wedge \mathbf{B}) = (\mathbf{d}\mathbf{A}) \wedge \mathbf{B} + (-1)^p \mathbf{A} \wedge (\mathbf{d}\mathbf{B})$$

Démonstration. On utilise la proposition VI.2.15 et donc, on a

$$\begin{aligned} \mathbf{A} &= \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} A_{i_1 \dots i_p} \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \\ \mathbf{B} &= \sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} B_{j_1 \dots j_q} \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \\ \Rightarrow \mathbf{A} \wedge \mathbf{B} &= \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} A_{i_1 \dots i_p} B_{j_1 \dots j_q} \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \wedge \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \end{aligned}$$

et donc en utilisant la proposition VI.3.30 et la linéarité de la différentiation extérieur on obtient,

$$\begin{aligned} \mathbf{d}\mathbf{A} \wedge \mathbf{B} &= \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} \mathbf{d}(A_{i_1 \dots i_p} B_{j_1 \dots j_q}) \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \wedge \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \\ &= \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} (B_{j_1 \dots j_q} \mathbf{d}A_{i_1 \dots i_p} + A_{i_1 \dots i_p} \mathbf{d}B_{j_1 \dots j_q}) \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \wedge \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \\ &= \left[\sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \mathbf{d}(A_{i_1 \dots i_p}) \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \right] \wedge \left[\sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} B_{j_1 \dots j_q} \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \right] \\ &\quad + \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} A_{i_1 \dots i_p} \mathbf{d}B_{j_1 \dots j_q} \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \wedge \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \\ &= \left[\sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \mathbf{d}(A_{i_1 \dots i_p}) \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \right] \wedge \left[\sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} B_{j_1 \dots j_q} \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \right] \\ &\quad + (-1)^p \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \sum_{1 \leq j_1 \leq \dots \leq j_q \leq n} A_{i_1 \dots i_p} \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_p} \wedge \mathbf{d}B_{j_1 \dots j_q} \wedge \mathbf{d}x^{j_1} \wedge \dots \wedge \mathbf{d}x^{j_q} \\ &= (\mathbf{d}\mathbf{A}) \wedge \mathbf{B} + (-1)^p \mathbf{A} \wedge (\mathbf{d}\mathbf{B}). \end{aligned}$$

où le passage à l'avant dernière lignes est assuré par l'anti-symétrie du produit extérieur. ■

Rotationnel

Une quantité souvent rencontré en dimension 3 est le rotationnel d'un champs. Cette quantités est construite à l'aide de toute les constructions effectué jusqu'à présent.

Définition VI.3.32 (Rotationnel). *Pour toute variété Riemannienne (\mathcal{M}, g) de dimension 3, pour tout champs de vecteur $\vec{v} \in \mathcal{T}\mathcal{M}$ lisse, on appelle rotationnel de \vec{v} le dual métrique de $\star d\vec{v}$. On le note $\text{rot } \vec{v} \in \mathcal{T}\mathcal{M}$.*

Propriété VI.3.33. Pour toute variété Riemannienne (\mathcal{M}, g) de dimension 3, pour tout champs de vecteurs $\vec{v} \in \mathcal{T}\mathcal{M}$ lisse, on a,

$$\text{rot } \vec{v} = \varepsilon_{\lambda}^{\mu\gamma} \nabla_{\gamma}(v^{\lambda}) \partial_{\mu}$$

Ce qui dans un système de coordonnée orthogonal donne.

$$\text{rot } \vec{v} = \sum_{\sigma, \tau, \gamma} \frac{h_{\gamma}[\sigma\tau\gamma]}{h_1 h_2 h_3} \partial_{\sigma} (h_{\tau} v^{\hat{\tau}}) e_{\gamma}$$

Démonstration. On obtient en utilisant le théorème VI.3.27

$$\begin{aligned} (\bar{d}v)_{\alpha\beta} &= \partial_{\alpha}v_{\beta} - \partial_{\beta}v^{\alpha} = \nabla_{\alpha}v_{\beta} - \nabla_{\beta}v^{\alpha} \\ \star(\bar{d}v)_{\mu} &= \frac{1}{2} \varepsilon_{\mu_1\mu_2\mu} g^{\mu_1\nu_1} g^{\mu_2\nu_2} (\bar{d}v)_{\nu_1\nu_2} \\ &= \varepsilon_{\mu_1\mu_2\mu} g^{\mu_1\nu_1} g^{\mu_2\nu_2} \nabla_{\nu_1}v_{\nu_2} \\ &= \varepsilon_{\mu}^{\gamma\sigma} \nabla_{\gamma}v_{\sigma} \\ \text{rot } \vec{v} &= g^{\mu\lambda} \varepsilon_{\mu}^{\gamma\sigma} \nabla_{\gamma}v_{\sigma} \partial_{\lambda} \\ &= \varepsilon^{\gamma\sigma\mu} \nabla_{\gamma}v_{\sigma} \partial_{\mu} \\ &= \varepsilon^{\gamma\sigma\mu} \nabla_{\gamma}(g_{\sigma\lambda}v^{\lambda}) \partial_{\mu} \\ &= \varepsilon^{\mu\gamma\sigma} g_{\sigma\lambda} \nabla_{\gamma}(v^{\lambda}) \partial_{\mu} \text{ métrique parrallèle} \\ &= \varepsilon_{\lambda}^{\mu\gamma} \nabla_{\gamma}(v^{\lambda}) \partial_{\mu} \end{aligned}$$

Dans un système de coordonnées orthogonale, reprenons de $\text{rot } \vec{v} = \varepsilon^{\gamma\sigma\mu} \nabla_{\gamma}v_{\sigma} \partial_{\mu}$ et remplaçons la dérivée covariante par la dérivée partielle (justifié par le théorème VI.3.27) alors,

$$\begin{aligned} \text{rot } \vec{v} &= \varepsilon^{\gamma\sigma\mu} \partial_{\gamma}v_{\sigma} \partial_{\mu} \\ &= \sum_{\sigma, \tau, \gamma} \frac{[\gamma\sigma\mu]}{h_1 h_2 h_3} \partial_{\gamma}v_{\sigma} h_{\mu} e_{\mu} \\ &= \sum_{\sigma, \tau, \gamma} \frac{h_{\gamma}[\sigma\tau\gamma]}{h_1 h_2 h_3} \partial_{\sigma} (h_{\tau} v^{\hat{\tau}}) e_{\gamma}. \end{aligned}$$

En effet, on a $v_{\sigma} = h_{\sigma}^2 v^{\sigma} = h_{\sigma} v^{\hat{\sigma}}$ ■

Exercice 41. — Rotationnel en coordonnée parabolique.

On rappelle que dans $\mathbb{R}^3 - O_y$ les coordonnées paraboliques sont définie par

$$\begin{aligned} \psi &: \mathbb{R} \times \mathbb{R}_{+,*} \times \mathbb{R} \rightarrow \mathbb{R}^3 - O_y \\ (u, v, z) &\mapsto \left(uv, \frac{u^2 - v^2}{2}, z \right) \end{aligned}$$

Montrer alors que le rotationnel vaut,

$$\begin{aligned} \text{rot } \vec{V} &= \frac{1}{\sqrt{u^2 + v^2}} \left(\partial_v V^{\hat{z}} - \sqrt{u^2 + v^2} \partial_z V^{\hat{v}} \right) e_u \\ &+ \frac{1}{\sqrt{u^2 + v^2}} \left(\sqrt{u^2 + v^2} \partial_z V^{\hat{u}} - \partial_u V^{\hat{z}} \right) e_v \\ &+ \frac{1}{u^2 + v^2} \left(\partial_u (\sqrt{u^2 + v^2} V^{\hat{v}}) - \partial_v (\sqrt{u^2 + v^2} V^{\hat{u}}) \right) e_v \end{aligned}$$

Lemme de Poincaré

Pour tout fonction scalaire, il est facile de montrer que,

$$\mathbf{d}^2 f = \mathbf{d}\mathbf{d}f = (\partial_\alpha \mathbf{d}f_\beta - \partial_\beta \mathbf{d}f_\alpha) \mathbf{d}x^\alpha \otimes \mathbf{d}x^\beta = (\partial_\alpha \partial_\beta f - \partial_\beta \partial_\alpha f) \mathbf{d}x^\alpha \otimes \mathbf{d}x^\beta = \mathbf{0}$$

cette propriété découlant du théorème de Schwarz s'étendra facilement p -formes différentielles. La question posé par le Lemme de Poincaré est la question de l'existence d'une "primitive" c'est-à-dire de savoir si $\mathbf{d}\Omega = \mathbf{0}$ implique l'existence de ω tel que $\Omega = \mathbf{d}\omega$.

Definition des formes exact et fermé Commençons par un peu de vocabulaire.

Définition VI.3.34 (Forme différentielle exacte et fermé). *Soit pour tout $p \geq 1$, une p -forme différentielle $\Omega \in \mathcal{A}_p \mathcal{M}$ alors on dis que Ω :*

- est fermé si et seulement si $\mathbf{d}\Omega = \mathbf{0}$
- est exacte si et seulement si il existe $\omega \in \mathcal{A}_{p-1} \mathcal{M}$ tel que $\Omega = \mathbf{d}\omega$, on dira alors que ω est une primitive

Propriété VI.3.35. Toute p -forme exact est fermé

Démonstration. Pour toute p -forme Ω exacte il existe ω tel que $\omega \in \mathcal{A}_{p-1} \mathcal{M}$ alors on note,

$$\omega = \sum_{1 \leq i_1 \leq \dots \leq i_{p-1} \leq n} \omega_{i_1, \dots, i_{p-1}} \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_{p-1}}$$

Ce qui implique,

$$\begin{aligned} \Omega &= \mathbf{d}\omega = \sum_{1 \leq i_1 \leq \dots \leq i_{p-1} \leq n} \mathbf{d}\omega_{i_1, \dots, i_{p-1}} \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_{p-1}} \\ &= \mathbf{d}\omega = \sum_{1 \leq i_1 \leq \dots \leq i_{p-1} \leq n} \partial_i \omega_{i_1, \dots, i_{p-1}} \mathbf{d}x^i \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_{p-1}} \end{aligned}$$

D'où,

$$\begin{aligned} \mathbf{d}\Omega &= \mathbf{d}^2 \omega = \sum_{1 \leq i_1 \leq \dots \leq i_{p-1} \leq n} \mathbf{d}\partial_i \omega_{i_1, \dots, i_{p-1}} \wedge \mathbf{d}x^i \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_{p-1}} \\ &= \sum_{1 \leq i_1 \leq \dots \leq i_{p-1} \leq n} \partial_{ij}^2 \omega_{i_1, \dots, i_{p-1}} \mathbf{d}x^j \wedge \mathbf{d}x^i \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_{p-1}} \\ &= \sum_{1 \leq i_1 \leq \dots \leq i_{p-1} \leq n} \underbrace{(\partial_{ij}^2 \omega_{i_1, \dots, i_{p-1}} \mathbf{d}x^j \wedge \mathbf{d}x^i)}_{=0 \text{ Schwarz}} \wedge \mathbf{d}x^{i_1} \wedge \dots \wedge \mathbf{d}x^{i_{p-1}} \end{aligned}$$

■

Circulation des 1-formes Les 1-formes sont susceptible de permettre un calcul intégral sur une courbe paramétrée incluse dans la variété qui ne sera caractéristique, à un signe près qu'à la "courbe géométrique" associé.

Définition VI.3.36 (Circulation d'une 1-forme). *Pour toute 1-forme, $\omega \in \mathcal{A}_1 \mathcal{M}$ et pour toute courbe paramétrée fermé $\gamma : [a, b] \rightarrow \mathcal{M}$ ($a \leq b$), la circulation de ω le long de γ est :*

$$\int_{\gamma} \omega \hat{=} \int_a^b \omega(\gamma(t)) \left(\frac{d\gamma}{dt} \right) dt$$

Propriété VI.3.37. La circulation d'une 1-forme différentielle ω le long d'une courbe paramétrée fermé $\gamma : [a, b] \rightarrow \mathcal{M}$ ($a \leq b$) est invariante par changement de paramétrage admissible conservant l'orientation. Elle est multiplié par -1 si on utilise un changement de paramétrage admissible changeant l'orientation.

Démonstration. Un changement de paramètre admissible est tout simplement la composition de γ avec une fonction lisse $\theta : [\alpha, \beta] \rightarrow [a, b]$ réel strictement monotone, et dont la dérivée ne s'annule pas, alors $\Gamma = \gamma \circ \theta$,

$$\begin{aligned}\int_{\Gamma} \omega &= \int_{\alpha}^{\beta} \omega(\gamma \circ \theta(t)) \left(\frac{d\gamma \circ \theta}{dt} \right) dt \\ &= \int_{\alpha}^{\beta} \omega(\gamma \circ \theta(t)) \left(\frac{d\gamma}{dt}(\theta(t)) \right) \dot{\theta}(t) dt \\ &= \int_{\theta(\alpha)}^{\theta(\beta)} \omega(\gamma(u)) \left(\frac{d\gamma}{dt}(u) \right) du\end{aligned}$$

Si θ est croissante alors l'orientation est préservé et $\theta(\alpha) = a$ et $\theta(\beta) = b$. Et on obtient le résultat. Si elle est décroissante alors $\theta(\alpha) = b$ et $\theta(\beta) = a$ et on obtient le résultat multiplié par -1 ■

Remarque VI.3.38. La circulation d'une forme différentielle sur un chemin orienté est donc un nombre qui ne dépend que du chemin orienté choisi.

Propriété VI.3.39. Soit une 1-forme différentielle ω exacte et f une primitive alors la circulation de ω le long de $\gamma : [a, b] \rightarrow \mathcal{M}$ vaut,

$$\int_{\omega} \omega = [f \circ \gamma]_a^b = f(\gamma(b)) - f(\gamma(a))$$

Ainsi pour une courbe paramétrée fermé $\gamma(a) = \gamma(b)$ on a,

$$\oint_{\omega} \omega = 0$$

Démonstration. Dérivons $t \mapsto f \circ \gamma(t)$,

$$\frac{df \circ \gamma}{dt} = df(\gamma)(\dot{\gamma}(t)) = df(\gamma)(\dot{\gamma}(t)) = \omega(\gamma)(\dot{\gamma}(t))$$

Ce qui justifie les résultats. ■

Corrolaire VI.3.40. Soit une 1-forme différentielle ω exacte et f et g deux primitives sur une variété connexe par arc. Alors $f - g$ est constante sur \mathcal{M} .

Démonstration. Posons $h = f - g$ puisque la variété est connexe par arc $\forall M, N \in \mathcal{M}$ il existe une courbe paramétrée $\gamma : [a, b] \rightarrow \mathcal{M}$ lisse tel que $\gamma(a) = N$ et $\gamma(b) = M$. Sachant que $\mathrm{dh} = df - dg = \omega - \omega = 0$. On a,

$$0 = \int_{\omega} \mathrm{dh} = h(M) - h(N)$$

ce qui justifie le résultat. ■

Lemme de Poincaré

Définition VI.3.41 (Ouvert étoilé). *Une ouvert \mathcal{U} de \mathbb{R}^n est étoilé si et seulement si $\exists M_0$ tel que $\forall M \in \mathcal{U}$ le segment $[M_0 M] \subset \mathcal{U}$. On dit alors que \mathcal{U} est étoilé par rapport à M_0 .*

Théorème VI.3.42 (Lemme de Poincaré). *Soit (\mathcal{M}, g) une variété Riemannienne de dimension n décrite par un atlas composé d'une seule carte, c'est-à-dire qu'il existe $\mathcal{U} \subset \mathbb{R}^n$ et un difféomorphisme $\psi : \mathcal{U} \rightarrow \mathcal{M}$. On supposera également que \mathcal{U} soit étoilé. Alors toute $p+1$ -forme différentielle fermé de \mathcal{M} est exact.*

Démonstration. On ne démontrera ce théorème uniquement pour les 1-forme différentielle. Prenons donc ω une 1-forme différentielle exacte. On confond comme d'ordinaire $M \in \mathcal{M} \mapsto \omega(M)(.)$ et $m \in \mathcal{U} \mapsto \omega(\psi(m))(.)$. On note alors,

$$\omega(m \in \mathcal{U})(.) = P_{\alpha}(m) \mathrm{d}x^{\alpha}(.)$$

Puisque \mathcal{U} est étoilé, on peut (puisque on peut composer ϕ au départ avec une translation) sans perte de généralité supposer que \mathcal{U} est étoilé par rapport à $o = 0_{\mathbb{R}^n}$. On pose alors, pour tout $m = (x^1, \dots, x^n) \in \mathcal{U}$,

$$f(m) = f(o) + \int_0^1 x^\alpha P_\alpha(o(1-t) + tm) dt$$

$$f(x^1, \dots, x^n) = f(o) + \int_0^1 x^\alpha P_\alpha(tx^1, \dots, tx^n) dt$$

Montrons que $\mathbf{d}f = \omega$. La fonction f est une intégrale à paramètre. On pose $F : (t, m) \in [0, 1] \times \mathcal{U} \mapsto x^\alpha P_\alpha(tm) \in \mathbb{R}$. Pour tout $m_0 \in \mathcal{U}$, on posera \mathcal{C}_0 un compacte inclus dans \mathcal{U} et contenant m et on considérera la fonction F restreinte à $[0, 1] \times \mathcal{C}_m$, alors $\forall i \in \llbracket 1, n \rrbracket$

1. Pour tout $(t, m) \in [0, 1] \times \mathcal{C}_0$ $\partial_i F(t, m) = P_i(tm) + tx^\alpha \partial_i P_\alpha(tm)$ existe puisque les P_α sont lisses.
2. Pour tout $m \in \mathcal{U}$, la fonction $t \in [0, 1] \mapsto \partial_i F(t, m)$ est lisse puisque les P_α le sont.
3. Puisque elle est continue définit sur le compacte $[0, 1] \times \mathcal{C}_0$ la fonction $t \in [0, 1] \mapsto |\partial_i F|(t, m)$ est majoré.

On entre donc dans les hypothèses de la dérivation sous l'intégrale et donc f est différentiable. Ainsi, pour tout k ,

$$\partial_k f(m) = \int_0^1 P_k(tm) + tx^\alpha \partial_k P_\alpha(tm) dt.$$

Puisque ω est exact, pour tout $k, l = 1 \dots n$ on a $\partial_k P_l = \partial_l P_k$ et donc,

$$\partial_k f(m) = \int_0^1 P_k(tm) + tx^\alpha \partial_\alpha P_k(tm) dt = \int_0^1 \frac{d}{dt} [tP_k(tm)] dt = [tP_k(tm)]_0^1 = P_k(tm)$$

Ce qui reviens à dire que $\omega = P_\alpha \mathbf{d}x^\alpha = \partial_\alpha f \mathbf{d}x^\alpha = df = \mathbf{d}f$ ■

Remarque VI.3.43. Ce théorème peut être étendu à des variétés contractiles (voir : wiki). En fait l'important est qu'il n'y ai pas de "trou".

Dans les variété de dimension 3 On se place dans une variété de dimension 3.

Propriété VI.3.44 (Tout champs irrotationnel dérive d'un gradient). On dis d'un champ $\mathbf{v} \mathcal{T} \mathcal{M}$ qu'il est irrotationnel si et seulement si $\text{rot } \vec{v} = \mathbf{0}$, alors si \mathcal{M} est une variété répondant au critère du Lemme de Poincaré, \mathbf{v} dérive d'un gradient, c'est-à-dire qu'il existe un scalaire f lisse tel que $\mathbf{v} = \nabla f$

Démonstration. On raisonne par équivalence dans une variété répondant au critère du Lemme de Poincaré,

$$\begin{aligned} \text{rot } \vec{v} = \mathbf{0} &\Leftrightarrow \star \mathbf{d}\underline{v} = 0 \\ &\Leftrightarrow \mathbf{d}\underline{v} = 0 \quad \text{Poincaré} \\ &\Leftrightarrow \exists f \text{ scalaire } \underline{v} = \mathbf{d}f = df \\ &\Leftrightarrow \exists f \text{ scalaire } \vec{v} = \nabla f \end{aligned}$$

Propriété VI.3.45 (Tout champ incompressible dérive d'un rotationnel). On dis d'un champ $\mathbf{B} \mathcal{T} \mathcal{M}$ qu'il est incompressible si et seulement si $\text{div } \vec{B} = \mathbf{0}$, alors si \mathcal{M} est une variété répondant au critère du Lemme de Poincaré, \mathbf{v} dérive d'un rotationnel, c'est-à-dire qu'il existe un champ de vecteur \mathbf{A} lisse tel que $\mathbf{B} = \text{rot } \vec{A}$.

Démonstration. Soit Ω une 2-forme, alors $\mathbf{d}\Omega$ est une 3-forme et,

$$\mathbf{d}\Omega_{\alpha, \beta, \gamma} = \nabla_\alpha \Omega_{\beta\gamma} + \nabla_\beta \Omega_{\gamma\alpha} + \nabla_\gamma \Omega_{\alpha\beta}$$

et donc $\star \mathbf{d}\Omega$ est un scalaire valant,

$$\begin{aligned} \star \mathbf{d}\Omega &= \frac{1}{3!} \varepsilon^{\alpha\beta\gamma} [\nabla_\alpha \Omega_{\beta\gamma} + \nabla_\beta \Omega_{\gamma\alpha} + \nabla_\gamma \Omega_{\alpha\beta}] \\ &= \frac{1}{3} \left[\nabla_\alpha \left(\frac{1}{2} \varepsilon^{\alpha\beta\gamma} \Omega_{\beta\gamma} \right) + \nabla_\beta \left(\frac{1}{2} \varepsilon^{\alpha\beta\gamma} \Omega_{\gamma\alpha} \right) + \nabla_\gamma \left(\frac{1}{2} \varepsilon^{\alpha\beta\gamma} \Omega_{\alpha\beta} \right) \right] \\ &= \frac{1}{3} [\nabla_\alpha \star \vec{\Omega}^\alpha + \nabla_\beta \star \vec{\Omega}^\beta + \nabla_\gamma \star \vec{\Omega}^\gamma] \\ &= \operatorname{div} \vec{\Omega} \end{aligned}$$

où le passage à la deuxième ligne est permis par le fait que $\nabla \varepsilon = \mathbf{0}$. Ainsi on a,

$$\boxed{\mathbf{d}\Omega = \star \star \mathbf{d}\Omega = \star \operatorname{div} \vec{\Omega} = \operatorname{div} \vec{\Omega} \varepsilon}. \quad (\text{VI.158})$$

Ainsi pour tout vecteur \mathbf{B} , on pose alors $\Omega = \star \underline{\mathbf{B}}$ qui est une deux forme. Ce qui nous autorise à utiliser la relation précédente,

$$\mathbf{d} \star \underline{\mathbf{B}} = \operatorname{div} \star \star \vec{\mathbf{B}} b m \varepsilon = \operatorname{div} \vec{\mathbf{B}} \varepsilon$$

Ce qui implique que $\mathbf{d} \star \underline{\mathbf{B}} = \mathbf{0}$ puisque \mathbf{B} est incompressible. Le lemme de Poincaré s'applique et puisque $\star \underline{\mathbf{B}}$ est une deux forme, il existe une une forme $\underline{\mathbf{A}}$ tel que $\star \underline{\mathbf{B}} = \mathbf{d} \underline{\mathbf{A}}$ et donc $\underline{\mathbf{B}} = \star \mathbf{d} \underline{\mathbf{A}}$. Ce qui en passant au dual métrique implique $\mathbf{B} = \operatorname{rot} \vec{\mathbf{A}}$. ■

Exercice en rab

Exercice 42. — De l'importance de l'étoilé.

Montrer que

$$\omega = \frac{x \mathbf{d}y - y \mathbf{d}x}{x^2 + y^2}$$

définie sur $\mathbb{R}^2 - \{\mathbf{0}_{\mathbb{R}^2}\}$ est fermé mais qu'elle n'est pas exacte. (Indication la circulation sur le cercle unité)

Exercice 43. — Condition fermeture.

Chercher l'ensemble des fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ tel que la forme,

$$\omega = \left(\frac{x^2 + y^2 - y}{xy} \mathbf{d}x + \frac{\mathbf{d}y}{y} \right) f \left(\frac{y}{x} \right)$$

soit exact.

VI.4 Intégration

L'on vient de voir comment on pouvait intégrer une 1-forme différentielle sur une courbe paramétrée plongé dans la variété sur lequel étais défini la forme différentielle. Cette construction à un sens car elle est indépendante du paramétrage orienté de la courbe paramétrée. Le but de cette partie sera de généraliser ce genre de construction. Afin de se simplifier la vie, on travaillera dans une variété de dimension n pour lequel il existe un atlas qui soit composé que d'une seule carte $\psi : \mathcal{U} \rightarrow \mathcal{M}$ et $\phi = \psi^{-1} : \mathcal{M} \rightarrow \mathcal{U}$. La plupart des théorèmes de cette section sont assez technique et ne serons pas démontré. Ce qui est important est que l'élève puisse savoir quoi faire lorsqu'il aura à calculer des intégrales dans une variétés.

VI.4.1 Intégration sur un volume

Dans la sous-section VI.2.3 on vu que l'action de la forme différentielle de Levi-Civita sur n vecteur correspondait au volume orienté du parallélogramme construit sur ces vecteurs. Ainsi le tenseur de Levi-Civita et les construction associées jouerons un grand rôle dans le calcul d'intégrale.

Théorème (et définition) VI.4.1 (Maillage volumique Volumes élémentaire). Pour une variété \mathcal{M} paramétrée par $\psi : \mathcal{U} \rightarrow \mathcal{M}$, nous avons une base naturel $\partial_i = \partial_i \psi$ alors pour tout $i \in [1, n]$ on pose $\ell_i = dx^i \partial_i$ où il n'y a pas de sommation et $dx^i \in \mathbb{R}$ est un accroissement élémentaire. La famille $(\ell_i)_{i=1 \dots n}$ est alors appelé *maillage* de \mathcal{M} .

On définira également l'intégrande,

$$d^n V = \varepsilon(\ell_1, \dots, \ell_n) = \sqrt{\pm g} dx^1 \dots dx^n$$

Soit un compact non vide $\mathcal{V} \subset \mathcal{M}$ (c'est-à-dire que $\mathcal{V} = \phi(\mathcal{V})$ est un compact). Pour tout scalaire, $f : \mathcal{M} \rightarrow \mathbb{R}$ et pour une n -forme Ω , ($\Omega = \star \Omega \varepsilon$), on définit alors :

$$\left\{ \begin{array}{l} \int_{\mathcal{V}} f \hat{=} \int_{\mathcal{V}} f \circ \psi(x^1, \dots, x^n) \varepsilon(\ell_1, \dots, \ell_n) \\ \hat{=} \int_{\mathcal{V}} f \circ \psi(x^1, \dots, x^n) \varepsilon(\partial_1, \dots, \partial_n) dx^1 \dots dx^n \\ = \int_{\mathcal{V}} f \circ \psi(x^1, \dots, x^n) \sqrt{\pm g} dx^1 \dots dx^n \\ \\ \int_{\mathcal{V}} \Omega \hat{=} \int_{\mathcal{V}} \star \Omega \\ = \int_{\mathcal{V}} \star \Omega(x^1, \dots, x^n) \varepsilon(\partial_1, \dots, \partial_n) dx^1 \dots dx^n \\ = \int_{\mathcal{V}} \Omega(\partial_1, \dots, \partial_n) \sqrt{\pm g} dx^1 \dots dx^n \end{array} \right.$$

L'intégrale sur \mathcal{V} de f ou de Ω . Cette quantités sont invariantes par changement de paramétrage conservant l'orientation.

Démonstration. Un changement de paramétrage admissible est un changement de coordonnée. C'est-à-dire la composition de ψ avec un difféomorphisme $\theta : \tilde{\mathcal{U}} \rightarrow \mathcal{U}$ dont la Jacobienne est partout strictement positive sur $\tilde{\mathcal{U}}$. On notera \tilde{x}^α les nouvelles coordonnées, on a donc $\theta(\tilde{x}^1, \dots, \tilde{x}^n) = (x^1, \dots, x^n)$, dans ces coordonnées la base naturelle devient,

$$, \tilde{\partial}_i = \frac{\partial x^j}{\partial \tilde{x}^i} d\tilde{x}^i .$$

Posons, $\tilde{\mathcal{V}} = \theta^{-1} \circ \phi(\mathcal{V})$ et donc $\theta(\tilde{\mathcal{V}}) = \mathcal{V}$.

Donc dans le nouveau paramétrage l'intégrale est égale à

$$\begin{aligned} \int_{\mathcal{V}} f &= \int_{\tilde{\mathcal{V}}} f \circ \psi \circ \theta(\tilde{x}^1, \dots, \tilde{x}^n) \varepsilon(\tilde{\partial}_1, \dots, \tilde{\partial}_n) d\tilde{x}^1 \dots d\tilde{x}^n \\ &= \int_{\theta^{-1}(\phi(\mathcal{V}))} f \circ \psi(\theta(\tilde{x}^1, \dots, \tilde{x}^n)) \varepsilon\left(\frac{\partial x^{j_1}}{\partial \tilde{x}^1} \partial_{j_1}, \dots, \frac{\partial x^{j_n}}{\partial \tilde{x}^n} \partial_{j_n}\right) d\tilde{x}^1 \dots d\tilde{x}^n \\ &= \int_{\theta^{-1}(\phi(\mathcal{V}))} f \circ \psi(\theta(\tilde{x}^1, \dots, \tilde{x}^n)) \sqrt{\pm g(\theta(\tilde{x}^1, \dots, \tilde{x}^n))} [j_1, \dots, j_n] \frac{\partial x^{j_1}}{\partial \tilde{x}^1} \dots \frac{\partial x^{j_n}}{\partial \tilde{x}^n} d\tilde{x}^1 \dots d\tilde{x}^n \\ &= \int_{\theta^{-1}(\phi(\mathcal{V}))} f \circ \psi(\theta(\tilde{x}^1, \dots, \tilde{x}^n)) \sqrt{\pm g(\theta(\tilde{x}^1, \dots, \tilde{x}^n))} \det d\theta(\tilde{x}^1, \dots, \tilde{x}^n) d\tilde{x}^1 \dots d\tilde{x}^n \\ &= \int_{\theta^{-1}(\phi(\mathcal{V}))} f \circ \psi(\theta(\tilde{x}^1, \dots, \tilde{x}^n)) \sqrt{\pm g(\theta(\tilde{x}^1, \dots, \tilde{x}^n))} |\det d\theta(\tilde{x}^1, \dots, \tilde{x}^n)| d\tilde{x}^1 \dots d\tilde{x}^n \end{aligned}$$

où le passage à la dernière ligne est permis par la conservation de l'orientation (: $\det d\theta(\tilde{x}^1, \dots, \tilde{x}^n)$

est partout positive). On peut donc faire un changement de variable et obtenir :

$$\begin{aligned}
 \int_{\mathcal{V}} f &= \int_{\boldsymbol{\theta}^{-1}(\phi(\mathcal{V}))} f \circ \psi(\boldsymbol{\theta}(\tilde{x}^1, \dots, \tilde{x}^n)) \sqrt{\pm g(\boldsymbol{\theta}(\tilde{x}^1, \dots, \tilde{x}^n))} |\det d\boldsymbol{\theta}(\tilde{x}^1, \dots, \tilde{x}^n)| d\tilde{x}^1 \dots d\tilde{x}^n \\
 &= \int_{\boldsymbol{\theta}(\boldsymbol{\theta}^{-1}(\phi(\mathcal{V})))} f \circ \psi(x^1, \dots, x^n) \sqrt{\pm g(\boldsymbol{\theta}(x^1, \dots, x^n))} dx^1 \dots dx^n \\
 &= \int_{\phi(\mathcal{V})} f \circ \psi(x^1, \dots, x^n) \sqrt{\pm g(\boldsymbol{\theta}(x^1, \dots, x^n))} dx^1 \dots dx^n \\
 &= \int_{\mathcal{V}} f \circ \psi(x^1, \dots, x^n) \sqrt{\pm g(\boldsymbol{\theta}(x^1, \dots, x^n))} dx^1 \dots dx^n
 \end{aligned}$$

ce qui donne le résultat attendu. ■

Remarque VI.4.2. Le volume de la partie \mathcal{V} est tout simplement $\int_{\mathcal{V}} 1$.

VI.4.2 Intégration sur une sous variété à bords de dimension p

L'intégration peut également se définir sur des sous-ensemble de la variété qui sont de dimension inférieur : des sous-variétés. Dans ce qui suivra, on notera $\mathcal{M}_n = \mathcal{M}$.

Sous variété à bords et orientation du bords

Pour tout p , on pose,

$$\begin{aligned}
 \mathcal{B}_1^p &= \left\{ (x^1, \dots, x^p) \in \mathbb{R}^p \mid \sum x^{k^2} \leq 1 \right\} \\
 \mathcal{B}_{1/2}^p &= \left\{ (x^1, \dots, x^p) \in \mathbb{R}^p \mid \sum x^{k^2} \leq 1 \text{ et } x^p \geq 0 \right\} \partial \mathcal{B}_{1/2}^p = \left\{ (x^1, \dots, x^p) \in \mathbb{R}^p \mid \sum x^{k^2} \leq 1 \text{ et } x^p = 0 \right\}
 \end{aligned}$$

Commençons par définir une sous-variété à bords,

Définition VI.4.3 (Sous-variété à bords). *Une partie $\mathcal{M}_p \subset \mathcal{M}_n$ est une sous-variété de \mathcal{M} si et seulement si $\tilde{\mathcal{M}}_p = \phi(\mathcal{M}_p) \subset \mathcal{U}$ est une variété à bords de dimension p plongé dans $\mathcal{U} \subset \mathbb{R}^n$, c'est-à-dire $\exists (\mathcal{V}_i^I)_{i \in I}, \exists (\mathcal{V}_k^K)_{k \in K}, \exists \varphi_i^I : \mathcal{B}_1^p \rightarrow \mathcal{V}_i^I$ pour tout $i \in I$ et enfin $\exists \varphi_k^K : \mathcal{B}_{1/2}^p \rightarrow \mathcal{V}_k^K$ tel que*

VI.4.3.1. I et K sont dénombrable

VI.4.3.2. Les $(\bigcup_{i \in I} \mathcal{V}_i^I) \cup (\bigcup_{k \in K} \mathcal{V}_k^K) = \tilde{\mathcal{M}}_p$

VI.4.3.3. $\forall (i, k) \in I \times K \varphi_i^I$ et φ_k^K sont des difféomorphismes immersif

VI.4.3.4. Les applications de recollement que l'ont peut construire sont également des difféomorphismes

Alors pour tout points $P \in \mathcal{M}_p$ on dis que P appartient au bords $\partial \mathcal{M}_p$ si $\exists k \in K$ tel que $\phi(P) \in \mathcal{V}_k^K$ et $(\varphi_k^K)^{-1}(\phi(P)) \in \partial \mathcal{B}_{1/2}^p$ et à l'intérieur $\text{Int } \mathcal{M}_p$ sinon.

Remarque VI.4.4. On pourrait, comme dans la section V.4, montrer qu'il y a partition du bord et de l'intérieur d'une sous variété à bord. C'est-à-dire que pour tout point $P \in \partial \mathcal{M}_p$ appartenant au bord, quelque soit $i \in I$, $\phi(P)$ n'appartient à aucun ouvert du type \mathcal{V}_i^I et pour tout $k \in K$ $\phi(P)$ ne peut appartenir à l'intérieur de \mathcal{V}_k^K . Ce genre de démonstration, basé sur la conservation de la structure topologique via des difféomorphisme astucieusement construit, est assez technique : nous l'admettrons.

Nous admettrons également que le bord de la variété $\partial \mathcal{M}_p$ est lui même une sous variété de dimension $p - 1$.

Dans la suite on travaillera (sans perte de généralités) avec une sous-variété compact à bords \mathcal{M}_p de \mathcal{M}_n dont l'atlas n'est constitué que d'une seule carte et dont le bords n'est pas vide. Comme le bords n'est pas vide, alors K est constitué d'un élément et $I =$ on notera $\psi \circ \varphi : \mathcal{B}_{1/2}^p \rightarrow \mathcal{M}_p$.

Définition VI.4.5 (Décomposition de l'espace tangent). *On considère une sous-variété \mathcal{M}_p compact à bords non vide dont l'atlas n'est constitué que d'une seule carte $\psi \circ \varphi : \mathcal{B}_{1/2}^p \rightarrow \mathcal{M}_p$. Alors pour tout point $M = \psi \circ \varphi(m) \in \mathcal{M}_p$, l'application $\psi \circ \varphi$ induit un espace tangent $\mathcal{T}_M \mathcal{M}_p = \text{im } d\psi \circ \varphi(m) \subset \mathcal{T}_m \mathcal{M}_n$, qui est de dimension p vue que φ est un difféomorphisme.*

Elle induit également d'une base naturelle $(\tilde{\partial}_i = \partial_i \psi \circ \varphi)_{i \in \llbracket 1, p \rrbracket} \in \mathcal{T}\mathcal{M}_n$ de cette espace tangent. L'existence d'un supplémentaire orthogonal en dimension finie assure que,

$$\forall M \in \mathcal{M}_p \quad \mathcal{T}_M \mathcal{M}_n = \mathcal{T}_M \mathcal{M}_p \overset{\perp}{\oplus} (\mathcal{T}_M \mathcal{M}_p)^\perp$$

On admettra qu'il existe $n - k$ champs de vecteurs $(\mathbf{n}_k)_{k \in \llbracket p+1, n \rrbracket} \in \mathcal{T}\mathcal{M}_n$ lisse définie sur \mathcal{M}_p tel que $\forall M \in \mathcal{M}_p$ les vecteurs $(\mathbf{n}_k(P))_{k \in \llbracket p+1, n \rrbracket}$ soit une base orthonormé de $(\mathcal{T}_M \mathcal{M}_p)^\perp$ et que $(\tilde{\partial}_1, \dots, \tilde{\partial}_p, \mathbf{e}_{p+1}, \dots, \mathbf{e}_n)$ soit une base direct :

$$\varepsilon(\tilde{\partial}_1, \dots, \tilde{\partial}_p, \mathbf{e}_{p+1}, \dots, \mathbf{e}_n) \geq 0$$

Les points images de $\partial \mathcal{B}_{1/2}^p$ par $\psi \circ \varphi$ forment le bord de \mathcal{M}_P . Les vecteurs $(\tilde{\partial}_i)_{i \in \llbracket 1, p-1 \rrbracket} \in \mathcal{T}\mathcal{M}_n$ évalué sur $\partial \mathcal{M}_p$ engendre l'espace tangents du bord $\mathcal{T}\partial \mathcal{M}_p$ il est de dimension $p - 1$.

Intégrale d'une n-p-forme où d'une p-forme

Nous sommes maintenant armé pour intégrer des p et des $n - p$ forme s différentielles.

Théorème (et définition) VI.4.6 (Intégrale des p et n-p formes). Pour une variété \mathcal{M}_n paramétrée par $\psi : \mathcal{U} \rightarrow \mathcal{M}$, contenant une sous-variété à bord non vide \mathcal{M}_p constitué d'une seule carte $\psi \circ \varphi : \mathcal{B}_{1/2}^p \rightarrow \mathcal{M}_p$. Nous avons donc une base naturel $\tilde{\partial}_{i=\llbracket 1, p \rrbracket}$ de l'espace tangent $\mathcal{T}\mathcal{M}_p$ à la sous variété et une base $(\mathbf{n}_k(P))_{k \in \llbracket p+1, n \rrbracket}$ de son supplémentaire orthogonal tel qu'introduit dans la définition VI.4.5. Alors pour tout $i = \llbracket 1, p \rrbracket$, on pose $\ell_i = dx^i \partial_i$ (où il n'y a pas de sommation et $dx^i \in \mathbb{R}$ est un accroissement élémentaire) un maillage de la sous-variété \mathcal{M}_p .

On définira également l'intégrande,

$$d^p V = \varepsilon(\tilde{\partial}_1, \dots, \tilde{\partial}_p, \mathbf{e}_{p+1}, \dots, \mathbf{e}_n) dx^1 \dots dx^p$$

Soit un compacte $\mathcal{V}_p \subset \mathcal{M}_p$ de \mathcal{M}_p (c'est-à-dire que $\mathcal{V}_p = \varphi^{-1} \circ \phi(\mathcal{V})$ est compact de $\mathcal{B}_{1/2}^p$ d'intérieur non vide). Alors pour tout p-forme différentielle ω , pour toute $n - p$ -forme différentielle Ω , on définit :

$$\left\{ \begin{array}{l} \int_{\mathcal{V}_p} \omega \hat{=} \int_{\mathcal{V}_p} \omega \circ \psi \circ \varphi(x^1, \dots, x^p)(\ell_1, \dots, \ell_p) \\ \quad = \int_{\mathcal{V}_p} \omega \circ \psi \circ \varphi(x^1, \dots, x^p)(\tilde{\partial}_1, \dots, \tilde{\partial}_p) dx^1 \dots dx^p \\ \int_{\mathcal{V}} \Omega \hat{=} \int_{\mathcal{V}} \star \Omega \end{array} \right.$$

l'intégrale sur \mathcal{V}_p de ω ou de Ω . Ces quantités sont invariantes par changement de paramétrage conservant l'orientation. On à également,

$$\left\{ \begin{array}{l} \int_{\mathcal{V}_p} \omega = \int_{\mathcal{V}_p} \star \omega (\psi \circ \varphi(x^1, \dots, x^p)) (\mathbf{e}_{p+1}, \dots, \mathbf{e}_n) d^p V \\ \int_{\mathcal{V}_p} \Omega = \int_{\mathcal{V}_p} \star \Omega (\psi \circ \varphi(x^1, \dots, x^p)) (\tilde{\partial}_1, \dots, \tilde{\partial}_p) d^p V \end{array} \right.$$

La preuve de l'invariance vis-à-vis des paramétrages de la sous variété est du même type en plus technique que celles de VI.4.1. Enfin, les dernières relations proviennent des propriétés concernant les formes différentielles et du tenseur de Levi-Civita

Cas des surfaces de l'espace ordinaire de dimension 3 Considérons le cas d'une nappes $\varphi : \mathcal{B}_{1/2}^2 \rightarrow \Sigma \subset \mathbb{R}^3$ à bords plongé dans \mathbb{R}^3 , soit \vec{B} un vecteur de dual métrique \underline{B} qui est une

1-forme différentielle. Alors les relations Eq.VI.4.6 impliquent,

$$\begin{aligned}
 \int_{\Sigma} \star \underline{\mathbf{B}} &= \int_{\mathcal{B}_{1/2}^2} \underline{\mathbf{B}}(\varphi(x, y))(\mathbf{e}_3) \epsilon(\tilde{\partial}_1, \tilde{\partial}_2, \mathbf{e}_3) dx dy \\
 &= \int_{\mathcal{B}_{1/2}^2} \vec{\mathbf{B}}(\varphi(x, y)) \cdot \mathbf{e}_3 (\tilde{\partial}_1 \wedge \tilde{\partial}_2) \cdot \mathbf{e}_3 dx dy \\
 &= \int_{\mathcal{B}_{1/2}^2} \vec{\mathbf{B}}(\varphi(x, y)) \cdot \mathbf{n} ||\tilde{\partial}_1 \wedge \tilde{\partial}_2|| dx dy \\
 &= \int_{\mathcal{B}_{1/2}^2} \vec{\mathbf{B}}(\varphi(x, y)) \cdot \mathbf{d}^2 S
 \end{aligned} \tag{VI.159}$$

où on a posé $\mathbf{n} = \mathbf{e}_3 = \tilde{\partial}_1 \wedge \tilde{\partial}_2 / ||\tilde{\partial}_1 \wedge \tilde{\partial}_2||$ et $\mathbf{d}^2 S = ||\tilde{\partial}_1 \wedge \tilde{\partial}_2|| dxdy \mathbf{n}$.

VI.4.3 Théorème de Stockes

Dans le cas où une p -forme différentiel est exact, l'expression de son intégrale se simplifie. C'est l'objet du théorème de Stockes.

Formulation générale

Théorème (et définition) VI.4.7 (Theoreme de Stockes). Soit pour $p \geq 1$ et \mathcal{M}_p une sous variété à bords de dimension p compacte et une $p - 1$ -forme différentielle $\omega \in \mathcal{A}_{p-1}\mathcal{M}$.

$$\int_{\mathcal{M}_p} \mathbf{d}\omega = \int_{\partial\mathcal{M}_p} \omega$$

où la seconde intégrale est définie à l'aide des vecteurs $(\tilde{\partial}_i)_{i \in \llbracket 1, p-1 \rrbracket}$ induisant l'orientation du bord de la sous-variété \mathcal{M}_p

Dans l'espace ordinaire de dimension 3

Comment se traduit ce théorème en dimension 3 ?

Stockes Soit une surface régulière Σ compacte à bords régulier compacte. Le théorème V.4.7 a permis de montrer que le bord de cette surface était une réunion finie de courbe paramétrée fermé simple. On note $\partial\Sigma = \bigcup_{i=1}^N \partial\Sigma_i$. Pour toute 2-forme différentielle \mathbf{B} exacte, on pose \mathbf{A} la une 1-forme différentielle primitive $\mathbf{B} = \mathbf{d}\mathbf{A}$ et $\vec{\mathbf{A}}$ sont dual métrique. On posera $\underline{\mathbf{V}} = \star \underline{\mathbf{B}}$ et $\vec{\mathbf{V}}$ le dual métrique associé on a donc $\underline{\mathbf{V}} = \star \mathbf{d}\mathbf{A}$ et donc $\vec{\mathbf{V}} = \text{rot } \vec{\mathbf{A}}$, alors les résultats obtenue en Eq.VI.159 impliquent,

$$\int_{\Sigma} \mathbf{B} = \int_{\Sigma} \vec{\mathbf{V}}(\varphi(x, y)) \cdot \mathbf{d}^2 S = \int_{\Sigma} \text{rot } \vec{\mathbf{A}} \cdot \mathbf{d}^2 S$$

Le théorème de Stockes implique,

$$\int_{\Sigma} \mathbf{B} = \sum_{i=1}^k \int_{\partial\Sigma_i} \mathbf{A} = \sum_{i=1}^k \oint_{\partial\Sigma_i} \vec{\mathbf{A}} \cdot \mathbf{d}\ell_i$$

où $d\ell$ est le maillage associé au paramétrage du bord Σ_i . Comment l'orienté ? En effet dans un paramétrage φ local au bord, $d\ell \parallel \partial_x \varphi$ car localement le bord est l'image de la droite des abscisses. Afin d'avoir $d\ell, \partial_x \varphi, \mathbf{n}$ orienté directement on choisira $d\ell \cdot \partial_x \varphi \geq 0$. Ce qui implique.

$$\int_{\Sigma} \text{rot } \vec{\mathbf{A}} \cdot \mathbf{d}^2 S = \sum_{i=1}^k \oint_{\partial\Sigma_i} \vec{\mathbf{A}} \cdot \mathbf{d}\ell_i \tag{VI.160}$$

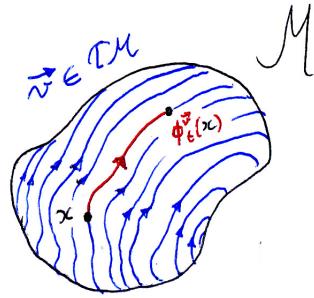


FIGURE VI.8 – Illustration de l'action du flot sur un point de \mathcal{M} . Le flot d'un champ de vecteurs transporte un point le long des lignes de champ de ce vecteur.

Green-Ostrogradski On considère un compacte connexe par arc d'intérieur non vide \mathcal{U} dont le bord est une surface compact $\partial\mathcal{U}$ fermé sans bord, alors pour toutes 2-forme différentielle $\underline{\mathbf{A}}$, $\mathbf{d}\underline{\mathbf{A}}$ est une 3-forme différentielle. Et donc en posant $\star\underline{\mathbf{A}} = \underline{\mathbf{v}}$ et en utilisant l'Eq.VI.158,

$$\int_{\partial\mathcal{U}} \underline{\mathbf{A}} = \int_{\mathcal{U}} \mathbf{d}\underline{\mathbf{A}} = \int_{\mathcal{U}} \star \mathbf{d}\underline{\mathbf{A}} d^3V = \int_{\mathcal{U}} \operatorname{div} \vec{\mathbf{v}} d^3V = \int_{\partial\mathcal{U}} \star \underline{\mathbf{v}}$$

Soit, en utilisant Eq.VI.159,

$$\boxed{\int_{\mathcal{U}} \operatorname{div} \vec{\mathbf{v}} d^3V = \int_{\partial\mathcal{U}} \vec{\mathbf{v}} \cdot \mathbf{d}\mathbf{S}} \quad (\text{VI.161})$$

VI.5 Flots et dérivée de Lie

VI.5.1 Flot d'un champ de vecteurs

Équation différentielle dans une variété

On sait que tout vecteur peut être considéré comme la dérivée d'un arc paramétré en un point : $\forall \vec{\mathbf{v}} \in T_x \mathcal{M}, \exists \gamma : I \rightarrow \mathcal{M} / \gamma(0) = x, \vec{\gamma}'(0) = \vec{\mathbf{v}}$.

Définition VI.5.1. Soit $\vec{\mathbf{v}} \in \mathcal{T}\mathcal{M}$. On dit que γ , arc paramétré, est une ligne de champ de $\vec{\mathbf{v}}$ si $\forall \lambda \in I, \vec{\gamma}'(\lambda) = \vec{\mathbf{v}}(\gamma(\lambda))$.

Soit $\vec{\mathbf{v}} \in \mathcal{T}\mathcal{M}$, soit $x \in \mathcal{M}$. On sait que le problème de Cauchy $\vec{\mathbf{y}}' = \vec{\mathbf{v}}(y), y(0) = x$ admet une unique solution maximale. On note J_x l'intervalle de définition de la solution maximale. On note

$$D = \{(t, x) | x \in \mathcal{M}, t \in J_x\} \quad (\text{VI.162})$$

l'ensemble de vie du champ $\vec{\mathbf{v}}$, c'est-à-dire le plus grand sous-ensemble de $\mathbb{R} \times \mathcal{M}$ sur lequel on puisse transporter des points le long des lignes du champ $\vec{\mathbf{v}}$. Sous réserve d'existence, pour $t \in \mathbb{R}$, on note

$$D_t = \{x | (t, x) \in D\} \quad (\text{VI.163})$$

l'ensemble des points qui ont été transportés par les lignes de $\vec{\mathbf{v}}$ au bout d'un temps t .

Définition VI.5.2. On appelle flot local maximal de $\vec{\mathbf{v}}$, ou plus simplement flot de $\vec{\mathbf{v}}$, l'application définie ainsi

$$\begin{aligned} \phi^{\vec{\mathbf{v}}} : D &\rightarrow \mathcal{M} \\ (t, x) &\mapsto \phi_t^{\vec{\mathbf{v}}}(x) \end{aligned} \quad (\text{VI.164})$$

où $t \mapsto \phi_t^{\vec{\mathbf{v}}}(x)$ est l'unique solution maximale du problème de Cauchy $\vec{\mathbf{y}}' = \vec{\mathbf{v}}(t), y(0) = x$.

Nous représentons figure VI.8 le flot d'un champ de vecteurs appliqué en un point.

Définition VI.5.3. Si $D = \mathbb{R} \times \mathcal{M}$, on dit que $\phi^{\vec{\mathbf{v}}}$ (ou $\vec{\mathbf{v}}$) est complet.

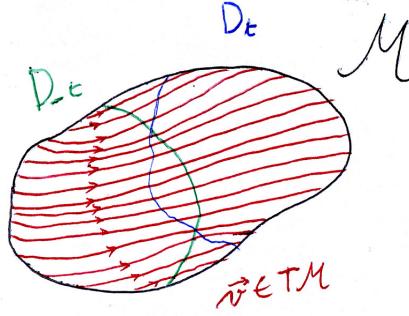


FIGURE VI.9 – Représentation de D_{+t} et D_{-t} pour un flot (ou un champ de vecteurs) donné. Pour simplifier la représentation, on a choisi \mathcal{M} comme étant un ouvert de \mathbb{R}^2 .

Voici quelques propriétés élémentaires du flot.

- D est un ouvert de $\mathbb{R} \times \mathcal{M}$ et $\{0\} \times \mathcal{M} \subset D$.
- Sous réserve d'existence pour $t \in \mathbb{R}$, $\phi_t^{\vec{v}}$ est un \mathcal{C}^1 difféomorphisme de D_{-t} dans D_t . Son inverse est $\phi_{-t}^{\vec{v}}$. (Voir figure VI.9.)
- Si \vec{v} est complet, alors $D_t = \mathcal{M}$. Dans ce cas, pour tout t , $\phi_t^{\vec{v}}$ est un \mathcal{C}^1 difféomorphisme de \mathcal{M} dans \mathcal{M} . Le flot de \vec{v} définit alors un groupe continu à un paramètre qui opère un morphisme de groupes entre $+$ et \circ : $\forall t, s \in \mathbb{R}, \phi_t^{\vec{v}} \circ \phi_s^{\vec{v}} = \phi_{t+s}^{\vec{v}}$. On a donc affaire à un groupe commutatif.

Équations aux variations

La notion de flot consiste à interpréter l'unique solution du problème de Cauchy non plus seulement comme une fonction du temps, mais aussi comme une fonction de la condition initiale. Ainsi, comme pour toute bonne fonction qui se respecte, on peut se demander comment varie le flot lorsque les conditions initiales varient légèrement. Si le flot est différentiable, cela donne

$$\phi_t^{\vec{v}}(x + \vec{h}) = \phi_t^{\vec{v}}(x) + d\phi_t^{\vec{v}}(x) \vec{h} + o(\|\vec{h}\|) \quad (\text{VI.165})$$

En dérivant par rapport au temps, il vient

$$\frac{d\phi_t^{\vec{v}}(x + \vec{h})}{dt} = \vec{v}(\phi_t^{\vec{v}}(x)) + \frac{d}{dt} d\phi_t^{\vec{v}}(x) \vec{h} + o(\|\vec{h}\|) = \vec{v}(\phi_t^{\vec{v}}(x)) + d\vec{v}(\phi_t^{\vec{v}}(x)) \vec{h} + o(\|\vec{h}\|) \quad (\text{VI.166})$$

d'où l'on tire que $d\vec{v}(\phi_t^{\vec{v}}(x)) = \frac{d}{dt} d\phi_t^{\vec{v}}(x)$. En effectuant une intégration, on vérifie donc au passage que si le champ \vec{v} est différentiable sur \mathcal{M} , alors le flot l'est aussi. Nous représentons figure VI.10 la géométrie de l'équation aux variations.

Une question qu'on peut souvent être amené à se poser est de savoir comment varie la solution d'une équation différentielle si l'on change légèrement sa condition initiale. En supposant que l'on reste dans le domaine de linéarité, et si l'on suppose $\phi_t^{\vec{v}}(x)$ connu, on voudrait savoir comment varie \vec{h} lorsque la condition initiale n'est pas x_0 mais $x_0 + \vec{h}_0$. Dans ce cas, un simple calcul permet de montrer que

$$\frac{d\vec{h}}{dt} = d\vec{v} \vec{h} \quad (\text{VI.167})$$

mais on peut faire mieux. En effet, puisque l'on a

$$x(t) + \vec{h}(t) = \phi_t^{\vec{v}}(x_0 + \vec{h}_0) + o(\|\vec{h}\|) = \phi_t^{\vec{v}}(x_0) + d\phi_t^{\vec{v}}(x_0) \vec{h}_0 + o(\|\vec{h}\|) \quad (\text{VI.168})$$

on en déduit ce résultat intéressant

$$\vec{h}(t) = d\phi_t^{\vec{v}}(x_0) \vec{h}_0 \quad (\text{VI.169})$$

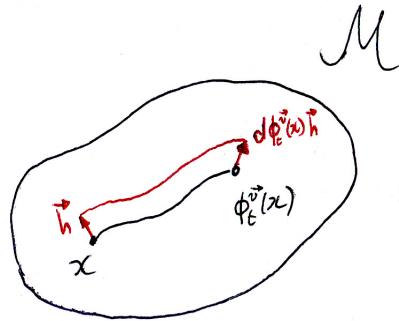


FIGURE VI.10 – Action du flot sur l'espace tangent. Il faut prendre sa différentielle.

Du point de vue des simulations numériques, ce résultat peut avoir un grand intérêt. En effet, on peut souvent calculer numériquement les dérivées partielles des quantités calculées numériquement au temps t en fonction des paramètres du modèle et des conditions initiales à $t = 0$. En effet, en supposant qu'on sache calculer numériquement ϕ_t^v via des intégrations numériques, alors on aura

$$\frac{\partial \phi_t^v(x_0)}{\partial x_0^i} \approx \frac{1}{2\delta} \left[\phi_t^v(x_0^1, \dots, x_0^i + \delta, \dots, x_0^n) - \phi_t^v(x_0^1, \dots, x_0^i - \delta, \dots, x_0^n) \right] \quad (\text{VI.170})$$

Une fois qu'on a calculé toutes les dérivées partielles (ce qui peut prendre du temps) on peut étudier très facilement les variations linéaires de toute quantité le long des trajectoires connues. Un point délicat sera le choix de δ : il faut qu'il soit suffisamment grand pour que la modification soit très grande devant le bruit numérique, mais suffisamment petit pour rester dans un voisinage linéaire du point considéré. L'existence d'un tel δ n'est pas assurée a priori.

Les équations aux variations nous permettent d'étudier la dynamique au voisinage des points d'équilibres pour lesquels $\vec{v}(x_{eq}) = \vec{0}$. En effet, au voisinage d'un tel point d'équilibre x_{eq} , si l'on s'intéresse à l'évolution de $x_{eq} + \vec{h}$, alors les variations ne se feront qu'à l'ordre 1, on aura

$$\forall t \in \mathbb{R}, \frac{d\vec{h}}{dt} = d\vec{v}(x_{eq}) \vec{h} \quad (\text{VI.171})$$

qui est une équation différentielle linéaire à coefficients constants (on sait que les solutions maximales des équations différentielles linéaires sont définies sur \mathbb{R} , voir le chapitre sur les équations différentielles). Plusieurs cas sont possibles. Si le rang de $d\vec{v}(x_{eq})$ est plus petit que n , alors il faudra pousser le calcul à des ordres supérieurs pour pouvoir dire quelque chose de la dynamique au voisinage de l'équilibre. En effet, si $\vec{h}(0)$ est dans le noyau de $d\vec{v}(x_{eq})$, on serait tenté de dire que $\forall t, \vec{h}(t) = \vec{h}(0)$, mais c'est potentiellement faux car sur le noyau de $d\vec{v}(x_{eq})$, l'approximation linéaire ne nous apprend rien du comportement des solutions.

Si le rang de $d\vec{v}(x_{eq})$ est égal à n , pour calculer son exponentielle et résoudre le système, on peut être tenté de le diagonaliser⁸. Supposons que $d\vec{v}(x_{eq})$ est diagonalisable et notons (λ_i) ses valeurs. On peut distinguer trois cas :

1. $\forall i, Re(\lambda_i) \leq 0$: le point d'équilibre est stable ($|e^{\lambda_i t}|$ est borné pour $t > 0$). C'est souvent le cas des points d'équilibres des systèmes dissipatifs (exemple : oscillateur harmonique amorti par un frottement fluide, circuits électriques RLC , systèmes gravitationnels avec marées, ...).
2. $\exists i / Re(\lambda_i) > 0$: le point d'équilibre est instable, il présente ce qu'on appelle une direction hyperbolique ($|e^{\lambda_i t}|$ n'est pas borné).
3. $\forall i, Re(\lambda_i) = 0$: le point d'équilibre est ce qu'on appelle un point elliptique ($|e^{\lambda_i t}| = 1$). La combinaison linéaire des exponentielles complexes devant rester réelle, on aura des signaux quasi-périodiques⁹. C'est le cas des systèmes gravitationnels conservatifs par exemple¹⁰.

8. C'est très souvent le cas dans les systèmes mécaniques puisqu'ils sont dérivés d'un lagrangien, ainsi la différentielle de la vitesse apparaît comme la hessienne d'un potentiel, symétrique, donc diagonalisable.

9. Un signal quasipériodique est un signal de la forme $s(t) = \sum_k s_k e^{i\lambda_k t}$ où les λ_k sont des réels quelconques.

10. La linéarisation des équations planétaires dans le Système Solaire donne ce genre de comportement. On consultera par exemple l'article de Laskar, Boué, Correia, *Tidal dissipation in multi-planet systems and constraints on orbit fitting*, A&A, 538, A105 (2012), disponible ici <http://dx.doi.org/10.1051/0004-6361/201116643>.

Exercice : pendule simple. On considère le système différentiel ordinaire

$$\frac{d}{dt} \begin{pmatrix} \theta \\ J \end{pmatrix} = \begin{pmatrix} J \\ -\sin \theta \end{pmatrix} = \vec{\mathbf{F}}(\theta, J) \quad (\text{VI.172})$$

(C'est simplement l'équation différentielle $y'' + \sin y = 0$ réécrite sous forme d'un système ordinaire du premier ordre). Étudier la dynamique au voisinage des points d'équilibres.

Solution. On voit que θ est une variable 2π périodique (on dit que c'est une variable cyclique). On étudiera la dynamique au voisinage des points d'équilibre en posant $(\theta, J) = (\theta_{eq}, J_{eq}) + \vec{h}$ $= (\theta_{eq}, J_{eq}) + (\delta\theta, \delta J)$ et on cherchera à étudier la dynamique linéaire de $(\delta\theta, \delta J)$.

Les points d'équilibre sont ceux pour lesquels le deuxième terme de l'équation s'annule (la vitesse sera nulle). Cela impliquera donc $J = 0$ et $\theta \equiv 0[\pi]$. On étudiera l'équilibre au voisinage de $(\theta_{eq}, J_{eq}) = (0, 0)$ puis $(\pi, 0)$. La différentielle de $\vec{\mathbf{F}}$ se calcule aisément

$$d\vec{\mathbf{F}}(\theta, J) = \begin{pmatrix} 0 & 1 \\ -\cos \theta & 0 \end{pmatrix} \quad (\text{VI.173})$$

Étudions la dynamique linéaire au voisinage de $(\theta_{eq}, J_{eq}) = (0, 0)$. On a

$$d\vec{\mathbf{F}}(0, 0) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (\text{VI.174})$$

De telle sorte que

$$\frac{d}{dt} \begin{pmatrix} \delta\theta \\ \delta J \end{pmatrix} = d\vec{\mathbf{F}}(0, 0) \begin{pmatrix} \delta\theta \\ \delta J \end{pmatrix} = \begin{pmatrix} \delta J \\ -\delta\theta \end{pmatrix} \quad (\text{VI.175})$$

On montre facilement que le point de coordonnées $(\delta\theta, \delta J)$ a un mouvement circulaire uniforme¹¹ dans le sens trigonométrique. On a affaire à un point elliptique qui correspond à un équilibre stable. Physiquement cela correspond à la configuration où la masse du pendule est en dessous du centre de rotation.. Au voisinage du point d'équilibre $(\theta_{eq}, J_{eq}) = (\pi, 0)$, on a

$$\frac{d}{dt} \begin{pmatrix} \delta\theta \\ \delta J \end{pmatrix} = \begin{pmatrix} \delta J \\ \delta\theta \end{pmatrix} \quad (\text{VI.176})$$

En injectant la première ligne dans la deuxième, on trouve $\delta\ddot{J} = \delta J$. La solution est de la forme $\delta J(t) = a\cosh t + b\sinh t$ où $a, b \in \mathbb{R}$ où cosh et sinh sont les fonctions cosinus et sinus hyperboliques. On aura par ailleurs $\delta\theta = \delta\dot{J} = a\sinh t + b\cosh t$. Ce point est un point d'équilibre instable car les deux variables divergent. Physiquement, cela correspond à la configuration où la masse du pendule est au dessus du point de rotation. Par ailleurs, on constate que $(\delta\theta)^2 - (\delta J)^2 = a^2 - b^2 = \text{cst}$, les trajectoires sont donc des hyperboles d'asymptotes $\delta J = \pm\delta\theta$. Cela justifie la terminologie "point hyperbolique". On donne une représentation du portrait de phase du pendule avec un zoom sur les points d'équilibre figure VI.11.

Intégrales premières et submersion

Définition VI.5.4. Soit $f : \mathcal{M} \rightarrow \mathbb{R}$, et soit ϕ le flot d'un champ de vecteurs. On dit que f est une intégrale première du flot si elle est invariante par le flot, c'est-à-dire $\forall t, f \circ \phi_t = f$.

Si f est une intégrale première, alors la solution du système différentiel considéré est contrainte de vivre dans une sous-variété. En effet, si l'on se dote d'une carte (x^i) , et que f est une intégrale première, alors on aura $f(x^i) = \text{cst}$. Si f est une submersion, alors toute solution est contrainte de vivre dans la sous-variété définie par submersion. S'il y a d'autres intégrales premières, alors cela définit des sous-sous-variétés définies par les intersections des sous-variétés. C'est une méthode classique : utiliser les intégrales premières pour restreindre l'ensemble dans lequel on cherche les solutions de l'équation différentielle à résoudre.

Un exemple qui sera vu abondamment en M2 DSG sera ce qu'on appelle les variables action-angle. Considérons un système à $2n$ degrés de liberté. On peut parfois réduire ce système à des variables

11. Par exemple, poser $z = \delta\theta + \beta\delta J$ et établir une équation différentielle sur z et voir comment il varie dans le plan complexe.

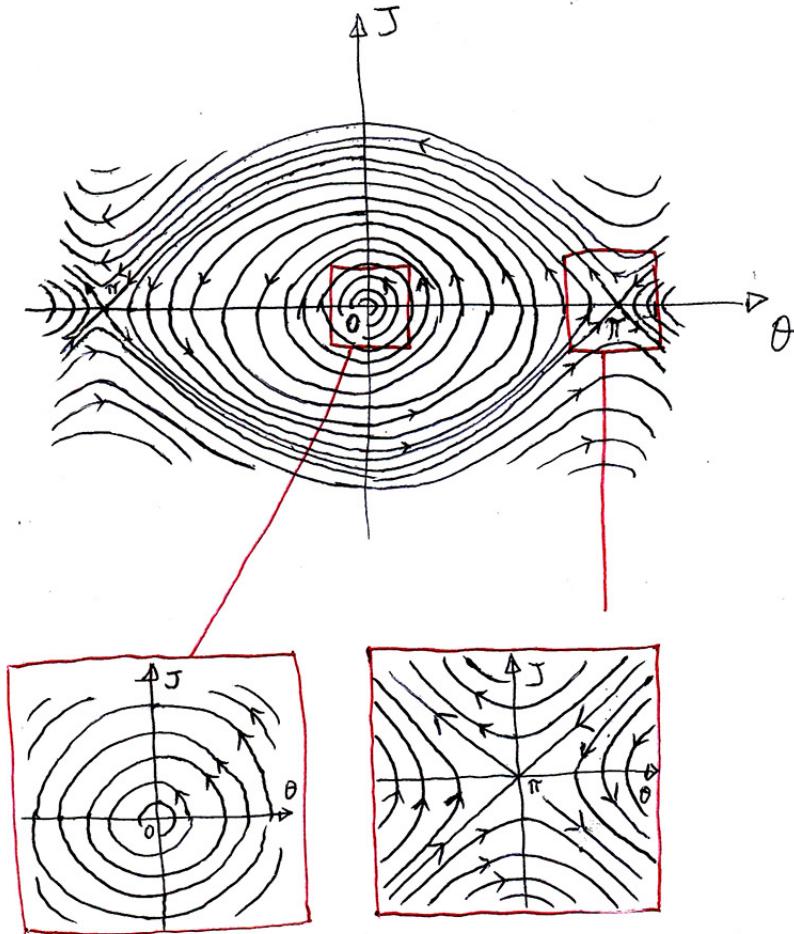


FIGURE VI.11 – Nous représentons le portrait de phase du pendule simple (pour cela il suffit de tracer les différentes courbes des niveaux d'énergie, à savoir $J^2/2 - \cos \theta = \text{cst}$). nous faisons un zoom sur les deux points d'équilibre. Nous y retrouvons les hyperboles pour le point hyperbolique (à droite) et les cercles pour le point elliptique (à gauche). Si nous avions pris des unités "physiques", il y aurait une affinité sur chacun des axes, ce qui justifierait la terminologie elliptique du point. La dynamique des séparatrices et de leur destruction lorsqu'on ajoute une perturbation non intégrable sera étudiée en M2 DSG.

action-angles définies de telle sorte qu'il existe un changement de variables dans lequel le système différentiel s'écrit ainsi

$$\frac{d}{dt} \begin{pmatrix} \theta \\ I \end{pmatrix} = \begin{pmatrix} 0 \\ \omega(I) \end{pmatrix} \quad (\text{VI.177})$$

On peut résoudre ce système ainsi : $I(t) = I_0$ et $\theta(t) = \omega(I)t + \theta_0$. La solution vit donc dans un tore de dimension n dont le i -ème rayon rayon est la composante i de I et dont le i -ème angle est la composante i de θ . Considérons par exemple l'oscillateur harmonique :

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix} \quad (\text{VI.178})$$

Si l'on effectue le changement de variables en posant $x = \sqrt{2I} \cos \theta$ et $y = \sqrt{2I} \sin \theta$, on peut facilement montrer qu'on a $\dot{I} = 0$, $\dot{\theta} = -1$.

VI.5.2 Dérivée de Lie

Minimum vital sur l'image réciproque du flot

La dérivée de Lie le long d'un vecteur sera simplement la dérivée temporelle du flot. Avant de passer à la dérivée de Lie en toute généralité, il faut voir comment agit le flot sur les tenseurs. On va le faire d'abord sur les champs scalaires, puis sur les vecteurs, et enfin sur les tenseurs.

Définition VI.5.5. Soit f un champ scalaire sur \mathcal{M} et soit $\vec{u} \in T\mathcal{M}$. On définit l'image réciproque de f sur \mathcal{M} par $\phi_t^{\vec{u}}$ le champ scalaire défini ainsi

$$\phi_t^{\vec{u}*} f = f \circ \phi_t^{\vec{u}} \quad (\text{VI.179})$$

Définition VI.5.6. Soient $\vec{u}, \vec{v} \in T\mathcal{M}$, et $x \in \mathcal{M}$. L'image réciproque de \vec{v} par $\phi_t^{\vec{u}}$ en x se définit ainsi

$$\phi_t^{\vec{u}*} \vec{v}_x = \left(d\phi_t^{\vec{u}} \right)^{-1}_{(\phi_t^{\vec{u}}(x))} \vec{v}_{(\phi_t^{\vec{u}}(x))} \quad (\text{VI.180})$$

Pour plus de clarté, on a mis en indice les points en lesquels les objets sont définis.

Définition VI.5.7. On définit l'image directe d'un champ de vecteur \vec{v} par $\phi_t^{\vec{u}}$ en x le champ de vecteur défini ainsi

$$\phi_{t*}^{\vec{u}} \vec{v}_{\phi_t^{\vec{u}}(x)} = d\phi_t^{\vec{u}} \vec{v}_x \quad (\text{VI.181})$$

En fait, on a $\phi_{t*}^{\vec{u}} = (\phi_t^{\vec{u}*})^{-1}$.

Définition VI.5.8. Soit $T \in T_s^0 \mathcal{M}$. Soit $\vec{u} \in T\mathcal{M}$. L'image réciproque de T par le flot $\phi_t^{\vec{u}}$ se définit ainsi

$$\phi_t^{\vec{u}*} T_x (\vec{v}_1, \dots, \vec{v}_s) = T_{(\phi_t^{\vec{u}}(x))} \left(d\phi_t^{\vec{u}} \vec{v}_1, \dots, d\phi_t^{\vec{u}} \vec{v}_s \right) \quad (\text{VI.182})$$

Remarque : on a

$$\phi_t^{\vec{u}*} \underline{\omega}_x (\vec{v}) = \underline{\omega}_{\phi_t^{\vec{u}}(x)} (d\phi_t^{\vec{u}} \vec{v}) = \mathbf{g} \left(d\phi_t^{\vec{u}} \vec{\omega}, d\phi_t^{\vec{u}} \vec{v} \right) = \phi_t^{\vec{u}*} \mathbf{g} (\vec{\omega}, \vec{v}) \quad (\text{VI.183})$$

Définition VI.5.9. Soit $\mathbf{T} \in T_s^r \mathcal{M}$. L'image réciproque de \mathbf{T} par le flot $\phi_t^{\vec{u}}$ se définit ainsi

$$\phi_t^{\vec{u}*} \mathbf{T}_x (\underline{u}_1, \dots, \underline{u}_r, \vec{v}_1, \dots, \vec{v}_s) = \mathbf{T}_{\phi_t^{\vec{u}}(x)} (\phi_t^{\vec{u}*} \underline{u}_1, \dots, \phi_t^{\vec{u}*} \underline{u}_r, d\phi_t^{\vec{u}} \vec{v}_1, d\phi_t^{\vec{u}} \vec{v}_s) \quad (\text{VI.184})$$

Propriétés de la dérivée de Lie

Définition VI.5.10. On définira alternativement la dérivée de Lie pour les champs scalaires, les vecteurs, et les tenseurs. Soit \vec{X} un champ de vecteur. On appelle dérivée de Lie le long de \vec{X} , et on note $L_{\vec{X}}$, l'opérateur qui à un élément de $T^{r_s} \mathcal{M}$ associe un autre élément de $T_s^r \mathcal{M}$, défini ainsi

$$L_{\vec{X}} = \frac{d\phi_t^{\vec{X}*}}{dt} \Big|_{t=0} \quad (\text{VI.185})$$

On aura en particulier

$$L_{\vec{X}} f = \frac{df \circ \phi_t^{\vec{X}}}{dt} \Big|_{t=0} \quad (\text{VI.186})$$

$$L_{\vec{X}} \vec{v} = \lim_{t \rightarrow 0} \frac{\phi_t^{\vec{X}*} \vec{v} - \vec{v}}{t} \quad (\text{VI.187})$$

$$L_{\vec{X}} \mathbf{T} = \lim_{t \rightarrow 0} \frac{\phi_t^{\vec{X}*} \mathbf{T} - \mathbf{T}}{t} \quad (\text{VI.188})$$

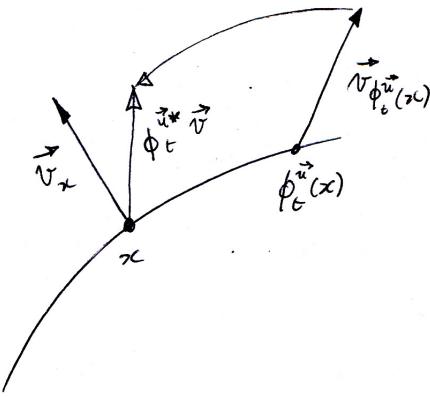


FIGURE VI.12 – Image réciproque d'un champ de vecteurs.

Proposition VI.5.11. Soient $\mathbf{T}, \mathbf{S} \in T_s^r \mathcal{M}$, $\lambda \in \mathbb{R}$. La dérivée de Lie vérifie les propriétés suivantes

- VI.5.11.1. $L_{\vec{\mathbf{X}}}(\mathbf{T} + \lambda \mathbf{S}) = L_{\vec{\mathbf{X}}}\mathbf{T} + \lambda L_{\vec{\mathbf{X}}}\mathbf{S}$ (linéarité).
- VI.5.11.2. $L_{\vec{\mathbf{X}}}(\mathbf{T} \otimes \mathbf{S}) = (L_{\vec{\mathbf{X}}}\mathbf{T}) \otimes \mathbf{S} + \mathbf{T} \otimes (L_{\vec{\mathbf{X}}}\mathbf{S})$ (règle de Leibniz).
- VI.5.11.3. $L_{\vec{\mathbf{X}}}$ commute avec la contraction. On aura donc en particulier $L_{\vec{\mathbf{X}}}(\mathbf{T}(\vec{u}, \dots)) = (L_{\vec{\mathbf{X}}}\mathbf{T})(\vec{u}, \dots) + \mathbf{T}(L_{\vec{\mathbf{X}}}\vec{u}, \dots)$.
- VI.5.11.4. $L_{\vec{\mathbf{X}}}f = \vec{\mathbf{X}}(f)$.
- VI.5.11.5. $L_{\vec{\mathbf{X}}} \vec{\mathbf{Y}} = [\vec{\mathbf{X}}, \vec{\mathbf{Y}}] = \vec{\mathbf{X}} \vec{\mathbf{Y}} - \vec{\mathbf{Y}} \vec{\mathbf{X}}$.

Démonstration. Seuls les points 4 et 5 n'apparaissent pas évident. Démontrons-les sous forme d'exercices. D'abord, signalons que l'on définit $[\vec{\mathbf{X}}, \vec{\mathbf{Y}}]$ comme le commutateur de $\vec{\mathbf{X}}$ et $\vec{\mathbf{Y}}$ par son action sur un champ scalaire. En effet, la notation $\vec{\mathbf{X}} \vec{\mathbf{Y}}$ peut sembler hérétique pour un physicien, pourtant il s'agit bel et bien d'un vecteur. En effet, si f est un champ scalaire, $\vec{\mathbf{Y}} f$ est un champ scalaire, donc rien n'interdit de lui appliquer un vecteur, par exemple $\vec{\mathbf{X}}$, ce qui nous donnera tout simplement $\vec{\mathbf{X}} \vec{\mathbf{Y}} f = \vec{\mathbf{X}}(\vec{\mathbf{Y}} f)$.

1. Calculer le développement limité de $\phi_t^* f$ au voisinage de $t = 0$, conclure et démontrer le point 4.
2. Calculer les coordonnées de $[\vec{\mathbf{X}}, \vec{\mathbf{Y}}]$ dans la base naturelle (ce n'est pas indispensable pour la démonstration mais ça permet de se faire une idée de ce qu'est ce commutateur).
3. Développer $L_{\vec{\mathbf{X}}}(\vec{\mathbf{Y}} f)$ à l'aide de la règle de Leibniz, d'une part, et le calculer comme l'action sur un champ scalaire, d'autre part. Conclure et démontrer le point 5.

Solution.

1.

$$\phi_t^* f(x) = f(\phi_t^*(x)) = f(x + t \vec{\mathbf{X}} + \mathbf{a}(t)) = f(x) + t f(x) \vec{\mathbf{X}} + o(t) = f(x) + t \vec{\mathbf{X}} f(x) + o(t) \quad (\text{VI.189})$$

la suite est évidente.

2. On pourrait faire le calcul en regardant l'action sur un champ scalaire, mais avec un peu d'habitude on peut calculer directement

$$[\vec{\mathbf{X}}, \vec{\mathbf{Y}}] = X^i \vec{\partial}_i (Y^j \vec{\partial}_j) - Y^p \vec{\partial}_p (X^q \vec{\partial}_q) \quad (\text{VI.190})$$

$$= X^i \vec{\partial}_i (Y^j) \vec{\partial}_j + X^i Y^j \vec{\partial}_i \vec{\partial}_j - Y^p \vec{\partial}_p (X^q) \vec{\partial}_q - Y^p X^q \vec{\partial}_p \vec{\partial}_q \quad (\text{VI.191})$$

$$= X^i Y^j \vec{\partial}_j - Y^p X^q \vec{\partial}_q \quad (\text{VI.192})$$

où l'on a posé $\partial_i A = A_{,i}$ pour toute quantité A . La compensation des dérivées secondes vient du théorème de Schwarz. En regroupant les indices muets on obtient finalement

$$[\vec{X}, \vec{Y}] = (X^k Y_{,k}^i - Y^k X_{,k}^i) \vec{\partial}_i \quad (\text{VI.193})$$

ce qui indique les coordonnées du commutateur dans la base naturelle.

3. D'une part, la règle de Leibniz implique

$$L_{\vec{X}}(\vec{Y}f) = L_{\vec{X}}(\vec{Y})f + \vec{Y}(L_{\vec{X}}f) = (L_{\vec{X}}\vec{Y})f + \vec{Y}\vec{X}f. \quad (\text{VI.194})$$

D'autre part, en considérant $\vec{Y}f$ comme un champ scalaire, on a

$$L_{\vec{X}}(\vec{Y}f) = \vec{X}\vec{Y}f \quad (\text{VI.195})$$

L'identification des deux égalités permet de conclure. ■

On a également les propositions naturelles suivantes.

Proposition VI.5.12. Soient $\vec{X}, \vec{Y} \in \mathcal{T}\mathcal{M}$, $\lambda \in \mathbb{R}$. Alors :

$$\text{VI.5.12.1. } L_{\vec{X}+\lambda\vec{Y}} = L_{\vec{X}} + \lambda L_{\vec{Y}}$$

$$\text{VI.5.12.2. } L_{[\vec{X}, \vec{Y}]} = [L_{\vec{X}}, L_{\vec{Y}}] = L_{\vec{X}} \circ L_{\vec{Y}} - L_{\vec{Y}} \circ L_{\vec{X}}$$

Théorème VI.5.13. Soient $\vec{X}, \vec{Y} \in \mathcal{T}\mathcal{M}$. Il y a équivalence entre les trois propositions suivantes.

$$\text{VI.5.13.1. } [\vec{X}, \vec{Y}] = 0$$

$$\text{VI.5.13.2. } [L_{\vec{X}}, L_{\vec{Y}}] = 0$$

$$\text{VI.5.13.3. } \forall t, s, \phi_t^{\vec{X}} \circ \phi_s^{\vec{Y}} = \phi_s^{\vec{Y}} \circ \phi_t^{\vec{X}}$$

On aura besoin de deux lemmes préliminaires.

Lemme VI.5.14. Soit ψ un automorphisme différentiable de \mathcal{M} . Soit $\vec{X} \in \mathcal{T}\mathcal{M}$. On note formellement $\phi_t^{\vec{X}} = e^{t\vec{X}}$ en raison de l'expression du développement en série entière de Taylor d'un champ scalaire. Alors on a

$$\forall t, \psi(e^{t\vec{X}})(x) = e^{t\psi_*\vec{X}}(\psi(x)) \quad (\text{VI.196})$$

Démonstration.

$$\frac{d}{dt} \psi(e^{t\vec{X}}(x)) = \psi_* \vec{X}_x = \frac{d}{dt} e^{t\psi_*\vec{X}}(\psi(x)) \quad (\text{VI.197})$$

Les deux flots vérifient la même équation différentielle et vérifient tous deux la même condition initiale puisque

$$\psi(e^{0\cdot\vec{X}}(x)) = \psi(x) = \psi(e^{0\cdot\vec{X}}(x)) \quad (\text{VI.198})$$

Le théorème de Cauchy-Lipschitz permet de conclure. ■

Lemme VI.5.15. Si $[\vec{X}, \vec{Y}] = 0$, alors $\forall t, e_*^t \vec{Y} \vec{X} = \vec{X}$.

Démonstration.

$$\frac{d}{dt} e_*^t \vec{Y} \vec{X} = -L_{\vec{Y}} \vec{X} = [\vec{X}, \vec{Y}] = 0 \quad (\text{VI.199})$$

par intégration le long du flot, on obtient le résultat annoncé. ■

Désormais la démonstration du théorème est beaucoup plus simple.

Démonstration. (2) \Rightarrow (1) est évident. (3) \Rightarrow (2) s'obtient en dérivant l'équation par rapport à s puis t . Montrons que (1) \Rightarrow (3). Posons $\psi = e^{s\vec{Y}}$. Supposons que le commutateur soit nul. Il vient

$$e^{s\vec{Y}}(e^{s\vec{X}}(x)) = e^{te_*^s\vec{Y}}(e^{s\vec{Y}}(x)) = e^{t\vec{X}}e^{s\vec{Y}}(x) \quad (\text{VI.200})$$

La première égalité vient du premier lemme et la deuxième du second. Ce qui achève la preuve. ■

Un contre-exemple important : le groupe unitaire de \mathbb{R}^3 , noté $O(3)$, constitué des rotations et des symétries axiales et centrales, n'est pas un groupe commutatif. Il est facile de montrer que dans \mathbb{R}^3 la rotation autour de l'axe x et celle autour de l'axe y ne commutent pas. Faites l'expérience avec des objets autour de vous (et pas des objets présentant trop de symétries bien sûr!). L'ordre dans lequel on effectue les transformations changera le résultat, sauf si les générateurs infinitésimaux des transformations ont un commutateur nul.

Par ailleurs ce théorème exprime que l'exponentiation d'opérateurs n'effectue pas un morphisme de groupe entre $+$ et \circ sauf si le commutateur est nul. La formule générale du commutateur de la somme d'opérateurs s'exprime grâce à la formule de Baker-Campbell-Hausdorff que nous donnons pour information mais qui sort complètement du cadre de ce cours. Nous ne donnerons pas de démonstration.

Théorème VI.5.16 (Formule de Baker-Campbell-Hausdorff). Si A et B sont deux opérateurs, alors nous avons l'identité suivante :

$$e^A e^B = e^C \quad (\text{VI.201})$$

où

$$C = \sum_{n>0} \frac{(-1)^{n-1}}{n} \sum_{\substack{r_i+s_i>0 \\ 1 \leq i \leq n}} \frac{[A^{r_1} B^{s_1} A^{r_2} B^{s_2} \dots A^{r_n} B^{s_n}]}{(\sum_{i=1}^n (r_i + s_i)) r_1! s_1! \dots r_n! s_n!} \quad (\text{VI.202})$$

où s_n et r_n sont des entiers positifs, et où nous avons noté :

$$[A^{r_1} B^{s_1} \dots A^{r_n} B^{s_n}] = [\underbrace{A, [A, \dots [A}_{A^{r_1}}, \underbrace{B, [B, \dots [B}_{B^{s_1}}, \dots \underbrace{A, [A, \dots [A}_{A^{r_n}}, \underbrace{B, [B, \dots [B]}_{B^{s_n}}] \dots]. \quad (\text{VI.203})$$

Dans le cas où A et B sont du même ordre et petits devant l'identité, dans le sens où $\mathcal{O}(A) = \mathcal{O}(B) = o(1)$, cette formule donne (ou alors par un "simple" développement de l'exponentielle) à l'ordre 4 :

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[A, [A, B]] + \frac{1}{24}[A, [[A, B], B]] + \mathcal{O}(A^5). \quad (\text{VI.204})$$

On conclut avec un théorème "évident"¹² mais qui exprime tout l'intérêt de la dérivée de Lie.

Théorème VI.5.17. Soit $\vec{X} \in \mathcal{T}\mathcal{M}$, et $T \in T_s^r \mathcal{M}$. Les deux propositions sont équivalentes.

VI.5.17.1. $\forall t, (e^{t\vec{X}})^* T = T$.

VI.5.17.2. $L_{\vec{X}} T = 0$.

Coordonnées naturelles de la dérivée de Lie

Sous forme d'exercice. Soit $\vec{X} = X^i \vec{\partial}_i$, et $T = T^{i_1 \dots i_r}_{j_1 \dots j_s} \vec{\partial}_{i_1} \otimes \dots \otimes \vec{\partial}_{i_r} \otimes \underline{dx}^{j_1} \otimes \dots \otimes \underline{dx}^{j_s}$.

1. Calculer les coordonnées de $L_{\vec{X}} \underline{dx}^i$ dans la base des formes coordonnées.
2. Calculer les coordonnées de $L_{\vec{X}} \vec{\partial}_i$ dans la base naturelle.
3. En utilisant la règle de Leibniz et la commutation avec la contraction, en déduire les coordonnées de $L_{\vec{X}} T$ en fonction des coordonnées de T .

Solution.

12. Pour ne pas dire "trivial"...

1. Calculons l'action de $L_{\vec{X}} \underline{\mathbf{dx}}^i$ sur un vecteur \vec{Y} . En utilisant la règle de Leibniz et la commutation avec la contraction, on a

$$L_{\vec{X}} (\underline{\mathbf{dx}}^i (\vec{Y})) = L_{\vec{X}} (\underline{\mathbf{dx}}^i) \vec{Y} + \underline{\mathbf{dx}}^i (L_{\vec{X}} \vec{Y}) = L_{\vec{X}} (\underline{\mathbf{dx}}^i) \vec{Y} + (X^k Y_{,k}^i - Y^k X_{,k}^i) \quad (\text{VI.205})$$

Et d'autre part, en considérant $\underline{\mathbf{dx}}^i (\vec{Y}) = Y^i$ comme un champ scalaire, on obtient

$$L_{\vec{X}} (\underline{\mathbf{dx}}^i (\vec{Y})) = \vec{X}(Y^i) = X^j Y_{,j}^i \quad (\text{VI.206})$$

d'où

$$L_{\vec{X}} (\underline{\mathbf{dx}}^i) \vec{Y} = \vec{X}(Y^i) = X^j Y_{,j}^i - (X^k Y_{,k}^i - Y^k X_{,k}^i) = X_{,k}^i Y^k = X_{,k}^i \underline{\mathbf{dx}}^k (\vec{Y}) \quad (\text{VI.207})$$

donc on a

$$L_{\vec{X}} \underline{\mathbf{dx}}^i = X_{,k}^i \underline{\mathbf{dx}}^k \quad (\text{VI.208})$$

2.

$$L_{\vec{X}} \vec{\partial}_i = [\vec{X}, \vec{\partial}_i] = X^j \vec{\partial}_j \vec{\partial}_i - X_{,i}^j \vec{\partial}_j - X^j \vec{\partial}_i \vec{\partial}_j = -X_{,i}^j \vec{\partial}_j \quad (\text{VI.209})$$

3. On utilise exactement la même méthode, en utilisant le fait que

$$\begin{aligned} L_{\vec{X}} (T (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s})) &= (L_{\vec{X}} T) (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s}) \\ &\quad + \sum_{\ell=1}^r T (\underline{\mathbf{dx}}^{i_1}, \dots, L_{\vec{X}} \underline{\mathbf{dx}}^{i_\ell}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s}) \\ &\quad + \sum_{m=1}^s T (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, L_{\vec{X}} \vec{\partial}_{j_m}, \dots, \vec{\partial}_{j_s}) \end{aligned} \quad (\text{VI.210})$$

$$\begin{aligned} &= (L_{\vec{X}} T) (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s}) \\ &\quad + \sum_{\ell=1}^r T (\underline{\mathbf{dx}}^{i_1}, \dots, X_{,k}^i \underline{\mathbf{dx}}^k, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s}) \\ &\quad + \sum_{m=1}^s T (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, -X_{,j_m}^k \vec{\partial}_k, \dots, \vec{\partial}_{j_s}) \end{aligned} \quad (\text{VI.211})$$

$$\begin{aligned} &= (L_{\vec{X}} T)^{i_1 \dots i_r}_{j_1 \dots j_s} \\ &\quad + \sum_{\ell=1}^r T^{i_1 \dots k \dots i_r}_{j_1 \dots j_s} X_k^{i_\ell} \\ &\quad - \sum_{m=1}^s T^{i_1 \dots i_r}_{j_1 \dots k \dots j_s} X_{j_m}^k \end{aligned} \quad (\text{VI.212})$$

où k est à la place du ℓ ème indice dans la première somme et à la place du m ème indice dans la deuxième somme. Mais on a aussi

$$L_{\vec{X}} (T (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s})) = \vec{X} (T (\underline{\mathbf{dx}}^{i_1}, \dots, \underline{\mathbf{dx}}^{i_r}, \vec{\partial}_{j_1}, \dots, \vec{\partial}_{j_s})) \quad (\text{VI.213})$$

$$= \vec{X} (T^{i_1 \dots i_r}_{j_1 \dots j_s}) \quad (\text{VI.214})$$

$$= X^k T^{i_1 \dots i_r}_{j_1 \dots j_s, k} \quad (\text{VI.215})$$

On en déduit l'expression des coordonnées de la dérivée de Lie d'un tenseur dans la base naturelle.

$$(L_{\vec{X}} t)^{i_1 \dots i_r}_{j_1 \dots j_s} = X^k T^{i_1 \dots i_r}_{j_1 \dots j_s, k}$$

$$\begin{aligned} & - \sum_{\ell=1}^r T^{i_1 \dots k \dots i_r}_{\quad \quad j_1 \dots j_s} X^{i_\ell}_{,k} \\ & + \sum_{m=1}^s T^{i_1 \dots i_r}_{\quad \quad j_1 \dots k \dots j_s} X^k_{,j_m} \end{aligned} \tag{VI.216}$$

Bibliographie

- AUDOLY, B. (2000). « Elasticité et géométrie : de la rigidité des surfaces à la délamination en fil de téléphone ». Theses. Université Pierre et Marie Curie - Paris VI. URL : <https://tel.archives-ouvertes.fr/tel-00002515>.
- CARMO, M. P. do (1976). *Differential geometry of curves and surfaces*. Prentice Hall, p. I–VIII, 1–503. URL : <http://www2.ing.unipi.it/griff/files/dC.pdf>.
- GOURGOULHON, E. (2010). *Relativité restreinte : Des particules à l'astrophysique*. Savoirs actuels. EDP Sciences. URL : <https://books.google.fr/books?id=gsYRsXj-Gl8C>.
- GOURGOULHON, E. (2014). « Relativité générale ». Lecture. Année universitaire 2013-2014
Observatoire de Paris. URL : <https://cel.archives-ouvertes.fr/cel-00366315>.
- GUEDJ, V. (2015). *Géométrie différentielle*. URL : <https://www.math.univ-toulouse.fr/~guedj/fichierspdf/GeomDiff2015.pdf>.
- HELEIN, F. (2016). *Systèmes différentiels extérieurs*. URL : <https://webusers.imj-prg.fr/~frederic.helein/cours/sdem1.pdf>.
- POL, D. (2012). *Classification des surfaces topologiques compactes*. URL : <https://www.math.univ-angers.fr/~pol/memoirem1.pdf>.
- STRAUMANN, N. (1984). *General relativity and relativistic astrophysics*. Texts and monographs in physics. German version available : Allgemeine Relativitätstheorie und relativistische Astrophysik - Berlin, Springer, 1981. Berlin : Springer. URL : <https://cds.cern.ch/record/103736>.
- VOEDTS, J. (2002). *Cours de Mathématiques*. Ellipses.