né deheuvels

tenseurs et spineurs

puf

MATHÉMATIQUES

Tenseurs et spineurs

COLLECTION DIRIGÉE PAR PAUL DEHEUVELS

MATHÉMATIQUES

Tenseurs et spineurs

RENÉ DEHEUVELS

Professeur honoraire à l'Université de Paris VI
Président honoraire
de l'Ecole Pratique des Hautes Etudes
Professeur à l'Ecole Polytechnique de 1956 à 1980

PRESSES UNIVERSITAIRES DE FRANCE

A Fleur, Sophie, Camille et Aurore

ISBN 2 13 044940 9
ISSN 0246-3822

Dépôt légal — 1re édition : 1993, avril

© Presses Universitaires de France, 1993
108, boulevard Saint-Germain, 75006 Paris

SOMMAIRE

Préface

Ce livre a pour ambition d'exposer les bases algébriques *des calculs tensoriels et spinoriels et d'en indiquer quelques applications.*

Nous sommes partis du fait que les définitions partielles ou sommaires que l'on donne habituellement des tenseurs et des spineurs, pour exactes qu'elles soient dans le contexte où on les utilise, ont souvent pour effet d'engendrer des idées fausses résultant d'une volonté excessive de simplification.

Nous avons donc entrepris d'explorer leur structure sans faire de concession à la facilité.

Suivant les principes de cette collection nous partons de notions élémentaires pour construire peu à peu la théorie en l'appuyant sur des exemples.

La majeure partie de cet ouvrage a été exposée dans des cours à l'Université de Paris 6 et à l'Ecole Polytechnique.

René Deheuvels

CHAPITRE PREMIER

Nature mathématique des grandeurs physiques Bibliographie

Introduction

Les lois physiques expriment des relations entre des grandeurs dont les concepts ont été préalablement dégagés et définis de façon précise. Les grandeurs les plus simples sont *de type scalaire*, ce qui signifie qu'une fois choisies des grandeurs étalons comme unités, elles s'expriment par un nombre unique qui les « mesure ». Il en est ainsi, par exemple, des intervalles de temps, des dimensions, du volume ou de la masse d'un corps, de la pression ou de la température d'un fluide en un point, de la vitesse ou de l'accélération d'un point mobile sur une droite de la charge électrique, etc...

Certaines de ces grandeurs scalaires se définissent à partir d'autres : le volume à l'aide de la longueur, la vitesse, à l'aide de la longueur et du temps, etc...

Ayant choisi des grandeurs indépendantes comme fondamentales, par exemple la longueur L, la masse M, le temps T, chaque grandeur qui s'en déduit est caractérisée par une « dimension », monôme en les trois variables L, M, T, qui permet, en particulier, de définir l'unité de cette grandeur à partir des unités des grandeurs fondamentales. Exemples de « dimension » : surface L^2, volume L^3, accélération $L\,\overset{-2}{T}$, densité $M\,\overset{-3}{L}$. Une égalité exprimant une loi physique doit évidemment être indépendante du choix des unités des grandeurs qu'elle relie : les « dimensions » de ses deux membres doivent être les mêmes. Cela revient à dire que l'on peut multiplier les longueurs par un même nombre λ, les temps par τ, les masses par μ sans changer la relation. Cette condition suffit souvent à imposer la forme de la loi. Par exemple, sachant que la période T des petites oscillations d'un pendule est une fonction φ de sa longueur $\mathtt{l}$ et l'accélération de la pesanteur $\mathtt{g}$, on doit avoir :

$\tau \mathrm{T} = \varphi(\lambda \mathtt{l}, \lambda \overset{-2}{\tau} \mathtt{g})$ quels que soient τ et λ ; d'où, en prenant $\lambda = \dfrac{1}{\mathtt{l}}$ et $\lambda \overset{-2}{\tau} = \dfrac{1}{\mathtt{g}}$:

$\mathrm{T} = \dfrac{1}{\tau} \varphi(\lambda \mathtt{l}, \lambda \overset{-2}{\tau} \mathtt{g}) = \varphi(1,1)\sqrt{\dfrac{\mathtt{l}}{\mathtt{g}}} = k \sqrt{\dfrac{\mathtt{l}}{\mathtt{g}}}$ où k est une constante sans dimension.

L'usage de *vecteurs* s'avère indispensable si l'on veut représenter les grandeurs qui possèdent une direction et un sens : vitesses, accélérations, forces, champs électriques ou magnétiques en un point, etc... Si l'on fixe trois axes de coordonnées, le repérage d'un vecteur de notre espace se fait au moyen de trois nombres, ses composantes. Mais un vecteur ne peut naturellement pas être assimilé à trois scalaires : les composantes n'ont pas de signification intrinsèque. Elles subissent une transformation linéaire lorsqu'on change d'axes de référence alors que les scalaires sont invariants. Il s'agit donc d'un être mathématique de nature différente.

Dès la fin du XIX$^{\mathrm{e}}$—siècle, on s'est vu contraint d'ajouter aux précédentes des grandeurs d'un nouveau type. En effet, si l'on veut exprimer les tensions en un point à l'intérieur d'un solide élastique déformé, un raisonnement simple qu'on verra plus loin, fournit dans un système d'axes de référence donné, un ensemble de six nombres. On peut penser qu'ils représentent six scalaires ou deux vecteurs. Mais si l'on change d'axes de coordonnées, la transformation subie par ces six quantités ne correspond ni à l'un ni à l'autre cas. On est en présence d'un être mathématique nouveau que son origine a fait appeler tenseur. Il s'avère d'ailleurs que ce « tenseur des tensions » est de nature assez particulière ; il est « symétrique » et « du second ordre ». Or il existe des tenseurs de tous ordres, les vecteurs formant les tenseurs d'ordre un et les scalaires les tenseurs d'ordre zéro. On a ainsi à sa disposition une vaste famille d'êtres mathématiques, naturellement attachés à la géométrie de l'espace, et il n'est pas surprenant que de nombreuses grandeurs physiques ne puissent être représentées ni par des scalaires, ni par des vecteurs, mais qu'elles exigent l'emploi de tenseurs, d'ordres divers. Par exemple, la véritable nature du champ électromagnétique est celle d'un tenseur d'ordre deux de l'espace temps, ce qui fait apparaître le caractère indissoluble du couple : champ électrique, champ magnétique ; la grandeur essentielle de la relativité générale est le tenseur de courbure de l'espace, qui est un tenseur d'ordre quatre !

Nous avons observé que pour tester la nature d'une grandeur physique, on procède à un changement du système de référence et on regarde ce qui se passe.

Cette façon de faire est analogue à l'estimation de la « dimension » d'une grandeur scalaire : on change les unités de référence et on examine ce qu'est devenue la mesure de la grandeur dans les nouvelles unités.

Essayons de préciser la notion de tenseur. Il s'agit d'un « objet mathématique » attaché à un espace donné, comme notre espace à trois dimensions, où l'espace-temps. Cet objet a une existence intrinsèque, mais est décrit, et déterminé, dans chaque base de l'espace, par un ensemble de composantes : une pour un scalaire, n, dimension de l'espace, pour un vecteur, etc... A défaut d'une conceptualisation intrinsèque d'un tenseur T, c'est-à-dire indépendante du choix d'un système d'axes de référence, on peut donc toujours le définir, pragmatiquement, comme un objet possèdant dans chaque base e de l'espace un ensemble de composantes $\mathrm{T}(e)$, ces composantes subissant, lors du passage de la base e à une autre base e', une transformation $\mathrm{R}(e', e)$, qui permet de calculer les composantes de $\mathrm{T}(e')$ à partir de celles de $\mathrm{T}(e)$. Symboliquement : $\mathrm{T}(e') = \mathrm{R}(e', e)\mathrm{T}(e)$. On doit alors avoir : $\mathrm{R}(e, e')\mathrm{R}(e', e) =$ Identité, autrement dit, $\mathrm{R}(e, e')$ est la transformation inverse de $\mathrm{R}(e', e)$. On doit aussi avoir, si e'' est une troisième base :

$$\mathrm{T}(e'') = \mathrm{R}(e'', e')\mathrm{T}(e') = \mathrm{R}(e'', e')\mathrm{R}(e', e)\mathrm{T}(e) = \mathrm{R}(e'', e)\mathrm{T}(e),$$

soit : $\mathrm{R}(e'', e) = \mathrm{R}(e'', e')\mathrm{R}(e', e)$. C'est la nature de ces transformations $\mathrm{R}(e', e)$ qui caractérise la « nature » ou le « type » du tenseur T.

On passe d'une base e de l'espace E à une base e' par une transformation linéaire g, élément du « groupe linéaire » $\mathrm{Gl(E)}$ des transformations linéaires inversibles de l'espace E : $e' = g(e)$. Or, si e_1 est n'importe quelle autre base de E, elle est tranformée par g en une base $e'_1 = ge_1$. On a : $\mathrm{T}(e') = \mathrm{R}(e', e)\mathrm{T}(e)$ et $\mathrm{T}(e'_1) = \mathrm{R}(e'_1, e_1)\mathrm{T}(e_1)$. Puisque c'est la même transformation linéaire g de l'espace E qui permet de passer de e à e', ou de e_1 à e'_1, on doit avoir $\mathrm{R}(e', e) = \mathrm{R}(e'_1, e_1)$ sinon T se comporterait différemment lors du passage de e à e', et de e_1 à e'_1, c'est-à-dire dépendrait du système de référence choisi.

Les transformations tensorielles $\mathrm{R}(e', e)$ ne dépendent donc que de la transformation g faisant passer de e à e' : $\mathrm{R}(e', e) = \mathrm{R}(g)$. La correspondance : $g \in \mathrm{Gl(E)} \to \mathrm{R}(g)$, constitue une *représentation du groupe* Gl(E), ce qui signifie de façon précise que $\mathrm{R}(g' \circ g)$ $= \mathrm{R}(g') \circ \mathrm{R}(g)$. R(Identité) = Identité et $\mathrm{R}(\overset{-1}{g}) = \mathrm{R}^{-1}(g)$.

On serait donc presque tenté de définir les tenseurs d'un type donné, attachés à un espace E comme les éléments d'un espace vectoriel de dimension finie F, muni d'une représentation R du groupe GL(E) dans le groupe des automorphismes linéaires Gl(F), autrement dit de confondre les notions de tenseurs et de représentations de dimension finie du groupe linéaire de l'espace. Nous allons voir que la notion de tenseur est plus riche et contient des éléments de structure supplémentaires qui d'ajoutent à celle de représentation.

Observons aussi qu'il est rare que l'on soit en présence dans les applications, d'un espace E « nu », c'est-à-dire dépourvu de structure autre que sa structure linéaire d'espace vectoriel (cf. S II.1). La plupart du temps E est muni d'une « métrique ». Cela peut être une métrique euclidienne, qui attribue à chaque vecteur une longueur, et, à partir de là, permet de définir toutes les notions métriques : produit scalaire, angle, aire, volume etc... Cela peut être aussi une métrique pseudoeuclidienne comme celle de l'espace-temps de la relativité. Il arrive également que soit donné sur l'espace un produit scalaire antisymétrique, qui définit alors une structure et une géométrie, dites symplectiques.

Les objets mathématiques liés à un espace E doté d'une structure particulière doivent évidemment être considérés en liaison avec elle. Le groupe fondamental n'est plus Gl(E) mais le sous-groupe U(E) que forment les transformations linéaires conservant la structure : groupe orthogonal O(E) pour un espace euclidien, groupe de Lorentz L(E) pour l'espace-temps, groupe symplectique Sp(E) pour un espace symplectique. Par exemple, un vecteur d'un espace euclidien, qui est un tenseur d'ordre un, possède une longueur déterminée qui reste la même par toute transformation orthogonale, alors que la notion de longueur disparaît dès que l'on effectue des transformations linéaires, arbitraires. On doit donc, nécessairement, considérer des *tenseurs euclidiens, lorentziens, symplectiques*.

On se trouve alors dans une situation analogue à celle qui a été brièvement décrite à propos des petites oscillations du pendule. Si une loi physique exprime une égalité entre grandeurs tensorielles, ces dernières doivent être de même nature, c'est-à-dire se transformer de la même façon par le groupe de structure. Cette condition n'est qu'une manifestation de son caractère intrinsèque. Cependant les contraintes de la covariance par le groupe de structure sont souvent suffisantes pour déterminer complètement sa forme. Autrement dit certaines lois physiques élémentaires ont un *caractère d'inévitabilité* ; elles sont une conséquence obligatoire de la covariance et de la géométrie de l'espace. Il en est ainsi de l'équation d'Einstein en relativité générale et de l'équation de Dirac de l'électron.

Poursuivons notre étude des grandeurs physiques. Les grandeurs, scalaires, vectorielles, tensorielles, dont nous venons de parler sont des grandeurs que l'on pourrait qualifier de « ponctuelles », ou élémentaires, et qui sont associées à l'abstraction théorique commode du « point matériel » : champ, tension, pression en un point, etc...

Si l'on considère par exemple l'ensemble des pressions en tous les points d'un fluide, on obtient une grandeur physique qui est maintenant une fonction numérique.

De la même façon les phénomènes électromagnétiques, ou de gravitation se manifestent toujours par des *champs* de vecteurs, ou de tenseurs : à chaque point x de l'espace, on associe un vecteur, ou un tenseur $E(x)$, représentant le champ en ce point, d'où une fonction : $x \to E(x)$ sur l'espace, à valeurs vectorielles, ou dans les tenseurs d'un type donné.

Ces remarques ne concernent pas seulement la physique « macroscopique ». En effet, si à notre échelle, on peut sans dommage utiliser dans les raisonnements des « points matériels » sans dimensions, il n'est plus de même à l'échelle des particules élémentaires.

Un point matériel doté d'une masse, ou d'une charge électrique, aurait une densité infinie ! Il faut donc changer complètement de concepts si l'on souhaite obtenir une description mathématique des phénomènes physiques au niveau de l'atome. C'est l'objet de la mécanique quantique et de la théorie quantique des champs. Les grandeurs physiques ne sont plus représentées par des fonctions, à valeurs scalaires ou tensorielles. Les fonctions et champs sont réservés aux « fonctions d'ondes » et plus généralement à la description des « objets physiques ». Les grandeurs deviennent des opérateurs agissant sur ces fonctions ou ces champs.

Certaines particules élémentaires, électrons, protons, neutrons, possèdent une propriété tout-à-fait inattendue, voire choquante pour l'intuition naïve : par une rotation de 360 degrés (2π) autour d'un axe, leur état ne revient pas coïncider avec l'état initial. On ne peut donc pas représenter l'état d'une telle particule à l'aide de vecteurs ou de tenseurs. Mais il existe bien des « objets mathématiques » attachés à notre espace euclidien, ou à l'espace-temps, qui possèdent cette propriété ! Ce sont les *spineurs*.

Leur existence avait été reconnues par les mathématiciens (Elie Cartan, 1914) mais ils n'avaient jamais été rencontrés ni en géométrie ni en physique.

Le physicien P.A.M. Dirac, par un trait de génie rarement égalé, les a retrouvés dans son équation de l'électron relativiste (1927) dont la fonction « d'onde » est un champ de spineurs de l'espace-temps. La puissance des mathématiques est telle que la logique interne de cette équation a fait prédire par son auteur l'existence de l'antimatière, et en particulier du positron, qui fut découvert cinq ans plus tard.

Il est également indispensable, lorsqu'on est en présence de champs de tenseurs, d'examiner l'effet des diverses opérations de : l'analyse dérivations ou intégrations.

C'est l'objet de *l'analyse tensorielle*.

Jusqu'à présent, nous avons supposé que l'on repérait les points de l'espace-temps, à l'aide d'un système de coordonnées rectilignes.

C'est, bien entendu, ce qui peut sembler le plus naturel et le plus simple. Mais il est d'abord évident que rapporter l'univers tout entier à un système d'axes rectilignes revient à faire le postulat audacieux qu'il a la structure géométrique d'un espace affine, et qu'en particulier il est infini, ce que l'on ignore complètement.

On voit ainsi que par le biais d'outils mathématiques trop simplistes, d'insidieux postulats peuvent se glisser dans les théories physiques, en nuisant à leur rigueur, voire à leur exactitude.

Même l'usage local d'axes de coordonnées rectilignes suppose que les droites existent. On dérive cette existence du postulat métrique : l'espace physique est tel que l'on peut y définir les distances entre ses points.

Les droites sont alors les « plus courts chemins d'un point à un autre » ou géodésiques de l'espace.

Or, depuis le début de ce siècle, les mathématiciens ont effectué une étude complète et extrêmement détaillée des espaces où l'on peut définir une métrique, satisfaisant à des axiomes raisonnables et très généraux : ce sont les variétés riemanniennes (et plus généralement encore, pseudoriemanniennes). Si l'on veut faire le moins possible d'hypothèses *a priori* sur les propriétés locales ou globales de l'espace physique, c'est dans ce cadre qu'il faut se placer.

On peut d'ailleurs imaginer qu'imposer une métrique est encore une hypothèse trop forte, et considérer des variétés à structure plus faible : conforme, de Finsler, ou simplement munie de connexions (cf. Chapitre IX).

L'étude des propriétés de ces variétés constitue la géométrie différentielle. L'analyse tensorielle en est à la fois l'instrument et l'expression, et à son tour, la géométrie différentielle est l'instrument et l'expression de la relativité générale. Pour assurer à cet ouvrage un volume acceptable, il a fallu cependant résister à la tentation de trop s'engager dans ces deux domaines d'excellence de calcul tensoriel.

Plan d'étude

On vient de voir rapidement que l'utilisation des « tenseurs » est indispensable à la représentation de certaines grandeurs physiques. Mais de cette notion, nous n'avons encore que l'idée assez vague de quelque chose qui possède, dans chaque repère d'un espace donné, un ensemble de composantes. Nous devons donc, avant tout rechercher une définition précise. On souhaiterait, comme c'est le cas pour les vecteurs, un support intuitif, géométrique ou mécanique, permettant d'en véhiculer la compréhension et l'usage. Il faut se contenter de l'impression subjective que donne l'idée de contraintes, tensions, ou déformations d'un solide élastique ou de la courbure d'un espace de dimension $\geqslant 3$.

En fait, la nature des tenseurs est essentiellement algébrique, et c'est une intuition algébrique qu'il faut développer en ce qui les concerne (citons Hermann Weyl : « L'ange de la géométrie et le démon de l'algèbre luttent pour l'âme de chaque être mathématique »)

Nous dégagerons les notions de covariance et de contravariance en examinant les tenseurs élémentaires d'ordre un et deux au début du chapitre suivant.

BIBLIOGRAPHIE

[BA] A. BESSE, *Einstein manifolds* (Springer, 1987).

[B] BOURBAKI, *Algèbre multilinéaire* (Herrman).

[C] E.L. CORSON, *Introduction to tensors, spinors and relativistic wave equations* (Chelsea 1953).

[DR] R. DEHEUVELS, *Formes quadratiques et groupes classiques* (PUF 1981).

[D] P.A.M. DIRAC, *General theory of relativity* (Wiley 1975).

[E] L.P. EISEINHART, *Riemannian geometry* (Princeton 1925).

[GS] M. GÖCKELER & T. SCHÜCKER, *Differential geometry, gauge theories and gravity*, (Cambridge Uni. Press 1987).

[KN] S. KOBAYASHI S. & K. NOMIZU, *Foundations of diffential geometry* (Interscience – 1963).

[LD] D. LAWDEN, *An introduction to tensor calculus and relativity* (Chapman and Hall 1962).

[LB] H.B. LAWSON H.B. & M.L. MICHELSOHN, Spin Geometry (Princeton Univ. P. 1989).

[L1] A. LICHNEROWICZ, *Eléments de calcul tensoriel* (A. Colin 1950).

[MTW] C. MISNER K. THORNE, J.A. WHEELER, *Gravitation* (Freeman 1970).

[PS] S. PARROTT, *Relativistic electrodynamics and differential geometry* (Springer 1987).

[P] W. PAULI, *Theory of relativity* (Pergamon 1958).

[R] W. RINDLER, *Essential relativity* (Springer 1977).

[SS] J.L. SYNGE and A. SCHILD, *Tensor calculus* (University of Toronto Press 1956).

[S] J.L. SYNGE, *Relativity : the general theory* (North Holland 1958).

[WR] R. WALD, *General Relativity* (University of Chicago Press 1984).

[WS] S. WEINBERG, *Gravitation and cosmology* (Wiley 1972).

CHAPITRE II

Algèbre tensorielle : variance des composantes des tenseurs affines et produit tensoriel d'espaces vectoriels

Dans tout ce qui suit interviennent des « scalaires » c'est-à-dire des « nombres » qui, à l'exemple des nombres réels ou complexes, forment un *corps*. Rappelons qu'un corps est un ensemble muni de deux opérations : une addition pour laquelle c'est un groupe commutatif (donc avec un élément 0) et une multiplication pour laquelle les éléments non nuls forment un groupe, ayant pour élément neutre l'unité 1, la multiplication étant distributive par rapport à l'addition. *On supposera toujours la caractéristique du corps égale à zéro*, ce qui signifie que pour tout entier n non nul, $n \cdot 1 \neq 0$ donc inversible.

A partir du § II.2, et dans tout le reste de cet ouvrage, les corps de scalaires seront toujours supposés COMMUTATIFS.

Par convention, la sommation de termes relativement à des indices qui apparaissent en haut (contravariants) et en bas (covariants) sera indiquée *par le seul signe* $\sum$. Par exemple, au lieu de $\sum_{i,k} a_k^{ij} b_i^{lk}$, on écrira simplement $\sum a_k^{ij} b_i^{lk}$.

II.1 – Espaces vectoriels

Pendant longtemps, les vecteurs de la géométrie, de la mécanique ou de la physique (vitesses, forces, champs,...) ont été définis, dans notre espace ordinaire, comme des ensembles de trois composantes, car un système d'axes de référence était toujours présent et immuable.

Lorsqu'on sentait la nécessité d'une image intrinsèque du vecteur, indépendante des axes, le segment de droite terminé par un flèche suffisait à tous les besoins et à l'imagination. Cette image garde d'ailleurs encore et gardera toujours un rôle de support intuitif en algèbre linéaire.

Cependant, tous les raisonnements et calculs effectués sur les vecteurs dérivent de deux opérations.

1) l'addition, commutative, associative, avec existence d'un zéro et d'un opposé : $x+y=y+x$, $(x+y)+z=x+(y+z)$, $x+0=x$, $x+(-x)=0$.

2) la multiplication par les scalaires doublement distributive et associative : $(x+y)\lambda = x\lambda + y\lambda$, $x(\lambda+\mu) = x\lambda + x\mu$, $x(\lambda\mu) = (x\lambda)\mu$.

On a baptisé *espace vectoriel* sur un corps de scalaires K, tout ensemble muni de deux opérations du type précédent. (Si le corps des scalaires K est commutatif, on peut indifféremment noter $x\lambda$ ou λx la multiplication de x par λ). Un *vecteur* est alors simplement un élément d'un espace vectoriel, notion abstraite, et fort lointaine du segment de droite terminé par une flèche.

La notion d'espace vectoriel s'accompagne de celle d'*application linéaire* d'un espace vectoriel E dans un espace vectoriel F sur le même corps K : c'est une application f de E dans F qui respecte la structure : $f(x\lambda + y\mu) = f(x)\lambda + f(y)\mu$, quels que soient $x, y \in \mathrm{E}$, $\lambda, \mu \in \mathrm{K}$.

Des éléments $(x_i, i \in \mathrm{I})$ d'un espace vectoriel, où I désigne un ensemble d'indices quelconque (voire infini !) sont dits *indépendants* s'il n'existe entre eux aucune relation : une égalité telle que $x_{i1}\lambda_1 + x_{i2}\lambda_2 + \ldots + x_{ip}\lambda_p = 0$ entraîne nécessairement $\lambda_1 = \lambda_2 = \ldots = \lambda_p = 0$.

Une *base* d'un espace vectoriel E est un ensemble d'éléments indépendants maximal. Si $e = \{e_j ; j \in \mathrm{J}\}$ est une base, tout vecteur x de E s'écrit alors, puisque e est maximal, comme une combinaison linéaire finie de certains des vecteurs de e ; $x = \sum e_j x^j$ où x^j n'est différent de zéro que pour un nombre fini d'indices.

Cette écriture est unique puisque les e_j sont indépendants. Les x^j sont les composantes de x dans e.

Le cardinal de J dont on peut démontrer qu'il est le même pour toutes les bases de E, est la dimension de E.

Une application linéaire f de E dans F est entièrement déterminée par les images $f(e_j)$ des vecteurs d'une base $e = \{e_j; j \in \mathrm{J}\}$ de E. Réciproquement, à un choix arbitraire de vecteurs $\{f_j; j \in \mathrm{J}\}$ de F correspond une application linéaire f de E dans F telle que $f(e_j) = f_j$, $\forall j \in \mathrm{J}$.

Lorsque E et dimension finie n, une base e de E peut être écrite comme une matrice-ligne $e = \|e_1 e_2 \dots e_n\|$.

Les composantes d'un vecteur x dans e peuvent être écrites comme une matrice colonne X et l'écriture de x dans e est alors représentée comme un produit de matrices :

$$x = e\mathrm{X} = \|e_1 e_2 \dots e_n\| \begin{vmatrix} x^1 \\ x^2 \\ x^n \end{vmatrix} = e_1 x^1 + e_2 x^1 + \dots + e_n x^n.$$

Rappelons qu'une matrice $p \times q$: $\mathrm{A} = \|a_{ij}\|$ est un tableau rectangulaire à p lignes et q colonnes, où a_{ij} désigne l'élément qui se trouve dans la i^{e} ligne et la j^{e} colonne. La transposée ${}^t\mathrm{A}$ de A est la matrice à q lignes et p colonnes définies par : ${}^t\mathrm{A} = |a'_{ij}|$ avec $a'_{ij} = a_{ji}$.

Le produit d'une matrice $p \times q$: $\mathrm{A} = \|a_{ij}\|$ et d'une matrice $q \times r$: $\mathrm{B} = \|b_{jk}\|$, dont les éléments a et b peuvent être composés, est la matrice $\mathrm{C} = \|c_{ik}\|$ où $c_{ik} = \sum\limits_{j=1}^{q} a_{ij} b_{jk}$ est le produit terme à terme des éléments de la $i^{ème}$ ligne de A par ceux de la $k^{ième}$ colonne de B. Le produit de matrices est associatif, et on a : ${}^t(\mathrm{AB}) = {}^t\mathrm{B} \cdot {}^t\mathrm{A}$. Une matrice $n \times n$ A est dite inversible s'il existe une matrice, dite inverse de A, notée $\overset{-1}{\mathrm{A}}$, telle que $\overset{-1}{\mathrm{A}}\,\mathrm{A} = \mathrm{A}\,\overset{-1}{\mathrm{A}} = \mathrm{I}_n$: où I_n désigne la matrice $n \times n$:

$$\mathrm{I}_n = |\delta_{ij}| \quad \text{avec} \quad \delta_{ij} = 0 \text{ si } i \neq j,\ \delta_{ii} = 1,\ \forall i = 1, 2, \dots, n.$$

Si $e = \|e_1, e_2, \dots, e_n\|$ et $e' = \|e'_1, e'_2, \dots, e_n\|$ sont deux bases de l'espace vectoriel E, soit S la matrice $n \times n$ dont la j-ème colonne est celle des composantes $\{s^i_j; i = 1, 2, \dots, n\}$ du vecteur e'_j dans la base e (on écrit aussi s_{ij}). On a donc : $e' = e\mathrm{S}$ (produit de matrices). La matrice inversible S est la matrice de changement de bases : de la base e à la base e'. Si x est un vecteur de E, X et X$'$ les matrices colonnes de ses composantes dans e et e', on a :

$$x = e\mathrm{X} = e'\mathrm{X}' = e\mathrm{S}\mathrm{X}', \text{ d'où } \mathrm{X} = \mathrm{S}\mathrm{X}' \quad \text{et} \quad \mathrm{X}' = \overset{-1}{\mathrm{S}}\,\mathrm{X}.$$

Comparons le passage de e à e' et celui de X à X$'$. Dans le premier cas e et e' sont des matrices lignes dans le second des matrices colonnes.

Pour comparer des situations analogues, on va donc transposer les matrices X et X$'$.

Si $e' = e\mathrm{S}$, alors ${}^t\mathrm{X}' = {}^t\mathrm{X}\cdot{}^t\overset{-1}{\mathrm{S}} = {}^t\mathrm{X}\cdot\check{\mathrm{S}}$. La matrice $\check{\mathrm{S}} = {}^t(\overset{-1}{\mathrm{S}}) = ({}^t\overset{-1}{\mathrm{S}})$ est appelée la *contragrédiente* de la matrice carrée inversible S : $\mathrm{S} = \|s_j^k\|$.

On exprime ce fait en disant que les composantes d'un vecteur x subissent une transformation *contravariante* de celle de la base, ce qui est naturel puisque le produit $e\mathrm{X}$ reste fixe : $e\mathrm{X} = e'\mathrm{X}' = x$.

Une *forme linéaire* α sur un espace vectoriel E de dimension finie ou infinie est une application linéaire de E dans son corps K. On note $\langle\alpha, x\rangle$ la valeur de la forme α pour le vecteur x de E : $\langle\alpha, x\rangle$ est appelé le *crochet de dualité* de α et de x. Soit E* l'ensemble des formes linéaires sur E. Si $\alpha, \beta \in \mathrm{E}^*$ et $\lambda, \mu \in \mathrm{K}$, $\lambda\alpha + \mu\beta$ défini par $\langle\lambda\alpha + \mu\beta, x\rangle = \lambda\langle\alpha, x\rangle + \mu\langle\beta, x\rangle$ est une forme linéaire sur E : E* est un espace vectoriel appelé le *dual* de E.
$\langle\alpha, x\rangle$ est linéaire en x et en α, ce qui montre que tout vecteur x de E détermine « dualement » une forme linéaire sur E*, d'où $\mathrm{E} \subset (\mathrm{E}^*)^*$. On va voir que si la dimension de E est finie $\dim\mathrm{E} = \dim\mathrm{E}^* = \dim(\mathrm{E}^*)^*$ d'où $\mathrm{E} = (\mathrm{E}^*)^*$: E s'identifie à son bidual.

Dans une base e de E : $\langle\alpha, x\rangle = \langle\alpha, \sum\limits_j e_j x^j\rangle = \sum\limits_j \langle\alpha, e_j\rangle x^j$. La forme linéaire $x \in \mathrm{E} \to x^j \in \mathrm{K}$, $j^{ème}$ forme coordonnée, est notée e^{*j}, et α s'écrit donc :
$\alpha = \sum\limits_j \langle\alpha, e_j\rangle e^{*j} = \sum\limits_j \alpha_j e^{*j}$ ce qui prouve que $e^* = (e^{*1}, e^{*2}, \ldots, e^{*n})$ est une base de E*, dite duale de e. La matrice colonne X des composantes d'un vecteur x peut aussi être considérée comme la matrice colonne des formes e^{*j}.

On a $\langle e_j, e^{*k}\rangle = \delta_{k,j}$ = zéro si $j \neq k$, un si $j = k$. Puisque $\alpha_j = \langle \alpha, e_j\rangle$, e_j est la $j^{ème}$ forme coordonnée sur E^* relativement à la base e^* : e est la base duale de e^*, et $E \equiv (E^*)^*$.

Ecrivons les composantes de α dans e^* sous la forme d'une matrice ligne : $A = |\alpha_1\alpha_2 \ldots \alpha_n|$. $\langle \alpha, x\rangle$ s'écrit alors comme un produit de matrices : $\langle \alpha, X\rangle = \sum\limits_j \alpha_j x^j = AX$.

Si l'on passe de la base e à la base $e' = e\mathrm{S}$, on a : $\langle \alpha, x\rangle = AX = A'X' = ASX'$ d'où $A' = AS$. Les composantes d'une forme linéaire se tranforment comme la base de E : elles sont *covariantes* ce qui est naturel puisque AX est invariant. Si e^* est la matrice colonne des e^{*j}, puisque $X = SX'$, on a : $e^* = Se'^*$.

Lorsque E est de dimension finie, E et E$'$ jouent des rôles symétriques. Mais l'usage de l'expression « forme linéaire » n'est guère approprié pour en témoigner. C'est pourquoi *le calcul tensoriel préfère appeler « covecteurs » les éléments de l'espace dual* E^* ce que nous ferons le plus souvent possible. Si E et F sont deux espaces vectoriels sur le même corps commutatif K, l'ensemble des applications linéaires de E dans F est un espace vectoriel L(E; F) dont les opérations sont définies, si $a, b \in \mathcal{L}(E; F)$, $\lambda \in K$ et $x \in E$, par : $(a + b)(x) = a(x) + b(x)$; $(a\lambda)(x) = a(x\lambda) = a(x)\lambda$.

II.2 – Exemple : une première incursion en géométrie différentielle

Soient E un espace vectoriel *réel* de dimension finie n, f une fonction définie au voisinage d'un point x de E, à valeurs réelles, ou plus généralement à valeurs dans un espace vectoriel réel de dimension finie F. La dérivée de f en x suivant un vecteur u de E : $(D_u f)_x$, est la limite si elle existe.

$$(\mathrm{D}_u f)_x = \lim_{t\to 0} \frac{1}{t}\ \{f(x+tu) - f(x)\}.$$

On obtient immédiatement : $(\mathrm{D}_{\lambda u} f)_x = \lambda(\mathrm{D}_u f)_x$.

Si $(e_1, e_2, \ldots, e_n)$ est une base de E et si $(x^1, x^2, \ldots, x^n)$ sont les formes coordonnées correspondantes, on note habituellement :

$$(\mathrm{D}_{e_j} f)_x = \Big(\frac{\partial f}{\partial x^j}\Big)_x.$$

L'existence d'une dérivée suivant tout vecteur de E n'est pas très contraignante. Elle n'entraîne même pas la continuité de la fonction, ni l'existence de relations entre les dérivées D_u dans des directions différentes. La notion utile est celle de différentielle .f admet une *différentielle en* x : df_x, élément de E^* si f est à valeurs réelles, de $\mathcal{L}(E;F)$ si f est à valeur dans F, si l'on a pour tout vecteur u de E :

$$\underline{\overline{f(x+u) = f(x) + df_x(u) + \varepsilon(u)\|u\|}} \quad \text{avec} \quad \lim_{u\to 0}\varepsilon(u) = 0,$$

(où $\| \quad \|$ est une norme arbitraire sur E : elles sont toutes équivalentes).

f est dite *plate* en x si $df_x = 0$.

Si f est linéaire, elle coïncide avec sa différentielle en tout point : $df_x = f$ quel que soit $x \in E^*$. Il en est ainsi des formes linéaires $\alpha \in E^*$ et en particulier des formes linéaires coordonnées relatives à une base : $dx^j = x^j$. Or, la notation (x^j ; $j = 1, 2, \ldots, n$) étant généralement utilisée pour représenter les composantes d'un vecteur particulier x, c'est-à-dire les valeurs des fonctions : $x \to x^j$, la géométrie différentielle préfère la notation différentielle : $(dx^1, dx^2, \ldots, dx^n)$ pour représenter la base duale $(e^{*1}, e^{*2}, \ldots, e^{*n})$ de $(e_1, e_2, \ldots, e_n)$, et cet usage est d'ailleurs parfaitement adapté aux généralisations comme nous le verrons plus loin (Chap. V).

Lorsque f est différentiable en x, on a quel que soit $u \in E$, $(D_u f)_x = df_x(u)$ et $(D_u f)_x$ dépend alors linéairement de u. Ainsi

$$\begin{aligned}(d_u f)_x = df_x(u) = df_x(\sum_j e_j u^j) &= \sum_j u^j \Big(\frac{\partial f}{\partial x^j}\Big)_x \\ &= \sum_j \Big(\frac{\partial f}{\partial x^j}\Big) dx^j(u).\end{aligned}$$

On voit que chaque vecteur $u = \sum e_j u^j$ de E peut se représenter comme un opérateur de dérivation $\sum u^j \Big(\dfrac{\partial}{\partial x^j}\Big)_x$ et que la différentielle peut s'écrire : $df_x = \sum \dfrac{\partial f}{\partial x^j}\, dx^j$.

Nous allons réciproquement donner une définition directe, et intrinsèque des opérateurs de dérivation en un point x de E, qui seront appelés, en vue d'une généralisation ultérieure, vecteurs tangents en x :

Définition II.2 :

Un vecteur tangent en un point $x \in \mathrm{E}$ est une application X qui à chaque fonction numérique f différentiable en x, fait correspondre un nombre réel $\mathrm{X}f$, appelée dérivée de f par rapport à X, satisfaisant aux deux propriétés :

1) Linéarité : si $\lambda, \mu \in \Re$, $X \cdot (\lambda f + \mu g) = \lambda \mathrm{X} \cdot f + \mu \mathrm{X} \cdot g$.
2) X est nul sur les fonctions plates.

Exemple : Le vecteur tangent (ou vecteur vitesse) en un point x_0 d'un chemin différentiable : $t \to c(t) \in \mathrm{E}$, où $t \in [a, b] \subset \mathbb{R}$ et $x_0 = c(t_0)$, est le vecteur, noté $\left(\frac{dc}{dt}\right)_{x_0}$, défini pour toute fonction numérique f, différentiable au voisinage de x_0, par $\left(\frac{dc}{dt}\right)_{x_0} \cdot f = \left(\frac{df(c(t))}{dt}\right)_{t_0} = \lim_{h \to 0} \frac{1}{h} \left\{ f(c(t_0 + h)) - f(c(t_0)) \right\}$.

Les vecteurs tangents en x forment un espace vectoriel T_x appelé espace tangent en x. Chaque valeur u de E définit un vecteur tangent en x : $f \to df_x(u)$ et cette représentation est un isomorphisme de E sur T_x.

En effet, puisque la fonction $f - \sum \left(\frac{\partial f}{\partial x^j}\right)_x x^j$ est plate en x, l'action d'un vecteur tangent quelconque X en x peut s'écrire : $\mathrm{X} \cdot f = \sum_{,j} (\mathrm{X} \cdot x^j) \left(\frac{\partial f}{\partial x^j}\right)_x$. $\mathrm{X} = \sum (\mathrm{X} \cdot x^j) \left(\frac{\partial}{\partial x^j}\right)_x$ est donc une combinaison linéaire des $\left(\frac{\partial}{\partial x^j}\right)_x$, qui forment une base de T_x, image de la base (e_j) de E.

Nous avons ainsi atteint, le but de ce paragraphe qui était de représenter l'espace vectoriel réel (abstrait) E et son dual E^*, comme espaces vectoriels concrets *attachés à un point* x de E : T_x et son dual T^*_x. On a :

$$dx^i \left(\frac{\partial}{\partial x^j}\right) = \frac{\partial x^i}{\partial x^j} \doteq \delta^i_j.$$

Voyons comment, dans ces notations, s'exprime un changement de bases $e' = eS$, soit $e'j = \sum_j e_i s^i_j$.

On a $\mathrm{X} = \mathrm{SX}'$, soit $x^i = \sum_j s^i_j x'^j$ ou ${}^t\mathrm{X}' = {}^t\mathrm{X} \cdot \check{\mathrm{S}}$, soit $x'^j = \sum_j x^i \check{s}^j_i$. Il en résulte que :

$$s^i_j = \frac{\partial x^i}{\partial x'^j} \quad \text{et} \quad \check{s}^j_i = \frac{\partial x'^j}{\partial x^i}.$$

Les formules de changement de bases s'expriment donc de façon particulièrement simple, comme formules de calcul différentiel (formules de changement « de variables », les variables étant les fonctions coordonnées).

$$\frac{\partial}{\partial x'^j} = \sum \frac{\partial}{\partial x^i} \frac{\partial x^i}{\partial x'^j} \qquad \frac{\partial}{\partial x^i} = \sum \frac{\partial}{\partial x'^j} \frac{\partial x'^j}{\partial x^i},$$

$$dx'^j = \sum \frac{\partial x'^j}{\partial x^i}\, dx^i \qquad dx^i = \sum \frac{\partial x^i}{\partial x'^j}\, dx'^j.$$

Si f et g sont deux fonctions numériques différentiables en x, leur produit l'est aussi et $d(fg)_x = f(x)dg_x + g(x)df_x$.

Il en résulte que si X est un vecteur tangent en x : $\mathrm{X} \cdot fg = f(x)\mathrm{X}g + g(x)\mathrm{X}f$: X est une dérivation au sens algébrique. (Cf. §IV.13)

II.3 – Les tenseurs deux fois covariants sur un espace vectoriel de dimension finie

Les corps de scalaires sont désormais supposés *commutatifs* de caractéristique 0. Le *produit* de deux espaces vectoriels E, F sur le même corps K est l'ensemble $\mathrm{E} \times \mathrm{F}$ des couples (x, y) formés d'un vecteur x de E et d'un vecteur y de F, la *somme directe* de E et de F, notée $\mathrm{E} \oplus \mathrm{F}$, est l'espace vectoriel obtenu en munissant le produit $\mathrm{E} \times \mathrm{F}$ de l'addition :
$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ et de la multiplication par un scalaire λ définie par $\lambda(x, y) = (\lambda x, \lambda y)$. On définit de façon

analogue le produit et la somme directe de p espaces vectoriels sur un même corps.

Une forme bilinéaire f sur un espace vectoriel E est une fonction qui à tout couple (ordonné !) x, y de vecteurs de E, fait correspondre un scalaire $f(x, y) \in K$, dépendant linéairement de x et de y.

L'ensemble $\mathrm{T}_2(E)$ des formes bilinéaires sur E possède une structure d'espace vectoriel obtenue en définissant pour tous $f, g \in \mathrm{T}_2(E)$ et $\lambda \in \mathrm{K}$:

$$\{(f+g)(x,y) = f(x,y) + g(x,y), (\lambda f)(x,y) = \lambda \cdot f(x,y).$$

Un couple (α, β) de formes linéaires sur E définit une forme bilinéaire, appelée *produit tensoriel* de α et β et notée $\alpha \otimes \beta$, par :

$$(\alpha \otimes \beta)(x,y) = \alpha(x)\beta(y).$$

Le produit tensoriel est donc ici une application bilinéaire de $\mathrm{E}^* \times \mathrm{E}^*$ dans $\mathrm{T}_2(E)$. Son image est formée des formes bilinéaires décomposables en produit de formes linéaires. Elle engendre linéairement $\mathrm{T}_2(E)$ tout entier. En effet, si $e = (e_1, e_2, \ldots, e_n)$ est une base de E et $f \in \mathrm{T}_2(\mathrm{E})$:

$$\begin{aligned} f\Big(\sum e_i x^i, \sum e_j y^j\Big) &= \sum f(e_i, e_j) x^i y^j = \sum f_{ij} x^i y^j \\ &= \sum f_{ij} (e^{*i} \otimes e^{*j})(x,y) \text{ avec } f_{ij} = f(e_i, e_j). \end{aligned}$$

Ainsi, les n^2 produits tensoriels de formes coordonnées $e^{*i} \otimes e^{*j}$ engendrent $\mathrm{T}_2(\mathrm{E})$ et en forment une base puisque de $f = \sum f_{ij} e^{*i} \otimes e^{*j} = 0$, il résulte que $f(e_k, e_l) = f_{kl} = 0$ quels que soient k, l.

Les n^2 nombres f_{ij}, que l'on peut rassembler dans une matrice $\mathrm{F} = |f_{ij}|$, sont les composantes de f dans la base e et l'on peut écrire matriciellement $f(x,y) = {}^t\mathrm{XFY}$.

Nous allons déterminer la variance des f_{ij} c'est-à-dire leur comportement par changement de base. Si $e' = e\mathrm{S}$, soit $e'_k = \sum e_i s_{ik}$, en écrivant e^* et e'^* comme des matrices colonnes : $e^* = \mathrm{S}e'^*$ (cf. §II.1), soit $e^{*i} = \sum s_{ik} e'^{*k}$. En reportant dans l'expression de f :

$$\sum f_{ij} e^{*i} \otimes e^{*j} = \sum f_{ij} s_{ik} s_{jm} e'^{*k} \otimes e'^{*m} = \sum f'_{km} e'^{*k} \otimes e'^{*m}$$

d'où $f'_{km} = \sum_{i,j} f_{ij} s_{ik} s_{jm}$. Comparée à la formule de passage de e à e', cette relation conduit naturellement à dire que les composantes f_{ij} de $f \in \mathrm{T}_2(\mathrm{E})$ sont *deux fois covariantes*. On devait évidemment s'attendre à ce résultat puisque tout $f \in \mathrm{T}_2(\mathrm{E})$ est combinaison de produits de formes linéaires, et que les composantes d'une forme linéaire sont covariantes.

Définition II.3.A :

Les éléments de l'espace vectoriel $\mathrm{T}_2(\mathrm{E})$ que nous représenterons comme le produit tensoriel $\mathrm{E}^* \otimes \mathrm{E}^*$, de E^* par lui-même, sont appelés les *tenseurs deux fois covariants sur* E.
Supposons en effet qu'un objet associé à l'espace vectoriel E soit caractérisé dans chaque base $e = (e_1, e_2, \ldots, e_n)$ de E par n^2 composantes : $\{t_{ij}, 1 \leqslant i, j \leqslant n\}$ de telle sorte que par un changement de bases $e' = eS$, les composantes t'_{km} dans la base e' s'obtiennent à partir des t_{ij} par : $t'_{km} = \sum t_{ij} s^i_k s^j_m$. Autrement dit, les composantes de l'objet sont deux fois covariantes. Définissons un élément t de $\mathrm{T}_2(\mathrm{E})$ par $t = \sum t_{ij} e^{*j} \otimes e^{*j}$: t a alors les mêmes composantes que l'objet donné dans toutes les bases et peut donc le représenter.

Remarquons que le calcul de la relation entre les composantes de f dans deux bases e et $e' = e\mathrm{S}$ aurait pu s'effectuer matriciellement :

$$^t\mathrm{XFY} = {}^t(\mathrm{SX}')\mathrm{F}(\mathrm{SY}') = {}^t\mathrm{X}'{}^t\mathrm{SFSY}' = {}^t\mathrm{X}'\mathrm{F}'\mathrm{Y}'$$

d'où $\mathrm{F}' = {}^t\mathrm{SFS}$.

Mais l'usage des matrices ne sera plus possible dans l'étude des tenseurs d'ordre $\geqslant 3$, leurs composantes dépendant d'au moins trois indices ! Le rang de la matrice F, indépendant de la base, est appelé le rang de f. Les éléments de rang un sont les éléments décomposables, de la forme $\alpha \otimes \beta$, $\alpha, \beta \in \mathrm{E}^*$.

Si f est de rang maximum $n = \dim \mathrm{E}$, f est dit non-dégénéré.

A chaque vecteur y de E la forme bilinéaire f fait correspondre la forme linéaire sur E : $x \longrightarrow f(x, y)$ et définit ainsi une application linéaire ρ de E dans son dual E^*, dite *associée à droite de* f par : $f(x, y) = \langle x, \rho(y) \rangle$.

L'application γ de E dans E^*, *associée à gauche de* f est définie par : $f(x, y) = \langle \gamma(x), y \rangle$. Le bidual E^{**} étant identifié à E, on a : $\gamma = {}^t\rho$ et $\rho = {}^t\gamma$. Réciproquement, ρ ou γ détermine f. Si $f = \sum f_{ij} e^{*i} \otimes e^{*j}$, $\rho(y) = \sum f_{ij} e^{*i} y^j$ et $\gamma(x) = \sum f_{ij} x^i e^{*j}$. La matrice de ρ relativement à e est $\mathrm{F} = |f_{ij}|$. Celle de γ est ${}^t\mathrm{F}$. On a : rang f = rang ρ = rang γ.

Remarque : Si e et $e' = e\mathrm{S}$ sont deux bases de E, $\text{dét}\,\mathrm{F}' = (\text{dét}\,\mathrm{S})^2 \text{dét}\,\mathrm{F}$. Le nombre qui dans chaque base e de E est égal à dét F n'est pas un scalaire. C'est un pseudoscalaire suivant la :

Définition II.3.B :

Une fonction qui à chaque base e d'un espace vectoriel E associe un scalaire $u(e) \in \mathrm{K}$, de telle sorte que par un changement de bases $e' = e\mathrm{S}$, on aie $u(e') = u(eS) = (\text{dét}\,\mathrm{S})^p u(e)$, où p est un entier positif ou négatif, est appelée un *pseudoscalaire de variance scalaire p*.
Tout tenseur deux fois covariant non-dégénéré définit ainsi un pseudoscalaire de variance scalaire +2.

La transposée τf de $f \in \mathrm{T}_2(\mathrm{E})$ est définie par $\tau f(x,y) = f(x,y)$. Dans toute base la matrice des composantes de τf est évidemment la transposée de celle de f.
τ est un automorphisme involutif (τ^2 = identité) de l'espace vectoriel $\mathrm{T}_2(E)$. f est symétrique si $\tau f = f$ et antisymétrique si $\tau f = -f$. Toute forme bilinéaire f est somme de sa partie symétrique $\frac{1}{2}\,(f+\tau f)$ et de sa partie antisymétrique $\frac{1}{2}\,(f-\tau f)$. Pour que f soit symétrique (resp. antisymétrique) il est nécessaire et suffisant que la matrice de ses composantes dans une base particulière soit symétrique (resp. antisymétrique) et il en est alors ainsi dans toute base.
Une *forme quadratique* sur un espace vectoriel E est une fonction q sur E, à valeurs dans son corps des scalaires K, qui satisfait aux deux axiomes :

1) $q(\lambda x) = \lambda^2 q(x)$, $\forall x \in \mathrm{E}$, $\lambda \in \mathrm{K}$.

2) la fonction $q(x+y) - q(x) - q(y)$ est bilinéaire.

Toute forme bilinéaire f sur E définit une forme quadratique par $q(x) = f(x,x)$. La fonction correspondante $q(x+y) - q(x) - q(y) = f(x+y, x+y) - f(x,x) - f(y,y) = f(x,y) + f(y,x)$ est égale à $(f + \tau f)$, le double de partie symétrique de f. Pour cette raison, on préfère associer à la forme quadratique q la forme bilinéaire b définie par

$$\frac{1}{2}\,\{q(x+y)-q(x)-q(y)\}.\ \text{On a alors : } q(x)=b(x,x).$$

Une forme quadratique que E peut donc être considérée comme un tenseur deux fois covariant symétrique.

Lorsqu'une forme bilinéaire symétrique b sur E est non-dégénérée, on dit qu'elle définit sur R un *produit scalaire*. Lorsqu'une forme bilinéaire antisymétrique a sur E est non-dégénérée, on dit qu'elle définit sur E un *produit scalaire symplectique*, et dans ce cas la dimension de E est nécessairement paire.

Notation de la géométrie différentielle : une forme bilinéaire t sur l'espace vectoriel T_x se note $t = \sum t_{ij} dx^i dx^j$. La notation classique pour un produit scalaire sur T_x est $g = \sum g_{ij} dx^i dx^j$ avec $g_{ij} = g_{ji}$ et dét$\,|g_{ij}| \neq 0$.

Si t est antisymétrique, $t_{ij} = -t_{ji}$ et t s'écrit : $t = \sum\limits_{i<j} t_{ij}(dx^i dx^j - dx^j dx^i)$ avec une sommation restreinte aux couples d'indices $(i<j)$.

Si x et y sont deux vecteurs de T_x, on a

$$(dx^i dx^j - dx^j dx^i)(X,Y) = X^i Y^j - X^j Y^i = \text{dét}\begin{vmatrix} X^i & Y^i \\ X^j & Y^j \end{vmatrix}.$$

On convient de noter $dx^i \wedge dx^j$ la forme bilinéaire antisymétrique : $dx^i \wedge dx^j = dx^i dx^j - dx^i dx^i$, d'où pour une forme bilinéaire antisymétrique t :

$$t = \sum_{i<j} t_{ij} dx^i \wedge dx^j.$$

II.4 – Les tenseurs une fois covariant, une fois contravariant sur un espace vectoriel de dimension finie

Un opérateur linéaire a de E est une application linéaire de E dans lui-même. Son *image* est le sous-espace vectoriel de E que forme l'ensemble des images $a(x)$, $x \in$ E, son *noyau* est le sous-espace des vecteurs annulés par a. L'ensemble $\mathfrak{L}$(E) des opérateurs linéaires de E possède une structure naturelle d'espace vectoriel :

si $a, b \in \mathfrak{L}(\mathrm{E})$, $\lambda, \mu \in K$, $\lambda a + \mu b$ défini par $(\lambda a + \mu b)(x) = \lambda a(x) + \mu b(x)$ est un opérateur linéaire. $\mathfrak{L}(\mathrm{E})$ possède même une structure plus riche. En effet, si $a, b \in \mathfrak{L}(\mathrm{E})$, le composé ab défini par $(ab)(x) = a(b(x))$ est un opérateur linéaire.

Cette multiplication, associative et distributive par rapport à l'addition, définit sur $\mathfrak{L}(\mathrm{E})$ une structure d'algèbre, à laquelle nous ne nous intéresserons que plus tard.

a est dit inversible s'il existe un opérateur, noté $\overset{-1}{a}$, tel que $a\overset{-1}{a} = \mathrm{I}$. Pour que a soit inversible, il suffit que l'image par a d'une base particulière soit encore une base, et c'est alors vrai de toute base.

Dans une base e de E, $a \in \mathfrak{L}(\mathrm{E})$ possède n^2 composantes qui sont les éléments de sa matrice $A = |a^i_j|$. Par définition a^i_j (indices décalés, le premier étant toujours celui de la ligne, le second celui de la colonne) est la $i^{ème}$ composante de l'image par a du $j^{ème}$ vecteur de base :
$a^i_j = a(e_j)^i = \langle e^{*i}, a(e_j)\rangle$. La $j^{ème}$ colonne de A est donc formée des composantes de $a(e_j)$. Dans la base e :

$$y = a(x) = a\Big(\sum_j e_j x^j\Big) = \sum_j a(e_j)x^j \text{ d'où } y^i = \sum_j a^i_j x^j.$$

ce qui s'écrit matriciellement : $\mathrm{Y} = \mathrm{AX}$.

Un couple (y, β) formé d'un vecteur $y \in \mathrm{E}$ et d'une forme linéaire $\beta \in \mathrm{E}^*$ définit un opérateur linéaire $y \otimes \beta$, appelé *produit tensoriel* de y et de β, par : $((y \otimes \beta)(x) = y\langle\beta, x\rangle$.

L'image de $y \otimes \beta$ est la droite portant y, son noyau est le noyau de β; si y et β sont non nuls, $y \otimes \beta$ est un opérateur de rang un et ses composantes dans une base e de E sont les produits $y^i\beta_j$ des composantes de y et de β.

Ce produit tensoriel est une application bilinéaire de $\mathrm{E} \times \mathrm{E}^*$ dans $\mathfrak{L}(\mathrm{E})$ dont l'image est évidemment formée des opérateurs de rang un, ou décomposables. Cette image engendre linéairement $\mathfrak{L}(\mathrm{E})$ tout entier. En effet si $e = (e_1, e_2, \ldots, e_n)$ est une base de E, e^* la base duale de E^*, on a pour tout $a \in \mathfrak{L}(\mathrm{E})$:

$$\begin{aligned} a(\textstyle\sum e_j x^j) &= \textstyle\sum a(e_j)x^j = \sum a(e_j)\langle e^{*j}, x\rangle \\ &= \textstyle\sum e_i a^i_j \langle e^{*j}, x\rangle = \sum a^i_j (e_i \otimes e^{*j})(x) \end{aligned}$$

soit : $$a = \sum a(e_j) \otimes e^{*j} = \sum a^i_j e_i \otimes e^{*j}.$$

Les n^2 produits tensoriel $e_i \otimes e^{*j}$ engendrent $\mathcal{L}(\mathrm{E})$ et en forment une base puisqu'une relation $\sum a^i_j e_i \otimes e^{*j} = 0$ signifie que l'opérateur de matrice a^i_j est l'opérateur nul, d'où $a^i_j = 0\ \forall i, j$.

Déterminons la variance des composantes a^i_j de a. Si $e' = e\mathrm{S}$, $e^* = \mathrm{S}e'^*$, d'où ${}^te = {}^t\overset{-1}{\mathrm{S}}\,{}^te' = \check{\mathrm{S}}{}^te'$, soient :

$$a = \sum a^i_j e_i \otimes e^{*j} = \sum a^i_j \check{s}^k_i s^j_m e'_k \otimes e'^{*m}$$
$$= \sum a'^k_m e'_k \otimes e'^{*m}.$$

d'où $$a'^k_m = \sum a^i_j \check{s}^k_i s^j_m$$

ce qui permet de dire que les composantes a^i_j de $a \in \mathcal{L}(\mathrm{E})$ sont *contravariantes dans le premier indice i, et covariantes dans le second indice j*. Leurs places respectives, en haut pour i, en base pour j rappellent cette variance. Matriciellement, $\mathrm{Y} = \mathrm{AX}$ s'écrit $\mathrm{SY}' = \mathrm{ASX}'$ d'où $A' = \overset{-1}{\mathrm{S}}\,\mathrm{AS}$.

Remarque : L'opérateur identique $\mathrm{I} = \sum\limits_i e_i \otimes e^{*i}$ a les mêmes composantes : $\delta^i_i = 1, \forall i$, $\delta^i_j = 0$ si $i \neq j$ dans toutes les bases, et la relation $\sum\limits_i s_{ik} s_{im} = \delta^k_m$ exprime simplement que ${}^t\check{\mathrm{S}} \cdot \mathrm{S} = \overset{-1}{\mathrm{S}}\,\mathrm{S} = \mathrm{I}$.

On peut également considérer les opérateurs linéaires de E^* qui forment l'espace vectoriel $\mathcal{L}(\mathrm{E}^*)$. A un couple (β, y), $\beta \in \mathrm{E}^*$, $y \in \mathrm{E}$ correspond leur *produit tensoriel* $\beta \otimes y \in \mathcal{L}(\mathrm{E}^*)$ par $\beta \otimes y(\alpha) = \beta\langle y, \alpha\rangle$, d'où une application bilinéaire de $\mathrm{E}^* \times \mathrm{E}$ dans $\mathcal{L}(\mathrm{E}^*)$. Si $b \in \mathcal{L}(\mathrm{E}^*)$, on a dans une base e de E et la base duale e^* de E^* :

$$b\Big(\sum \alpha_j e^{*j}\Big) = \sum \alpha_j b(e^{*j}) = \sum \langle e_j, \alpha\rangle b^j_i e^{*i} = \sum b^j_i e^{*i} \otimes e_j(\alpha)$$

où le premier indice i de b^j_i est celui de la ligne et le second j celui de la colonne. D'où : $b = \sum b^j_i e^{*i} \otimes e_j$. Par exemple les composantes du produit tensoriel $\beta \otimes y = \sum \beta_i y^j e^{*i} \otimes e_j$ sont les produits $\beta_i y^j$ de celles de β et de y. Un calcul analogue au précédent donne l'expression des composantes $b'^m_k = \sum b^j_i s_{ik} s_{jm}$.

ce qui permet de dire que les composantes b_i^j de $b \in \mathcal{L}(\mathrm{E}^*)$ sont *covariantes dans le premier indice i et contravariantes dans le second indice j.*

A tout opérateur linéaire $a \in \mathcal{L}(\mathrm{E})$ correspond son transposé ${}^t a \in \mathcal{L}(\mathrm{E}^*)$ et à tout opérateur linéaire $b \in \mathcal{L}(\mathrm{E}^*)$ correspond son transposé ${}^t b \in \mathcal{L}(\mathrm{E}^{**}) \equiv \mathcal{L}(\mathrm{E})$ au moyen des relations vérifiées quels que soient $x \in \mathrm{E}$ et $\alpha \in \mathrm{E}^*$:

$$\langle ax, \alpha\rangle = \langle x, {}^t a \cdot \alpha\rangle,$$
$$\langle x, b\alpha\rangle = \langle {}^t bx, \alpha\rangle \text{ d'où } {}^t({}^t a) = a,$$

que l'on peut aussi écrire en privilégiant le rôle de formes linéaires des éléments de E^* : $\alpha(ax) = ({}^t a \cdot \alpha)(x)$ et $(b\alpha)(x) = \alpha({}^t bx)$.

La transposition est un isomorphisme canonique des espaces vectoriels $\mathcal{L}(\mathrm{E})$ et $\mathcal{L}(\mathrm{E}^*)$ mais un antiisomorphisme pour les structures d'algèbres : ${}^t(u \circ v) = {}^t v \circ {}^t u$.

Calculons le transposé du produit tensoriel $y \otimes \beta \in \mathcal{L}(\mathrm{E})$. On a : $\langle y \otimes \beta \cdot x, \alpha\rangle = \langle x, \beta\rangle\langle y, \alpha\rangle = \langle x, \beta \otimes y \cdot \alpha\rangle$ soit ${}^t(y \otimes \beta) = \beta \otimes y$, et de même ${}^t(\beta \otimes y) = y \otimes \beta$.

En termes de produits tensoriels, la transposition est donc l'échange des facteurs. Si $a \in \mathcal{L}(\mathrm{E})$,

$$a = \sum a_j^i e_i \otimes e^{*j} \text{ et } {}^t a = \sum a_j^i e^{*j} \otimes e_i = \sum ({}^t a)_j^i e^{*j} \otimes e_i.$$

La matrice $|({}^t a)_j^i|$ du transposé est donc la transposée de la matrice $|a_j^i|$ (Rappelons que le premier indice est celui de la ligne et le second celui de la colonne).

Cette étude montre que l'on peut considérer $\mathcal{L}(\mathrm{E})$ comme l'espace vectoriel $\mathrm{T}_1^1(\mathrm{E})$ des tenseurs associés à E, contravariants dans le premier indice, covariants dans le second, $\mathcal{L}(\mathrm{E}^*)$ comme l'espace vectoriel $\mathrm{T}_1^1(\mathrm{E})$ des tenseurs associés à E, covariants dans le premier indice, contravariants dans le second.

$\mathrm{T}_1^1(\mathrm{E})$ peut être considéré comme le produit tensoriel $\mathrm{E} \otimes \mathrm{E}^*$ tout élément a de $\mathrm{T}_1^1(\mathrm{E})$ pouvant s'écrire comme une somme $\sum x_j \otimes \alpha_j$, et $\mathrm{T}_1^1(\mathrm{E})$ comme le produit tensoriel $\mathrm{E}^* \otimes \mathrm{E}$.

Ces deux espaces sont canoniquement isomorphes et l'on peut souvent les identifier. Par contre, rappelons que l'ordre des indices dans $\mathrm{T}^2(\mathrm{E}) = \mathrm{E}^* \otimes \mathrm{E}^*$ est strictement fixé !

Exemple : Tenseur des contraintes en un point d'un solide (stress tensor) : Pour analyser les tensions intérieures d'un solide soumis à des contraintes diverses (pression, torsion, étirement, etc...), on fait correspondre à un élément de surface orienté $d\mathrm{S}$, de vecteur normal unitaire $\overrightarrow{n}$, donc représenté par le vecteur $\overrightarrow{n}\, d\mathrm{S}$, la force $\overrightarrow{\mathrm{F}}\, d\mathrm{S}$ qu'exerce sur elle la partie du solide située du côté positif. On a ainsi un opérateur linéaire T défini par : $\overrightarrow{\mathrm{F}} = \mathrm{T}(\overrightarrow{n})$, qui est le tenseur des contraintes au point considéré. ■

Le déterminant de la matrice A d'un opérateur linéaire a dans une base e est indépendant de la base : c'est un scalaire, contrairement à ce qui se passe pour les formes bilinéaires. En effet dans deux bases e et $e' = eS$, les matrices de a sont reliées par $\mathrm{A}' = \overset{-1}{\mathrm{S}}\,\mathrm{AS}$ d'où $\det \mathrm{A}' = \det \mathrm{A}$.

La *trace* $\mathrm{Tr}\, a$ est un autre « invariant » d'un opérateur a (nous étudierons les invariants aux §III.2, III.14 et IV.13.

On sait que l'on appelle trace d'une matrice carrée $\mathrm{A} = |a_{ij}|$ la somme des éléments de sa diagonale principale : $\mathrm{Tr\,A} = \sum\limits_i a_{ii}$. Si $\mathrm{A} = |a_{ij}|$ est une matrice $n \times p$ et $\mathrm{B} = |b_{jk}|$ une matrice $p \times n$, AB et BA sont des matrices carrées, respectivement $n \times n$ et $p \times p$, et elles ont même trace car :

$$\mathrm{Tr\,AB} = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{ij} b_{ji} = \sum_{j=1}^{p} \sum_{i=1}^{n} b_{ji} a_{ij} = \mathrm{Tr\ BA}.$$

Il en résulte que l'on peut définir la trace d'un opérateur linéaire a comme étant la trace de sa matrice dans une base quelconque : $\mathrm{Tr}\, a = \mathrm{Tr\,A}$. En effet, si A et A$'$ sont les matrices de a dans les bases e et $e' = e\mathrm{S}$, on a $\mathrm{Tr\,A}' = \mathrm{Tr}(\overset{-1}{\mathrm{S}}\,\mathrm{AS}) = \mathrm{Tr}(\overset{-1}{\mathrm{S}}\,\mathrm{A})\mathrm{S} = \mathrm{Tr\,S}(\overset{-1}{\mathrm{S}}\,\mathrm{A}) = \mathrm{Tr\,A}$.

Cette façon de définir la trace est certainement assez artificielle. La cause profonde de son existence n'est pas cette simple propriété algébrique $\mathrm{Tr\,AB} = \mathrm{Tr\,BA}$ que nous avons utilisée mais une opération de nature tensorielle appelée la *contraction*.

Elle se manifeste de la façon suivante dans le cas des opérations linéaires : il existe une application linéaire naturelle du produit tensoriel $\mathrm{E} \otimes \mathrm{E}^*$, comme du produit tensoriel $\mathrm{E}^* \otimes \mathrm{E}$, à valeurs dans le corps K des scalaires, qui, au produit tensoriel $y \otimes \alpha$, ou

$\alpha \otimes y$, avec $y \in \mathrm{E}$ et $\alpha \in \mathrm{E}^*$ fait correspondre le crochet $\langle \alpha, y \rangle$ de α et de y, « contraction » de α et de y.

Un opérateur linéaire a, considéré comme élément de $\mathrm{E} \otimes \mathrm{E}^*$ peut s'écrire d'une infinité de façons différentes comme combinaison linéaire de produits tensoriels :
$a = \sum a^i_j e_i \otimes e^{*j}$ dans chaque base e, mais aussi, par exemple : $a = \sum a(e_j) \otimes e^{*j}$. Son image par la contraction est indépendante de son écriture, c'est sa trace :

$$\mathrm{Tr}\, a = \sum a^i_j \langle e_i, e^{*j} \rangle = \sum a^j_j = \sum \langle a(e_j), e^{*j} \rangle.$$

L'existence de la contraction sera *démontrée* ci-dessous.

II.5 – Intervention des groupes. Action du groupe linéaire

Les groupes jouent un rôle essentiel en calcul tensoriel. Ils participent de la nature même des tenseurs et des pseudotenseurs. Rappelons qu'un groupe G est un ensemble muni d'une loi de composition interne associative, avec élément neutre et existence de l'inverse. Si E est un espace vectoriel, l'ensemble des opérateurs linéaires inversibles de E forme un groupe, le groupe linéaire $\mathrm{Gl(E)}$. Cela signifie que : 1) si $a \in \mathrm{Gl(E)}$, $\overset{-1}{a} \in \mathrm{Gl(E)}$; 2) l'identité $I \in \mathrm{Gl(E)}$; 3) si $a, b \in \mathrm{Gl(E)}$, $ab \in \mathrm{Gl(E)}$. On a en effet $ab\, \overset{-1}{b}\, \overset{-1}{a} = I$, et $(a\, \overset{-1}{b}) = \overset{-1}{b} \cdot \overset{-1}{a}$.

Définition II.5 :

Une *représentation linéaire* d'un groupe G dans un espace vectoriel E est un homomorphisme ρ de G dans le groupe $\mathrm{Gl(E)}$, ce qui signifie que $\rho(gg') = \rho(g)\rho(g')$, $\forall g, g' \in \mathrm{G}$ et que ρ (élément neutre de G) = opérateur identique de E.

Muni d'une telle représentation E prend le nom de G-*module*. Un *invariant* d'un G-module E est un élément x de E qui reste fixe par les opérations de G : $\rho(g) \cdot x = x$, $\forall g \in \mathrm{G}$.

On a vu que l'application qui à chaque opérateur linéaire $a \in \mathrm{E}$ fait correspondre l'opérateur transposé ${}^t a$ de E^* est un antiautomorphisme de l'algèbre $\mathfrak{L}(\mathrm{E})$ sur l'algèbre $\mathfrak{L}(\mathrm{E}^*)$. Mais l'application qui à chaque $u \in \mathrm{Gl(E)}$ fait correspondre l'opérateur contragrédient : $\check{u} = {}^t\overset{-1}{u}$ de $\mathrm{Gl(E^*)}$ est un isomorphisme du groupe $\mathrm{Gl(E)}$ sur le groupe $\mathrm{Gl(E^*)}$ puisque

$$(\check{uv}) = {}^t(\overset{-1}{uv}) = {}^t\left(\overset{-1}{v}\,\overset{-1}{u}\right) = {}^t\overset{-1}{u}\,{}^t\overset{-1}{v} = \check{u}\cdot\check{v}.$$

La représentation de $\mathrm{Gl(E)}$ comme groupe linéaire de E^* est appelée la représentation contragrédiente. Elle fait opérer $\mathrm{Gl(E)}$ sur les tenseurs une fois covariant que sont les formes linéaires sur E.

Il faut remarquer que $\mathrm{Gl(E)}$ opère de façon directe sur les vecteurs de E, et de façon contragrédiente sur leurs composantes. Historiquement, les vecteurs et tenseurs étant toujours intervenus par leurs composantes, les vecteurs de E se sont trouvés baptisés contravariants. De façon duale $\mathrm{Gl(E)}$ opère de façon contragrédiente sur les formes linéaires, éléments de E^*, mais de façon directe sur leurs composantes. Les formes linéaires sont ainsi baptisées covariantes. Il faut s'habituer à cette terminologie un peu paradoxale consacrée par l'usage.

L'image $g\cdot\alpha$ d'une forme linéaire par $g \in \mathrm{Gl(E)}$ est la forme linéaire telle que sa valeur sur le vecteur gx soit la même que la valeur de α sur x : le crochet de dualité reste inchangé.

$$(g\cdot\alpha)(gx) = \langle g\cdot\alpha, gx\rangle = \langle {}^t\overset{-1}{g}\,\alpha, gx\rangle = \langle \alpha, \overset{-1}{g}\,gx\rangle = \langle\alpha, x\rangle.$$

D'une façon analogue, l'image $g\cdot a$ d'un opérateur linéaire par g est l'opérateur tel que son action sur le vecteur gx soit l'image par g de l'action de a sur x : $(g\cdot a)(gx) = g(a(x))$, soit en remplaçant x par $\overset{-1}{g}\,x$: $(g\cdot a)(x) = g(a(\overset{-1}{g}\,x))$ et $g\cdot a = ga\,\overset{-1}{g}$.

Remarque : Dire que le déterminant et la trace sont des invariants des opérateurs signifie que le déterminant et la trace sont invariants par les opérations de $\mathrm{Gl(E)}$: $\mathrm{dét}(g\cdot a) = \mathrm{dét}(ga\,\overset{-1}{g}) = \mathrm{dét}\,a$ et $\mathrm{Tr}(g\cdot a) = \mathrm{Tr}(ga\,\overset{-1}{g}) = \mathrm{Tr}\,a$, donc par les changements de base.

L'image $(g\cdot f)$ d'une forme bilinéaire f est la forme bilinéaire telle que sa valeur sur les vecteurs gx, gy soit la même que la valeur de f sur x, y : $(g\cdot f)(gx, gy) = f(x, y)$, d'où :

$$(g\cdot f)(x, y) = f(\overset{-1}{g}\,x, \overset{-1}{g}\,y) \text{ et } (g\cdot f) = f\circ(\overset{-1}{g}, \overset{-1}{g}).$$

Remarque : Sans avoir encore défini de façon précise les tenseurs, nous savons qu'ils possèdent des composantes dans chaque base.

Nous pouvons par analogie avec ce qui précède, définir l'image $(g \cdot t)$ d'un tenseur t par $g \in \mathrm{Gl(E)}$ comme le tenseur qui, dans la base $g \cdot e$ possède les mêmes composantes que t dans la base e. Les composantes de $(g \cdot t)$ dans la base e sont donc obtenues en effectuant sur les composantes de t la tranformation associée au changement de base e en $\overset{-1}{g}\, e$.

II.6 – Définition des tenseurs affines sur un espace vectoriel de dimension finie par l'effet des changements de base sur leurs composantes

Les espaces des deux paragraphes précédents, qui ont été identifiés à des espaces tensoriels, possèdaient déjà une définition directe, indépendante.

Par contre, il est curieux de constater que l'algèbre linéaire et la géométrie ne fournissent aucun objet naturel identifiable aux tenseurs deux fois contravariants !

Nous allons commencer par la définition des tenseurs qui est historiquement la première. Elle présente une indiscutable économie de moyens, mais elle reste superficielle, ignorant délibérément toutes les questions concernant leur nature intrinsèque, leur construction systématique et l'origine des diverses structures qu'ils présentent.

De plus, il ne s'agit que de tenseurs *affines* c'est-à-dire associés à un espace vectoriel nu.

Définition II.6.A des tenseurs affines « par leurs composantes » :

Soit E un espace vectoriel de dimension finie n. Soient p et q deux entiers $\geqslant 0$ et $m = p + q$. On appelle *tenseur sur* E, d'ordre m, p fois contravariant et q covariant, soit, en abrégé, de *variance* (p, q), la donnée pour chaque base e de E d'un ensemble de n^m composantes $t^{i_1 i_2 \dots i_p}_{j_1 j_2 \dots j_q}$ comportant p indices en haut, q indices en bas, telles que, si e' est une autre base de E, qui se déduit de e

par la matrice $\mathrm{S} = |s^i_k|$ de changement de bases : $e' = e\mathrm{S}$, les composantes de t dans e et e' soient liées par la relation :

$$t'^{i_1\ldots i_p}_{j_1\ldots j_q} = \sum t^{k_1\ldots k_p}_{l_1\ldots l_q} \check{s}^{i_1}_{k_1} \check{s}^{i_2}_{k_2} \ldots s^{l_1}{}_{j_1} s^{l_2}{}_{j_2},$$

où $\check{S} = {}^t\overset{-1}{\mathrm{S}} = |\check{s}^l_j|$ est la matrice contragrédiente de S.

Remarques et exemples :

La définition précédente est « individuelle ». Elle s'intéresse à un tenseur à la fois, et constitue de toutes façons un *critère pratique* pour décider si des composantes sont, ou non, celles d'un tenseur.

Les cas particuliers les plus simples ont été vus ci-dessus, mais sous la forme d'espaces vectoriels de tenseurs, avec les vecteurs $(p = 1, q = 0)$, les formes linéaires $(p = 0, q = 1)$ les formes bilinéaires $(p = 0, q = 2)$, les opérateurs linéaires $(p = q = 1)$.

Remarquons cependant que dans ce dernier cas, on avait fait la distinction entre les opérateurs linéaires de E et ceux de E^*, alors que dans la définition ci-dessus, il n'est question que de composantes t^i_j, sans que l'on se préoccupe ni de leur interprétation, ni de savoir si l'indice du haut précède ou non celui du bas, ce qui n'a pas de sens dans cette définition.

On peut construire immédiatement des exemples de tenseurs d'ordre quelconque. En effet, on vérifie immédiatement que p vecteurs de E : $x_{(1)}, x_{(2)}, \ldots, x_{(p)}$ et q formes linéaires sur E : $\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(q)}$ définissent un tenseur t, p fois contravariant et q fois covariant, dit tenseur *décomposable*, en posant :

$$t^{l_1 l_2\ldots l_p}_{j_1 j_2\ldots j_q} = x^{i_1}_{(1)} x^{i_2}_{(2)} \ldots x^{i_p}_{(p)} \alpha_{(1)j_1} \alpha_{(2)j_2} \ldots \alpha_{(p)j_p}.$$

Ce tenseur sera ultérieurement appelé *produit tensoriel* des p vecteurs et des q formes (en respectant l'ordre des vecteurs et celui de formes), et noté en conséquence : $t = x_{(1)} \otimes \ldots x_{(p)} \otimes \alpha_{(1)} \otimes \ldots \alpha_{(q)}$.

Nous allons maintenant définir des opérations extrêmement importantes sur les tenseurs : les *contractions* qui généralisent la trace. Elles ont une signification intrinsèque et ont la même définition, quelle que soit la base !

Proposition II.6 :

Soit t un tenseur p fois contravariant et q fois covariant avec $p \geqslant 1$, $q \geqslant 1$, sur un espace vectoriel E de dimension n.

Choisissons un indice contravariant, le b-ème $(1 \leqslant b \leqslant p)$ et un indice covariant, le c-ème $(1 \leqslant c \leqslant q)$. Définissons dans chaque base e de E les composantes :

$$u^{i_1 i_2 \ldots i_{b-1} i_{b+1} \ldots i_p}_{j_1 j_2 \ldots j_{c-1} j_{c+1} \ldots j_q} = \sum_{a=1}^{n} t^{i_1 \ldots i_{b-1} a i_{b+1} \ldots i_p}_{j_1 \ldots j_{c-1} a j_{c+1} \ldots j_q}.$$

Ce sont celles d'un tenseur u, $(p-1)$ fois contravariant. $(q-1)$ fois covariant (l'ordre de u : $p+q-2$ est donc de deux unités inférieur à celui de t) appelé *contracté* de t dans le b-ème et le c-ème indices.

Preuve : Il suffit, dans la formule du changement de base de la définition ci-dessus, de faire la sommation :

$$\sum_{a=1}^{n} t'^{i_1 \ldots i_{b-1} a i_{b+1} \ldots i_p}_{j_1 \ldots j_{c-1} a j_{c+1} \ldots j_q}.$$

et de remarquer dans le second membre que $\sum_{a=1}^{n} s^a_{k_b} \check{s}^{l_c}_a = 0$ si $k_b \neq l_c$ et $= 1$ si $k_b = l_c$ puisque le produit de matrices $\mathrm{S} \cdot {}^t\check{\mathrm{S}} = \mathrm{S} \cdot \overset{-1}{\mathrm{S}} = \mathrm{I}$.

On obtient alors la relation cherchée entre les composantes du tenseur u dans les bases e et e'.

Exemple : Le contracté ou trace d'un opérateur linéaire a ayant pour composantes a^i_j dans une base e est le scalaire (tenseur d'ordre zéro) : $\mathrm{Tr}\, a = \sum a^i_i$. Si $a = x \otimes \alpha$, sa trace est $\langle x, \alpha \rangle$. La contraction du i-ème indice contravariant et du j-ème indice covariant du tenseur décomposable de l'exemple 3) ci-dessus : $x_{(1)} \otimes \ldots x_{(p)} \otimes \alpha_{(1)} \otimes \ldots \alpha_{(q)}$ est le tenseur décomposable $\langle x_{(i)}, \alpha_{(j)} \rangle x_{(1)} \ldots \widehat{x}_{(i)} \otimes \otimes \ldots x_{(p)} \otimes \alpha_{(1)} \otimes \ldots \widehat{\alpha}_{(j)} \otimes \ldots \alpha_{(q)}$ où le signe $\widehat{}$ exprime que l'on a enlevé le terme situé par dessous.

A l'usage, la définition des tenseurs au moyen des composantes présente des infériorités techniques sérieuses. Par exemple :

- on voit mal le lien qui peut exister entre tenseurs quelconques et tenseurs décomposables;

- on peut définir formellement la somme de deux tenseurs quelconques. Mais cela n'a pas grand sens en soi : la somme, par exemple d'une vitesse et d'une force n'a aucune signification. Il est souhaitable de rassembler les tenseurs de même nature en « espaces tensoriels » ;
- l'opération de contraction de la proposition précédente semble tomber du ciel ;
- plus généralement, on ne voit pas très bien comment on peut organiser les relations qui doivent exister entre tenseurs de divers ordres.

Pour, toutes ces raisons, on ne peut se contenter de la définition « phénoménologique » ou « empirique » précédente.

En approfondissant la notion de tenseur, nous allons obtenir une définition constructive, qui ne s'intéresse plus à un tenseur à la fois, mais aux espaces tensoriels. Elle est basée sur une opération algébrique fondamentale : le produit tensoriel, autrefois appelé produit kroneckérien (du nom de l'algébriste Kronecker, 1823-1891).

Les propriétés de variance des composantes des tenseurs ne font plus partie de la définition. Elles sont démontrées, comme conséquence de celle-ci. De plus toutes les constructions sont valables pour des espaces vectoriels de dimension finies ou infinies.

II.7 – Produit tensoriel d'espaces vectoriels de dimensions finies ou infinies

Le produit tensoriel n'a rien de mystérieux. Il repose simplement sur une analyse de la notion d'application bilinéaire, et plus généralement d'application multilinéaire.

Les exemples que l'on a déjà vus pour un espace vectoriel E de dimension finie peuvent servir de référence à la définition ci-après :

Définition II.7.A :

Soient E et F deux espaces vectoriels sur le même corps des scalaires K. On appelle *produit tensoriel de* E *et de* F le couple formé d'un espace vectoriel T et d'une application bilinéaire, le produit tensoriel, de $\mathrm{E} \times \mathrm{F}$ à valeurs dans T, que l'on note à l'aide du signe $\otimes$ ($x \otimes y$ est donc l'élément de T, produit tensoriel de $x \in \mathrm{E}$ et de $y \in \mathrm{F}$) ces données satisfaisant à l'axiome suivant :

Pour au moins une base $e = (e_i)$, $i \in I$ de E et une base $f = (f_j)$, $j \in$ J de F, les produits tensoriels $e_i \otimes f_j$, dont les indices parcourent les couples $(i, j) \in I \times J$, forment une base de T, que l'on note $e \otimes f$.

Compléments et commentaires sur la définition précédente :

1) Nous verrons que la propriété pour les produits tensoriels des éléments d'une base de E et d'une base de F de former une base de T est satisfaite pour toutes les bases de E et de F : l'axiome $(\mathcal{T})$ ne dépend donc qu'en apparence du choix de bases particulières.

2) Nous verrons également que le produit tensoriel de E et de F est *unique* à un isomorphisme près de la structure. On peut donc utiliser pour T la notation $\mathrm{T} = \mathrm{E} \otimes \mathrm{F}$.

3) Si les dimensions de E et de F sont finies il résulte de l'axiome $(\mathcal{T})$ que $\dim(\mathrm{E} \otimes \mathrm{F}) = \dim \mathrm{E} \cdot \dim \mathrm{F}$.

4) Le produit tensoriel étant bilinéaire par définition si $x = \sum e_i x^i$ et $y = \sum f_j y^j$ sont des vecteurs de E et de F, on a : $x \otimes y = \sum (e_i \otimes f_j) x^i y^j$. Les éléments $x \otimes y$ qui forment l'image de l'application $\otimes$ sont dits *décomposables*, ou de rang un. Ils engendrent $\mathrm{E} \otimes \mathrm{F}$ d'après l'axiome $(\mathcal{T})$.

5) L'axiome $(\mathcal{T})$ assure immédiatement *l'existence* d'un produit tensoriel de E par F. En effet, un ensemble quelconque S peut toujours être considéré comme la base d'un espace vectoriel sur K : l'espace vectoriel, noté $\mathrm{K}^{(\mathrm{S})}$, des combinaisons linéaires finies $\left\{ \sum_{s \in \mathrm{S}} s\lambda^s \right\}$ d'éléments de S à coefficients dans K. Il suffit donc de prendre pour S l'ensemble des couples (e_i, f_j), $(i, j) \in \mathrm{I} \times \mathrm{J}$, pour $\mathrm{T} = \mathrm{E} \otimes \mathrm{F}$ l'espace vectoriel des combinaisons linéaires finies à coefficients dans K : $\sum (e_i, f_j) \lambda^{ij}$ et pour produit tensoriel l'application $(x, y) \rightarrow x \otimes y = \sum (e_i, f_j) x^i y^j$ ce qui entraîne $(e_i, f_j) = e_i \otimes f_j$.

L'axiome $(\mathcal{T})$ ainsi que la construction de $\mathrm{E} \otimes \mathrm{F}$ du numéro précédent ne sont pas intrinsèques, puisqu'ils font intervenir explicitement des bases particulières de E et de F.

Nous allons donner à l'axiome $(\mathcal{T})$ une autre forme entièrement équivalente pour ce qui concerne le produit tensoriel d'espace vectoriels mais présentant le triple avantage : d'être intrinsèque, d'être valable pour des produits tensoriels plus généraux (de modules sur un anneau) et surtout d'être un puissant instrument de démonstration.

Son inconvénient est qu'elle semble un peu plus abstraite, ce qui peut effrayer les utilisateurs des tenseurs, qui pensent souvent, à tort, que « l'abstraction » va en sens inverse de l'application.

Cette seconde forme de l'axiome $(\mathcal{T})$ repose sur l'observation suivante : si φ est une application bilinéaire de $\mathrm{E} \times \mathrm{F}$ à valeurs dans un espace vectoriel V quelconque, sur le même corps K, e une base de E, f une base de F, on a, quels que soient les vecteurs $x = \sum e_i x^i$ de E et $y = \sum f_j y^j$ de F, $\varphi(x, y) = \sum\limits_{i,j} \varphi(e_i, f_j) x^i y^j$ par bilinéarité. Cela montre que φ est entièrement déterminée par ses valeurs $\varphi(e_i, f_j)$ sur les couples (e_i, f_j) de vecteurs des bases e et f. Mais réciproquement, on peut se donner arbitrairement des éléments $\varphi(e_i, f_j)$ dans l'espace image V et cela détermine une application bilinéaire, unique, de $\mathrm{E} \times \mathrm{F}$ dans V en étendant par bilinéarité $\varphi(x, y) = \sum\limits_{i,j} \varphi(e_i, f_j) x^i y^j$.

Le comportement des applications bilinéaires de $\mathrm{E} \times \mathrm{F}$ dans V vis-à-vis des couples (e_i, f_j) est ainsi exactement le même que celui des applications linéaires vis-à-vis des éléments d'une base. Il résulte de l'axiome $(\mathcal{T})$ que les applications bilinéaires $\mathrm{E} \times \mathrm{F}$ dans V correspondent biunivoquement aux applications linéaires Φ de $\mathrm{E} \otimes \mathrm{F}$ dans V. En effet, φ détermine une application linéaire unique Φ en prenant $\Phi(e_i \otimes f_j) = \varphi(e_i, f_j)$ d'où il résulte par linéarité que $\Phi(x \otimes y) = \varphi(x, y)$: on obtient φ en composant Φ et l'application standard $\otimes$. Réciproquement, si Φ est une application linéaire de $\mathrm{E} \otimes \mathrm{F}$ dans V, par composition avec $\otimes$, elle donne une application bilinéaire φ de $\mathrm{E} \times \mathrm{F}$ dans V. On a ainsi obtenu la forme $(\mathcal{U})$, intrinsèque, de l'axiome $(\mathcal{T})$, caractérisant le produit tensoriel $\mathrm{E} \otimes \mathrm{F}$:

$(\mathcal{U})$ pour toute application bilinéaire φ de $\mathrm{E} \times \mathrm{F}$ dans un espace vectoriel quelconque V sur le même corps, il existe une application linéaire *unique* Φ de $\mathrm{T} = \mathrm{E} \otimes \mathrm{F}$ dans V telle que $\varphi = \Phi \circ \otimes$. On exprime cette propriété en disant que le diagramme suivant est « commutatif » :

$$\begin{tikzcd} E \times F \arrow[rr, "\varphi \text{ bilinéaire}"] \arrow[dr, "\otimes"'] & & V \\ & T = E\otimes F \arrow[ur, "\Phi \text{ linéaire unique}"'] & \end{tikzcd} \qquad \text{avec } \varphi = \Phi o \otimes .$$

Cette propriété s'exprime donc par une bijection, ou identification : $\mathrm{B}il(\mathrm{E}, \mathrm{F}; \mathrm{V}) \equiv \mathfrak{L}(\mathrm{E} \otimes \mathrm{F}, \mathrm{V})$.

L'application $\otimes$ apparaît comme une *application bilinéaire universelle*, et ce diagramme est souvent appelé diagramme universel du produit tensoriel. Une remarque importante est que l'unicité de Φ est équivalente au fait que l'image de l'application $\otimes$ engendre l'espace vectoriel $\mathrm{E} \otimes \mathrm{F}$.

Nous avons démontré ci-dessus que si $\mathrm{T} = \mathrm{E} \otimes \mathrm{F}$ satisfait à l'axiome $(\mathcal{T})$, il vérifie l'axiome $(\mathcal{U})$, mais il reste à démontrer la réciproque pour s'assurer de leur équivalence.

Cet exercice va permettre d'exploiter toutes les ressources du diagramme universel. On peut d'abord, dans ce diagramme prendre pour φ l'application $\otimes$ elle-même.

$$\begin{tikzcd} E \times F \arrow[rr, "\otimes = \varphi"] \arrow[dr, "\otimes"'] & & E\otimes F \\ & E\otimes F \arrow[ur, "\Phi"'] & \end{tikzcd} \qquad \text{avec } \Phi o \otimes = \otimes$$

L'axiome $(\mathcal{U})$ nous affirme que Φ est unique, donc, ne peut être que l'identité de $\mathrm{E} \otimes \mathrm{F}$: il n'y a pas d'autre opérateur linéaire de $\mathrm{E} \otimes \mathrm{F}$ qui soit l'identité sur les éléments décomposables $x \otimes y$ autre que l'identité.

Soient maintenant $\{\otimes; \mathrm{T} = \mathrm{E} \otimes \mathrm{F}\}$ et $\{\otimes'; \mathrm{T}' = \mathrm{E} \otimes' \mathrm{F}\}$ deux produits tensoriels de $\mathrm{E} \otimes \mathrm{F}$ satisfaisant à l'axiome $(\mathcal{U})$.

En vertu de $(\mathcal{U})$ il existe des applications linéaires uniquement déterminées Φ et Φ' telles que le diagramme :

$$\begin{tikzcd} E \times F \arrow[rr, "\otimes"] \arrow[dr, "\otimes'"'] & & E\otimes F \arrow[dl, shift right, "\Phi"'] \\ & E\otimes' F \arrow[ur, shift right, "\Phi'"'] & \end{tikzcd}$$

soit commutatif : $\otimes' = \Phi \circ \otimes$ et $\otimes = \Phi' \circ \otimes'$. En composant la première égalité avec Φ' et la seconde avec Φ on obtient : $\otimes = (\Phi' \circ \Phi) \circ \otimes$ et $\otimes' = (\Phi \circ \Phi') \circ \otimes'$ d'où il résulte d'après ce qui précède que $\Phi' \circ \Phi$ est l'identité de $\mathrm{E} \otimes \mathrm{F}$ et $\Phi \otimes \Phi'$ l'identité de $\mathrm{E} \otimes' \mathrm{F}$: Φ et Φ' sont donc des *isomorphismes préservant le produit tensoriel, uniquement déterminés entre* $\mathrm{E} \otimes \mathrm{F}$ et $\mathrm{E} \otimes' \mathrm{F}$. Si maintenant $\mathrm{T}' = \mathrm{E} \otimes' \mathrm{F}$ est un produit tensoriel de E et de F satisfaisant à l'axiome $(\mathcal{U})$, soit $\mathrm{T} = \mathrm{E} \otimes \mathrm{F}$ un produit tensoriel de E et de F satisfaisant à l'axiome $(\mathcal{T})$ relativement à des bases e de E et f de F. Puisque $\mathrm{E} \otimes \mathrm{F}$ satisfait aussi à $(\mathcal{U})$, on a un isomorphisme, unique, Φ de $\mathrm{E} \otimes \mathrm{F}$ sur $\mathrm{E} \otimes' \mathrm{F}$ tel que $\otimes' = \Phi \circ \otimes$, ce qui prouve que les éléments $e_i \otimes' f_j = \Phi(e_i \otimes f_j)$ forment une base de $\mathrm{E} \otimes' \mathrm{F}$ qui satisfait ainsi à $(\mathcal{T})$.

On aurait pu prendre pour $\mathrm{E} \otimes' \mathrm{F}$ un produit tensoriel satisfaisant à $(\mathcal{T})$ pour des bases e' de E et f' de F.

L'isomorphisme Φ' de $\mathrm{E} \otimes' \mathrm{F}$ sur $\mathrm{E} \otimes \mathrm{F}$ montre que pour ces autres bases, les $e'_k \otimes f'_l$ forment également une base de $\mathrm{E} \otimes \mathrm{F}$.

II.8 – Quelques propriétés du produit tensoriel

Un élément t du produit tensoriel $\mathrm{E} \otimes \mathrm{F}$ peut s'écrire, comme somme, ou combinaison linéaire, d'un nombre fini d'éléments décomposables, d'une infinité de façons. En effets, on peut toujours remplacer $x \otimes y$ par :

$$x \otimes y = (x + x') \otimes y - x' \otimes y = x\lambda \otimes y\,\frac{1}{\lambda} = x \otimes (y + y') - x \otimes y'.$$

Il est ainsi malcommode, voire impraticable, de définir directement une application linéaire de $\mathrm{E} \otimes \mathrm{F}$ dans un espace vectoriel V. En pratique, pour définir une telle application, on passe **toujours** par une application bilinéaire de $\mathrm{E} \times \mathrm{F}$ dans V, qui la détermine complètement. Nous allons voir de nombreux exemples de cette remarque :

Un isomorphisme canonique de $\mathrm{E} \otimes \mathrm{F}$ *sur* $\mathrm{F} \otimes \mathrm{E}$.

En effet l'application φ de $\mathrm{E} \times \mathrm{F}$ dans $\mathrm{F} \otimes \mathrm{E}$: $\varphi(x, y) = y \otimes x$ est bilinéaire et détermine donc une application linéaire unique Φ de $\mathrm{E} \otimes \mathrm{F}$ dans $\mathrm{F} \otimes \mathrm{E}$ telle que $\Phi(x \otimes y) = y \otimes x$. De la même façon

$\psi(y,x) = x \otimes y$ détermine une application linéaire ψ de $\mathrm{F} \otimes \mathrm{E}$ dans $\mathrm{E} \otimes \mathrm{F}$ et $\psi \circ \Phi$ et $\Phi \circ \psi$ sont des applications identiques d'après l'unicité Φ et ψ sont donc des isomorphismes inverses l'un de l'autre.

Un isomorphisme canonique de $\mathrm{E} \otimes \mathrm{K}$ *sur* E.

$\varphi(x, \lambda) = x\lambda$ détermine une application linéaire Φ de $\mathrm{E} \otimes \mathrm{K}$ sur E et $\psi(x) = x \otimes 1$ est une application linéaire de E dans $\mathrm{E} \otimes \mathrm{K}$, telle que $\Phi \circ \psi$ et $\psi \circ \Phi$ soient des applications identiques.

Tráce

Le crochet de dualité entre $x \in \mathrm{E}$ et $\alpha \in \mathrm{E}^*$ détermine une application bilinéaire $(x, \alpha) \to \langle \alpha, x \rangle$ de $\mathrm{E} \times \mathrm{E}^*$ dans K, d'où une application linéaire, la trace, de $\mathrm{E} \otimes \mathrm{E}^*$ dans K : $\mathrm{Tr}(x \otimes \alpha) = \langle x, \alpha \rangle$.

Si la dimension de E est finie, $e = (e_i)$ une base de E, e^* la base duale de E^*, la trace d'un élément a de $\mathrm{E} \otimes \mathrm{E}^*$ s'écrit :

$$\mathrm{Tr}\, a = \mathrm{Tr}\Big(\sum e_i \otimes e^{*j} a^i_j\Big) = \sum \langle e_i, e^{*j} \rangle a^i_j = \sum a^i_i.$$

Produit tensoriel d'applications et produits kroneckérien de matrices

Si $\mathrm{E}, \mathrm{E}', \mathrm{F}, \mathrm{F}'$ sont quatre espace vectoriels sur le même corps K, a une application linéaire de E dans E', b une application linéaire de F dans F', l'application bilinéaire : $\varphi(x,y) = a(x) \otimes b(y)$ de $\mathrm{E} \times \mathrm{F}$ dans $\mathrm{E}' \otimes \mathrm{F}'$ détermine une application linéaire unique de $\mathrm{E} \otimes \mathrm{F}$ dans $\mathrm{E}' \otimes \mathrm{F}'$ appelée produit tensoriel des applications a et b, et notée $a \bar{\otimes} b$. Une barre est nécessaire au-dessus du signe $\otimes$, pour distinguer l'élément $a \bar{\otimes} b \in \mathfrak{L}(\mathrm{E} \otimes \mathrm{F}, \mathrm{E}' \otimes \mathrm{F}')$ de $a \otimes b \in \mathfrak{L}(\mathrm{E}, \mathrm{E}') \otimes \mathfrak{L}(\mathrm{F}, \mathrm{F}')$. On a donc, par définition :

$$(a \bar{\otimes} b)(x \otimes y) = a(x) \otimes b(y).$$

En fait l'application bilinéaire : $\psi(a,b) = a \bar{\otimes} b$ détermine une application linéaire ψ telle que $\psi(a \otimes b) = a \bar{\otimes} b$, de $\mathfrak{L}(\mathrm{E}, \mathrm{E}') \otimes \mathfrak{L}(\mathrm{F}, \mathrm{F}')$ dans $\mathfrak{L}(\mathrm{E} \otimes \mathrm{E}', \mathrm{F} \bar{\otimes} \mathrm{F}')$, qui est un isomorphisme (canonique) lorsque les dimensions des espaces sont finies.

Dans ce dernier les choix de bases $e = (e_1)$ de E, $f = (f_j)$ de F, $e' = (e'_k)$ de E', $f' = (f'_1)$ de F' détermine une base $e \otimes f = (e_i \otimes f_j)$ de $\mathrm{E} \otimes \mathrm{F}$ indexée par les couples (i,j) et une base $e' \otimes f' = (e'_k \otimes f'_1)$ de $\mathrm{E}' \otimes \mathrm{F}'$, indexée par les couples (k,l).

Si $\mathrm{A}, \mathrm{B}, \mathrm{C} = \mathrm{A} \bar{\otimes} \mathrm{B}$ sont les matrices de a, b et $a \bar{\otimes} b$ dans ces bases, la matrice $\mathrm{C} = \mathrm{A} \bar{\otimes} \mathrm{B}$ est *le produit tensoriel ou kroneckérien des matrices* A et B. Elle a pour indices de ses colonnes les couples (i, j) que l'on peut ordonner par l'ordre lexicographique, et pour indices de ses lignes les couples (k, l).

$$\begin{aligned}(a \bar{\otimes} b)(e_i \otimes f_j) = a(e_i) \otimes b(f_j) &= \left(\sum e'_k a_i^k\right) \otimes \left(\sum f'_l b_j^l\right) \\ &= \sum (e'_k \otimes f'_l) a_i^k b_j^l.\end{aligned}$$

d'où : $C_{ij}^{kl} = a_i^k b_j^l$

Si maintenant a' est une application linéaire de E' dans E'', b' une application de F' dans F'', la composée des applications $a' \bar{\otimes} b'$ et $a \bar{\otimes} b$ s'écrit sur les éléments décomposables :

$$\begin{aligned}(a' \bar{\otimes} b')((a \bar{\otimes} b)(x \otimes y)) &= (a' \bar{\otimes} b')(a(x) \otimes b(y)) \\ &= (a'a)(x) \otimes (b'b)(y) = (a'a \bar{\otimes} b'b)(x \otimes y)\end{aligned}$$

ce qui entraîne l'égalité

$$(a' \bar{\otimes} b') \circ (a \bar{\otimes} b) = a'a \bar{\otimes} b'b.$$

En particulier, prenons $\mathrm{E}'' = \mathrm{E}$, $\mathrm{F}'' = \mathrm{F}$. Si a et b sont des isomorphismes, $a \bar{\otimes} b$ est également un isomorphisme, car :

$$(\overset{-1}{a} \bar{\otimes} \overset{-1}{b}) \circ (a \bar{\otimes} b) = \mathrm{Id}_\mathrm{E} \otimes \mathrm{Id}_\mathrm{F} = \mathrm{Id}_{\mathrm{E} \otimes \mathrm{F}},$$

d'où :

$$(\overset{-1}{a} \bar{\otimes} \overset{-1}{b}) = (a \overset{-1}{\bar{\otimes}} b).$$

Il en résulte aussi que si $a'' \in \mathfrak{L}(\mathrm{E}'', \mathrm{E}''')$ et $b'' \in \mathfrak{L}(\mathrm{E}'', \mathrm{E}''')$ on $\circ$ la propriété d'associativité :

$$(a'' \bar{\otimes} b'') \circ [(a') \bar{\otimes} b') \circ (a \bar{\otimes} b)] = [(a'' \bar{\otimes} b'') \circ (a' \bar{\otimes} b')] \circ (a \bar{\otimes} b).$$

Produit tensoriel de deux représentations K*-linéaires d'un groupe* G *:*

Une représentation K-linéaire d'un groupe G est un homorphisme de G dans le groupe des automorphismes d'un espace K-vectoriel $\mathrm{E} : \mathrm{G} \to \mathrm{Gl}(\mathrm{E})$.

Notons θ cet homomorphisme. Tout élément g de G est ainsi représenté par l'opérateur linéaire $\theta(g) \in \mathrm{Gl(E)}$. La fonction sur G à valeurs dans K : $\chi_\theta(g) = \text{Trace } \theta(g)$ est le *caractère* de θ.

θ est dite irréductible s'il n'existe aucun sous-espace propre A de E appliqué en lui-même par tous les $\theta(g)$, $g \in \mathrm{G}$ (qui fournirait alors une sous-représentation : $\mathrm{G} \to \mathrm{Gl(A)}$ par $g \to \theta(g)|_{\mathrm{A}}$.

Exemple : On a vu (§II.5) que le groupe linéaire Gl(E) possède des représentations linéaires naturelles dans les espaces E^*, $\mathrm{T}_2(\mathrm{E})$, $\mathrm{T}^1_1(\mathrm{E})$, ... On verra plus généralement que Gl(E) est représenté dans tous les « espaces tensoriels » sur E (cf III.2).

Si maintenant θ' est une représentation K-linéaire de G dans F, les propriétés du paragraphe précédent montrent que l'application : $g \to \theta(g) \mathbin{\bar{\otimes}} \theta'(g)$ dans $\mathrm{Gl(E \otimes F)}$ est une représentation de G, notée $\theta \otimes \theta'$ appelée *produit tensoriel de θ et θ'.*

Un élément t de $\mathrm{E \otimes F}$ peut s'écrire $t = \sum_j x_j \otimes y_j$ et $(\theta \otimes \theta')(g) \cdot t = \sum_j \theta(g)x_j \otimes \theta(g)y_j$.

Afin d'éviter les notations trop lourdes, on écrit le plus souvent $g \cdot x$, $g \cdot y$ pour $\theta(g) \cdot x$, $\theta'(g) \cdot y$ lorsqu'il n'y a pas d'ambiguité. On voit donc que l'opération de g sur $t \in \mathrm{E \otimes F}$ s'obtient en distribuant g dans une expression quelconque de t (il y en a une infinité !) :

$$g \cdot t = \sum_j g \cdot x_j \otimes g \cdot y_j.$$

On vérifie immédiatement que $\chi_{\theta \otimes \theta'} = \chi_\theta \cdot \chi_{\theta'}$.

Problème : θ et θ' étant données et irréductibles. Trouver toutes les sous-représentations irréductibles de $\theta \otimes \theta'$.

Lorsque G est le groupe des rotations de notre espace euclidien à trois dimensions, la solution de ce problème a de nombreuses applications en physique.

Complexifié d'un espace vectoriel réel

Soient E un espace vectoriel sur le corps complexe $\mathbb{C}$, ${}_{\mathbb{R}}\mathrm{E}$ l'espace vectoriel réel sous-jacent obtenu en restreignant les scalaires au sous-corps $\mathbb{R}$ de $\mathbb{C}$. $\mathbb{C}$ est un plan réel de base $(1, i)$ et la structure complexe de E est déterminée par une application $\mathbb{R}$-linéaire ψ de $\mathbb{C}$ dans $\mathfrak{L}({}_{\mathbb{R}}\mathrm{E})$ telle que $\psi(1) = \mathrm{Id}({}_{\mathbb{R}}\mathrm{E})$ et $\psi(i)^2 = -\mathrm{Id}({}_{\mathbb{R}}\mathrm{E})$. Il résulte de ces conditions que si z et $u \in \mathbb{C}$, on a automatiquement la propriété multiplicative : $\psi(zu) = \psi(z)\psi(u)$ et il suffit évidemment de se donner l'opérateur $\mathrm{J} = \psi(i)$.

Une application $\mathbb{C}$-linéaire de E dans F est une application $\mathbb{R}$-linéaire f de ${}_{\mathbb{R}}\mathrm{E}$ dans ${}_{\mathbb{R}}\mathrm{F}$ telle que $f \circ \psi_{\mathrm{E}}(z) = \psi_{\mathrm{F}}(z) \circ f$, $\forall z \in \mathbb{C}$.

Si V est un espace vectoriel réel, son *complexifié* $\mathrm{V}_{\mathbb{C}}$ est le produit tensoriel d'espaces vectoriels réels : $\mathrm{V}_{\mathbb{C}} = \mathrm{V} \otimes \mathbb{C}$. La structure complexe de $\mathrm{V}_{\mathbb{C}}$ est déterminée par l'application $\mathbb{R}$-linéaire ψ de $\mathbb{C}$ dans $\mathfrak{L}(\mathrm{V} \otimes \mathbb{C})$: $\psi(1) = \mathrm{Id}_{\mathrm{V}} \bar{\otimes} \mathrm{Id}_{\mathbb{C}}$ et $\psi(i) = \mathrm{Id}_{\mathrm{V}} \bar{\otimes}$ (multiplication par i) soit, pour un élément décomposable $(x \otimes u)$ de $\mathrm{V} \otimes \mathbb{C}$, $(x \in \mathrm{V}, u \in \mathbb{C})$, si $z \in \mathbb{C}$, $\psi(z)(x \otimes u) = x \otimes zu$.

Tout élément de $\mathrm{V}_{\mathbb{C}}$ s'écrit d'une façon et d'une seule : $x \otimes 1 + y \otimes i = x \otimes 1 + (\mathrm{Id}_{\mathrm{V}} \bar{\otimes}$ (multiplication par i) $(y \otimes 1)$.

Il en résulte qu'en identifiant V au sous-espace réel $\mathrm{V} \otimes 1$ de $\mathrm{V} \otimes \mathbb{C}$ tout élément de $\mathrm{V}_{\mathbb{C}}$ s'écrit et d'une seule $x + iy$ avec $x, y \in \mathrm{V}$. Si l désigne l'inclusion de V dans $\mathrm{V}_{\mathbb{C}}$: $l(x) = x \otimes 1$, le couple $(l, \mathrm{V}_{\mathbb{C}})$ possède la propriété universelle suivante : pour toute application $\mathbb{R}$-linéaire f de V dans un espace $\mathbb{C}$-vectoriel E, il existe une application $\mathbb{C}$-linéaire unique $\widetilde{f}$ de $\mathrm{V}_{\mathbb{C}}$ dans E qui prolonge f, c'est-à-dire que $\widetilde{f} \circ l = f$. En effet, l'application $\mathbb{R}$-bilinéaire de $\mathrm{V} \times \mathbb{C}$ dans E qui au couple (x, u) fait correspondre $\psi(u)(f(x))$ détermine une application $\mathbb{R}$-linéaire $\widetilde{f}$ de $\mathrm{V} \otimes \mathbb{C}$ dans E qui vérifie $\widetilde{f} \circ \psi(z) = \psi(z)\widetilde{f}$ et qui est donc une application $\mathbb{C}$-linéaire de $\mathrm{V}_{\mathbb{C}}$ dans E.

Si a est une application linéaire réelle de V dans V', $a_{\mathbb{C}} = a \bar{\otimes} \mathrm{Id}_{\mathbb{C}}$ est une application linéaire réelle complexe de $\mathrm{V}_{\mathbb{C}}$ dans $\mathrm{V}'_{\mathbb{C}}$ appelée la *complexifié de a.*

II.9 – Produit tensoriel de $p > 2$ espaces vectoriels

Soient p espaces vectoriels : $\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p$ sur le même corps K. On appelle produit tensoriel de $\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p$ le couple formé d'un espace vectoriel T, qu'en raison de l'unicité prouvée ci-après, on peut noter $\mathrm{T} = \mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \mathrm{E}_p$, et d'une application multilinéaire de $\mathrm{E}_1 \times \mathrm{E}_2 \ldots \mathrm{E}_p$, appelée produit tensoriel et notée : $(x_1, x_2, \ldots, x_p) \rightarrow x_1 \otimes x_2 \otimes \ldots x_p$ satisfaisant à l'un ou l'autre des axiomes équivalents suivants :

(T) il existe des bases $e_{(1)} = (e_{(1)i_1}, i_1 \in \mathrm{I}_1), \ldots, e_{(p)} = (e_{(p)i_p}; i_p \in \mathrm{I}_p)$ de $\mathrm{E}_1, \ldots, \mathrm{E}_p$, telles que les produits tensoriels : $e_{(1)i_1} \otimes e_{(2)i_2} \otimes \ldots e_{(p)i_p}$, indexés par le produit des ensembles d'indices $\mathrm{I}_1 \times \mathrm{I}_2 \times \ldots \mathrm{I}_p$, forment une base de $\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p$

(U) propriété universelle : pour toute application p-linéaire φ de $\mathrm{E}_1 \times \mathrm{E}_2 \times \ldots \mathrm{E}_p$ dans un espace vectoriel V, sur le même corps, il existe une application linéaire unique Φ de $\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p$ dans V telle que $\varphi = \Phi \circ \otimes$: (le diagramme est commutatif)

$$\begin{tikzcd}
\mathrm{E}_1 \times \mathrm{E}_2 \times \ldots \mathrm{E}_p \arrow[rr, "\varphi", "p\text{-linéaire}"'] \arrow[dr, "\otimes"'] & & \mathrm{V} \\
& \mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p \arrow[ur, "\phi \text{ linéaire unique}"'] &
\end{tikzcd}$$

Ainsi : $\varphi(x_1, x_2, \ldots, x_p) = \Phi(x_1 \otimes x_2 \otimes \ldots x_p)$ et l'application $\otimes$ peut être considérée comme une application p-linéaire universelle.

On démontre, comme dans le cas de deux espaces vectoriels, l'unicité à un isomorphisme près, l'existence, et le fait que la propriété (T) est valable quel que soit le choix des bases.

Les éléments $x_1 \otimes x_2 \otimes \ldots x_p$ de l'image de $\otimes$ sont dits décomposables.

Tout élément de $\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p$ est combinaison linéaire, ou somme, d'un nombre fini d'élément décomposables.

Chaque fois que l'on souhaite construire une application linéaire de $\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p$ dans un espace vectoriel V, *IL FAUT* pour cela passer par l'application p-linéaire associée, et c'est elle qu'il faut définir pour atteindre l'application linéaire.

Par exemple, le produit $a_1 \bar{\otimes}\, a_2 \bar{\otimes} \ldots a_p$ de p applications linéaires $a_1 : \mathrm{E}_1 \to \mathrm{E}'_1$, $a_2 : \mathrm{E}_2 \to \mathrm{E}'_2$, …, $a_p : \mathrm{E}_p \to \mathrm{E}'_p$ est l'application linéaire de $\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p$ dans $\mathrm{E}'_1 \otimes \mathrm{E}'_2 \otimes \ldots \mathrm{E}'_p$ associée à l'application p-linéaire :

$$(x_1, x_2, \ldots, x_p) \to a_1(x_1) \otimes a_2(x_2) \otimes \ldots a_p(x_p).$$

On a une application linéaire naturelle :

$$\mathfrak{L}(\mathrm{E}_1, \mathrm{E}'_1) \otimes \mathfrak{L}((\mathrm{E}_2, \mathrm{E}'_2) \otimes \ldots \mathfrak{L}(\mathrm{E}_p, \mathrm{E}'_p) \\ \to \mathfrak{L}(\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p; \mathrm{E}'_1 \otimes \mathrm{E}'_2 \otimes \ldots \mathrm{E}'_p),$$

qui envoie $a_1 \otimes a_2 \otimes \dots a_p$ sur $a_1 \bar{\otimes} a_2 \bar{\otimes} \dots a_p$, qui est un isomorphisme lorsque $E_1, E_2, \dots, E_p$ sont de dimensions finies.

Il en résulte, comme au paragraphe précédent, que si θ_1, θ_2, ... θ_p sont p représentations linéaires d'un même groupe G dans des espaces vectoriels $E_1, E_2, \dots, E_p$ sur le même corps K, l'application :

$$g \in E \to \theta_1(g) \bar{\otimes} \theta_2(g) \bar{\otimes} \dots \theta_p(g),$$

de G dans le groupe linéaire $Gl(E_1 \otimes E_2 \otimes \dots E_p)$ est une représentation de G que l'on appelle produit tensoriel des représentations $\theta_1, \theta_2, \dots, \theta_p$ et que l'on note $\theta_1 \otimes \theta_2 \otimes \dots \theta_p$.

Indiquons quelques propriétés immédiates du produit tensoriel de p espaces vectoriels, qui se démontrent exactement comme dans le cas du produit tensoriel de deux espaces, c'est-à-dire en passant par une application multilinéaire :

$K \otimes K \otimes \dots \otimes K$ est canoniquement isomorphe à K, tout élément s'écrivant : $\lambda_1 \cdot 1 \otimes \lambda_2 \cdot 1 \otimes \lambda_p \cdot 1 = \lambda_1 \lambda_2 \dots \lambda_p \cdot 1 \otimes 1 \otimes \dots 1$.

Il en résulte, d'après ce qui précède, que si $E_1 E_2, \dots, E_p$ sont de dimensions finies, on a :

$$E_1^* \otimes E_2^* \otimes \dots E_p^* \equiv (E_1 \otimes E_2 \otimes \dots E_p)^*.$$

Si dans un produit tensoriel l'un des facteurs est le corps K : $E_1 \otimes E_2 \otimes \dots E_{p-1} \otimes K \otimes E_{p+1} \otimes \dots E$, ce produit est canoniquement isomorphe au produit des espaces vectoriels distincts de celui-là, soit :

$$E_1 \otimes E_2 \otimes \dots E_{p-1} \otimes E_{p+1} \otimes \dots E_r.$$

Si σ est une permutation de $(1, 2, \dots, p)$, l'application $(x_1, x_2, \dots, x_p) \to x_{\sigma 1} \otimes x_{\sigma 2} \otimes \dots x_{\sigma p}$ de $E_1 \times E_2 \times \dots E_p$ dans $E_{\sigma 1} \otimes E_{\sigma 2} \otimes \dots E_{\sigma p}$ est p-linéaire et détermine une application linéaire Φ_σ de $E_1 \otimes E_2 \otimes \dots E_p$ dans $E_{\sigma 1} \otimes E_{\sigma 2} \otimes \dots E_{\sigma p}$. C'est un isomorphisme, appliquant les éléments décomposables de l'un sur ceux de l'autre, puisque $\Phi_{\sigma^1} \circ \Phi_\sigma$ est l'identité.

II.10 – Opérations tensorielles

Le calcul tensoriel est fondé sur deux opérations : la contraction et la multiplication tensorielle.

Composées, elles en donnent une troisième : la multiplication contractée.

La propriété essentielle de ces opérations est d'avoir la même forme quelle que soit *l'écriture des éléments*. De plus, si on les effectue sur des éléments décomposables, le résultat est un élément décomposable.

Définition II.10.A :

Soient p espaces vectoriels sur K : $E_1, E_2, \ldots, E_p$, et b une forme bilinéaire sur le couple (E_i, E_j) où l'on suppose $i < j$. L'application linéaire naturelle :

$$E_1 \otimes E_2 \otimes \ldots E_p \to E_1 \otimes \ldots \widehat{E}_i \otimes \ldots \widehat{E}_j \otimes \ldots E_p,$$

associée à l'application p-linéaire :

$$(x_1, x_2, \ldots, x_p) = b(x_i, x_j) x_1 \otimes \ldots \widehat{x}_i \otimes \ldots \widehat{x}_j \otimes \ldots x_p,$$

est appelée la *contraction par b* des éléments de $E_1 \otimes E_2 \otimes \ldots E_p$.

Le cas le plus fréquent est celui où l'un des deux espaces E_i, E_j est le dual de l'autre, b étant alors le crochet de dualité, et c'est dans ce cas, une extension de la trace.

Définition II.10.B :

Soient $(p+q)$ espaces vectoriels sur K : $E_1, E_2, \ldots, E_p, F_1, F_2, \ldots, F_q$. L'application bilinéaire naturelle μ : $(E_1 \otimes E_2 \otimes \ldots E_p) \times (F_1 \otimes F_2 \otimes \ldots F_q) \to E_1 \otimes \ldots E_p \otimes F_1 \otimes \ldots F_q$ dont l'existence est démontrée ci-dessous, est appelée la *multiplication tensorielle*.

L'application φ $(x_1, x_2, \ldots, x_p, y_1, y_2, \ldots, y_q) \to (x_1 \otimes x_2 \otimes \ldots x_p) \otimes (y_1 \otimes y_2 \otimes \ldots y_q)$ de $E_1 \times E_2 \ldots E_p \times F_1 \times \ldots F_q$ dans $P = (E_1 \otimes \ldots E_p) \otimes (F_1 \otimes \ldots F_q)$ est $(p+q)$-linéaire. Si l'on choisit des bases $\{e_{(1)i_1}\} \ldots \{e_{(p)i_p}\}\{f_{(1)j_1}\} \ldots \{f_{(q)j_q}\}$ dans les espaces $E_1, \ldots, F_q$, les images par φ : $(e_{(1)i_1} \otimes \ldots e_{(p)i_p}) \otimes (f_{(1)j_1}) \otimes \ldots f_{(q)j_q})$ forment évidemment une base de P qui satisfaisant ainsi à l'axiome (T), est avec φ un produit tensoriel de $E_1, \ldots, E_q$. Il en résulte un isomorphisme naturel Φ de P sur $E_1 \otimes E_2 \otimes \ldots E_q$. La multiplication tensorielle est la composée des deux applications :

$$(\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \dots \mathrm{E}_p) \times (\mathrm{F}_1 \otimes \mathrm{F}_2 \otimes \dots \mathrm{F}_q)$$
$$\xrightarrow{\otimes} (\mathrm{E}_1 \otimes \dots \mathrm{E}_p) \otimes (\mathrm{F}_1 \otimes \dots \mathrm{F}_q) \xrightarrow{\otimes} \mathrm{E}_1 \otimes \dots \mathrm{E}_p \otimes \mathrm{F}_1 \otimes \dots \mathrm{F}_q$$

Un raisonnement analogue montre que la multiplication tensorielle est associative ce qui se traduit par des isomorphismes canoniques entre $((\mathrm{E}_1 \otimes \dots \mathrm{E}_p) \otimes (\mathrm{F}_1 \otimes \dots \mathrm{F}_q)) \otimes (\mathrm{H}_1 \otimes \dots \mathrm{H})$, $(\mathrm{E}_1 \otimes \dots \mathrm{E}_p) \otimes ((\mathrm{F}_1 \otimes \dots \mathrm{F}_q) \otimes (\mathrm{H}_1 \otimes \dots \mathrm{H}))$ et $\mathrm{E}_1 \otimes \dots \mathrm{F}_1 \otimes \dots \mathrm{H}_1 \otimes \dots \mathrm{H}_r$. On ne distingue donc pas entre ces divers produits.

Définition II.10.C :

Soient $(p+q)$espaces vectoriels sur K : $\mathrm{E}_1, \mathrm{E}_2, \dots, \mathrm{E}_p, \mathrm{F}_1$, F_2, $\dots \mathrm{F}_q$ et b une forme bilinéaire sur le couple $(\mathrm{E}_i, \mathrm{F}_k)$. L'application composée de la multiplication tensorielle μ et de la contraction par b

$$(\mathrm{E}_1 \otimes \dots \mathrm{E}_p) \times (\mathrm{F}_1 \otimes \dots \mathrm{F}_q) \xrightarrow{\mu} \mathrm{E}_1 \otimes \dots \mathrm{E}_p \otimes \mathrm{F}_1 \otimes \dots \mathrm{F}_q$$
$$\longrightarrow \mathrm{E}_1 \otimes \dots \widehat{\mathrm{E}}_i \otimes \dots \widehat{\mathrm{F}}_k \otimes \dots \mathrm{F}_q,$$

est appelée *multiplication contractée.*

Exemples :

$a)$ soient E et F deux espaces vectoriels de dimension finie sur le même corps K. La multiplication contractée à droite : $(\mathrm{F} \otimes \mathrm{E}^*) \times \mathrm{E} \to \mathrm{F}$ permet de considérer un élément $a \in \mathrm{F} \otimes \mathrm{E}^*$ comme une application linéaire de E dans F par : $x \in \mathrm{E} \to a \cdot x \in \mathrm{F}$ en notant par un point cette multiplication contractée.

Si $a = \sum_{,j} y_j \otimes \alpha_j, \quad a \cdot x = \sum_{,j} y_j \langle \alpha_j, x \rangle$.

L'application φ qui au couple (y, α) de $\mathrm{F} \times \mathrm{E}^*$ fait correspondre l'application linéaire de E dans F : $x \in \mathrm{E} \to y\langle \alpha, x \rangle \in \mathrm{F}$, est une application bilinéaire de $\mathrm{F} \times \mathrm{E}^*$ dans $\mathfrak{L}(\mathrm{E}; \mathrm{F})$ satisfaisant à l'axiome (T) déterminant ainsi un isomorphisme naturel de $\mathrm{F} \otimes \mathrm{E}^*$ sur $\mathfrak{L}(\mathrm{E}; \mathrm{F})$.

La multiplication contractée à gauche de $\mathrm{F}^* \times (\mathrm{F} \otimes \mathrm{E}^*)$ dans E^* : $(\beta, a) \to \beta \cdot a$ est la transposée de la précédente. La double multiplication contractée :

$$\begin{aligned} \mathrm{F}^* \times (\mathrm{F} \otimes \mathrm{E}^*) \times \mathrm{E} &\to \mathrm{K} \\ &\to \beta \cdot a \cdot x \end{aligned}$$

est en effet associative, ce qu'il suffit de vérifier pour a décomposable, $a = y \otimes \alpha$:

$$\begin{aligned} &a \cdot x = y\langle \alpha, x \rangle; \ \beta \cdot a = \langle \beta, y \rangle \alpha, \\ \text{et} \quad &\beta(a \cdot x) = \langle \beta, y \rangle \langle \alpha, x \rangle = (\beta \cdot a) \cdot x \end{aligned}$$

b) si $\mathrm{E}, \mathrm{F}, \mathrm{H}$ sont trois espaces vectoriels de dimension finie sur le même corps K, la composition des applications linéaires de E dans F et de F dans H s'exprime, dans l'écriture tensorielle des applications linéaires, par une multiplication contractée :

$$(\mathrm{H} \otimes \mathrm{F}^*) \times (\mathrm{F} \otimes \mathrm{E}^*) \to \mathrm{H} \otimes \mathrm{E}^*,$$

$(z \otimes \beta) \cdot (y \otimes \alpha) = \langle \beta, y \rangle z \otimes \alpha$ est bien le composé de $x \to (y \otimes \alpha) \cdot x = y\langle \alpha, x \rangle$ et de $y \cdot \langle \alpha, x \rangle \to (z \otimes \beta) \cdot y \langle \alpha, x \rangle = z \langle \beta, y \rangle \langle \alpha, x \rangle = \langle \alpha, x \rangle = \langle \beta, y \rangle (z \otimes \alpha) \cdot x$.

L'écriture tensorielle rend immédiate une propriété de la trace. Si $a \in \mathfrak{L}(\mathrm{E}, \mathrm{F}) = \mathrm{F} \otimes \mathrm{E}^*$ et $b \in \mathfrak{L}(\mathrm{F}, \mathrm{E}) \equiv \mathrm{E} \otimes \mathrm{F}^*$, $a \circ b \in \mathfrak{L}(\mathrm{F})$, $b \circ a \in \mathfrak{L}(\mathrm{E})$, on a $\mathrm{Tr}(a \circ b) = \mathrm{Tr}(b \circ a)$. Il suffit de prendre a et b décomposables, le cas général s'en déduisant par linéarité. Si $a = y \otimes \alpha$ et $b = x \otimes \beta$, $a \circ b = y \otimes \beta \langle \alpha, x \rangle$; $b \circ a = \langle \beta, y \rangle x \otimes \alpha$, d'où :

$$\langle \beta, y \rangle \langle \alpha, x \rangle = \mathrm{Tr}(a \circ b) = \mathrm{Tr}(b \circ a).$$

c) si $\mathrm{E}_1, \mathrm{E}_2 \ldots, \mathrm{E}_p$ sont de dimension finie sur K, la multiplication p-fois contractée :

$$\begin{gathered} (\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p) \times (\mathrm{E}_1^* \otimes \mathrm{E}_2^* \ldots \mathrm{E}_p^*), \\ ((x_1 \otimes x_2 \otimes \ldots x_p), (\alpha_1 \otimes \alpha_2 \otimes \ldots \alpha_p)) \to \langle x_1, \alpha_1 \rangle \langle x_2, \alpha_2 \rangle \ldots \langle x_p, \alpha_p \rangle \end{gathered}$$

est l'expression du crochet de dualité, par l'identification vue précédemment de $\mathrm{E}_1^* \otimes \mathrm{E}_2^* \ldots \mathrm{E}_p^*$ avec $(\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \otimes \mathrm{E}_p)^*$, sur les éléments décomposables des deux espaces.

Il en résulte que les produits tensoriels de bases des E_j et des bases duales dans les E_j^* sont des bases duales de $\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p$ et de $\mathrm{E}_1^* \otimes \mathrm{E}_2^* \ldots \mathrm{E}_p^*$.

d) si E et F sont de dimensions finies sur le même corps K, une forme bilinéaire b sur $\mathrm{E} \times \mathrm{F}$ est équivalente à une forme linéaire sur $\mathrm{E} \otimes \mathrm{F}$, soit à un élément φ de $\mathrm{E}^* \otimes \mathrm{F}^*$.

On récupère la forme b par la double multiplication contractée : $\mathrm{E} \times (\mathrm{E}^* \otimes \mathrm{F}^*) \times \mathrm{F}$ soit : $b(x, y) = x \cdot \varphi \cdot y$ (car $x \cdot (\varphi \cdot y) = (x \cdot \varphi) \cdot y$). De façon précise, l'isomorphisme de $\mathrm{E}^* \otimes \mathrm{F}^*$ sur l'espace vectoriel $\mathcal{L}_2(\mathrm{E}, \mathrm{F}; \mathrm{K})$ des formes bilinéaires sur $\mathrm{E} \times \mathrm{F}$ est obtenu, à l'aide de la propriété universelle du produit tensoriel, à partir de l'application bilinéaire de $\mathrm{E}^* \times \mathrm{F}^*$ dans $\mathcal{L}_2(\mathrm{E}, \mathrm{F}; \mathrm{K})$ qui au couple α, β fait correspondre la forme bilinéaire $(x, y) \to \alpha(x)\beta(y)$.

L'application linéaire ρ *associé à droite* de b, de F dans E^*, fait correspondre à chaque $y \in \mathrm{F}$ la forme linéaire $x \to b(x, y)$ sur E, qui est aussi la contractée à droite de φ par y : $\rho(y) = \varphi \cdot y$.

Cette définition de ρ par contraction revient à l'identification de φ et de ρ comme l'élément de $\mathrm{E}^* \otimes \mathrm{F}^*$, puisque ce produit tensoriel $\mathrm{E}^* \otimes \mathrm{F}^*$ s'interprète aussi, comme on l'a vu au a) ci-dessus comme l'espace $\mathcal{L}(\mathrm{F}, \mathrm{E}^*)$.

De la même façon, l'application γ associée à gauche de b, de E dans F fait correspondre à $x \in \mathrm{E}$ la forme : $y \to b(x, y)$ sur F. On a $\gamma(x) = x \cdot \varphi$, ce qui revient à identifier aussi γ avec φ par l'interprètation de $\mathrm{E}^* \otimes \mathrm{F}^*$ comme $\mathcal{L}(\mathrm{F}, \mathrm{E}^*)$.

On voit donc que tensoriellement, c'est *le même élément* $\varphi \in \mathrm{E}^* \otimes \mathrm{F}^*$ qui s'interprète comme b, ρ ou γ.

On a : $b(x, y) = \langle x, \rho(y) \rangle = \langle \gamma(x), y \rangle$ qui montre également qu'avec l'identification $\mathrm{E}^{**} = \mathrm{E}$ et $\mathrm{F}^{**} = \mathrm{F}$ que $\gamma =^t \rho$ et $\rho = {}^t\gamma$.

Si e et f sont des bases de E et F, et $b_{ij} = b(e_i, f_j)$, $\varphi = \sum b_{ij} e^{*i} \otimes f^{*j}$.

ρ a même matrice $\mathrm{B} = |b_{ij}|$ que b tandis que γ a pour matrice ${}^t\mathrm{B}$.

Le rang b est celui de B, de ρ, de γ ou de φ. C'est en particulier le nombre minimum de termes qu'il faut pour écrire φ comme somme de tenseurs décomposables $\varphi = \sum \alpha_j \otimes \beta_j$. « b-non-dégénérée » équivaut à ρ (ou γ) est un isomorphisme.

Dans ce cas l'inverse $\overset{-1}{\rho}$ de ρ (ou $\overset{-1}{\gamma}$ de γ) définit une forme bilinéaire sur $\mathrm{F}^* \times \mathrm{E}^*$ appelée *inverse de b et notée* $\overset{-1}{b}$. On a évidemment $\overset{\overset{-1}{-1}}{b} = b$.

Si $(\beta,\alpha) \in \mathrm{F}^* \times \mathrm{E}^*$, $\overset{-1}{b}(\beta,\alpha) = \langle\beta, \overset{-1}{\rho}\alpha)\rangle = b(\overset{-1}{\gamma}\beta, \overset{-1}{\rho}\alpha) = \langle\overset{-1}{\gamma}\beta,\alpha\rangle$. Le nom de forme bilinéaire inverse est justifié par le fait que dans les multiplications contractées :

$$(\mathrm{E}^* \otimes \mathrm{F}^*) \times (\mathrm{F} \otimes \mathrm{E}) \to \mathrm{E}^* \otimes \mathrm{E} = \mathfrak{L}(\mathrm{E}^*),$$
$$(\mathrm{F} \otimes \mathrm{E}) \times (\mathrm{E}^* \otimes \mathrm{F}^*) \to \mathrm{F} \otimes \mathrm{F}^* = \mathfrak{L}(\mathrm{F}),$$

les produits des éléments $\varphi \otimes \mathrm{E}^* \otimes \mathrm{F}^*$ représentant b et $\overset{-1}{\varphi} \in \mathrm{F} \otimes \mathrm{E} = (\mathrm{F}^*)^* \otimes (\mathrm{E}^*)^*$ représentant $\overset{-1}{b}$ sont égaux aux unités (opérateurs identiques) de $\mathrm{E}^* \otimes \mathrm{E}$ et $\mathrm{F} \otimes \mathrm{F}^*$. C'est en effet, parfaitement évident si l'on interprète φ comme ρ et $\overset{-1}{\varphi}$ comme $\overset{-1}{\rho}$.

En prenant $\mathrm{F} = \mathrm{E}$, on voit qu'une forme bilinéaire non-dégénérée b sur E détermine automatiquement une forme bilinéaire non-dégénérée $\overset{-1}{b}$ sur le dual E^*. Cette propriété va jouer un rôle essentiel dans l'étude des tenseurs restreints (chapitre V).

II.11 – Produit tensoriel de formes bilinéaires

Soient $\mathrm{E}_1, \mathrm{F}_1, \mathrm{E}_2, \mathrm{F}_2, \ldots, \mathrm{E}_p, \mathrm{F}_p$, $2p$ espaces vectoriels de dimensions finies ou infinies sur le même corps K, b_1 une forme bilinéaire sur le couple $\mathrm{E}_1 \times \mathrm{F}_1$, b_2 sur $\mathrm{E}_2 \times \mathrm{F}_2$, ..., b_p sur $\mathrm{E}_p \times \mathrm{F}_p$. L'application qui aux $2p$ vecteurs : $(x_1, x_2, \ldots, x_p, y_1, y_2, \ldots, y_p)$ de $(\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p, \mathrm{F}_1, \mathrm{F}_2, \ldots, \mathrm{F}_p)$ associe le scalaire égal au produit : $b_1(x_1, y_1) b_2(x_2, y_2) \ldots b_p(x_p, y_p)$ est $2p$-linéaire et détermine donc une application linéaire B :

$$\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p \otimes \mathrm{F}_1 \otimes \mathrm{F}_2 \otimes \ldots \mathrm{F}_p \xrightarrow{\mathrm{B}} \mathrm{K},$$

par ses valeurs sur les éléments décomposables :

$$B(x_1 \otimes x_2 \otimes \ldots x_p \otimes y_1 \otimes y_2 \otimes \ldots y_p) = \prod b_j(x_j, y_j).$$

En composant avec la multiplication tensorielle μ :

$$(\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p) \times (\mathrm{F}_1 \otimes \mathrm{F}_2 \otimes \ldots \mathrm{F}_p) \to \mathrm{E}_1 \otimes \ldots \mathrm{E}_p \otimes \mathrm{F}_1 \ldots \mathrm{F}_p,$$

on obtient une forme bilinéaire $b = \mathrm{B} \circ \mu$ appelée produit tensoriel de $b_1, b_2, \ldots, b_p$ sur le couple :
$(\mathrm{E}_1 \otimes \ldots \mathrm{E}_p) \times (\mathrm{F}_1 \otimes \ldots \mathrm{F}_p)$ entièrement définie par ses valeurs sur les éléments *décomposables* :

$$b(x_1 \otimes \ldots x_p; y_1 \otimes \ldots y_p) = b_1(x_1, y_1) \ldots b_p(x_p, y_p).$$

En prenant tous les espaces E_j égaux à E, et tous les espaces F_j égaux à F, la construction précédente définit la *puissance tensorielle p-ème d'une forme bilinéaire b sur* $\mathrm{E} \times \mathrm{F}$.

C'est une forme bilinéaire sur $(\otimes^p \mathrm{E} \times (\otimes^p \mathrm{F})$, notée $\bar{\otimes}^p b$, entièrement définie par ses valeurs sur les tenseurs décomposables

$$\begin{aligned}(\bar{\otimes}^p b)(x_1 \otimes x_2 \otimes \ldots x_p; y_1 \otimes y_2 \otimes \ldots y_p)&\\ = b(x_1, y_1) b(x_2, y_2) \ldots b(x_p, y_p).&\end{aligned}$$

Les puissances tensorielles $\bar{\otimes}^p \rho$ et $\bar{\otimes}^p \gamma$ des applications associées à droite et à gauche de b sont évidemment les applications associées de $\bar{\otimes}\, pb$, puisque, sur des tenseurs décomposables :

$$\begin{aligned}(\bar{\otimes}^p b)&(x_1 \otimes x_2 \otimes x_2 \otimes \ldots x_p; y_1 \otimes y_2 \otimes \ldots y_p)\\ &= \langle x_1, \rho(y_1)\rangle \langle x_2, \rho(y_2)\rangle \ldots \langle x_p, \rho(y_p)\rangle\\ &= \langle x_1 \otimes x_2 \otimes \ldots x_p; \rho(y_1) \otimes \rho(y_2) \otimes \ldots \rho(y_p)\rangle\\ &= \langle x_1 \otimes x_2 \otimes \ldots x_p; (\bar{\otimes}^p \rho)(y_1 \otimes y_2 \otimes \ldots y_p)\rangle\end{aligned}$$

et la même chose pour γ.

En particularisant encore en prenant $\mathrm{F} = \mathrm{E}$, on voit qu'une forme bilinéaire b sur un espace vectoriel E de dimension finie définit des formes bilinéaires $\bar{\otimes}^p b$ sur toutes les puissances tensorielles de E. De plus, si b est non-dégénérée la forme inverse $\overset{-1}{b}$ sur E^* se prolonge par ses puissances tensorielles $(\bar{\otimes}^p \overset{-1}{b}$ sur les $\otimes^p \mathrm{E}^*$. Sur un espace de tenseurs affines quelconque $\otimes^{(v)}\mathrm{E}$ avec $|v| = m$, produit tensoriel de m espaces E ou E^*, en prenant b sur E et $\overset{-1}{b}$ sur E^*, leur produit tensoriel prolonge b en une forme bilinéaire $\bar{\otimes}^{(v)} b$ sur $\otimes^{(v)}\mathrm{E}$.

EXERCICES SUR LE CHAPITRE II

EX. II.1 :

Soit $t \in$ E$\otimes$E... Montrer que si dans une base de E, il existe un scalaire λ tel que $t^{ji} = \lambda t^{ij}$ quels que soient i, j les mêmes équations sont alors satisfaites dans toutes les bases. Quelles sont les seules valeurs possibles pour λ?

EX. II.2 :

Quelle condition, nécessaire et suffisante, doivent satisfaire les vecteurs $x_1, x_2, \ldots, x_p$ et $y_1, y_2, \ldots, y_p$, supposés non nuls, pour que les tenseurs décomposables $x_1 \otimes x_2 \otimes .x_p$ et $y_1 \otimes y_2 \otimes .y_p$ soient proportionnels.

EX. II.3 :

Soient $\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p$, p espaces vectoriels quelconques sur K. Montrer que si les vecteurs $x_1 \in \mathrm{E}_1$, $x_2 \in \mathrm{E}_2$, ..., $x_p \in \mathrm{E}_p$ sont nuls, leur produit tensoriel $x_1 \otimes x_2 \otimes \ldots p$ est non nul.

EX. II.4 :

Soient $(x_1, x_2, \ldots, x_n)$ et $(y_1, y_2, \ldots, y_n)$, $2n$ vecteurs d'un espace vectoriel à n dimensions.

Montrer que pour que le déterminant : dét$|t^{ij}|$ des composantes du tenseurs : $t = x_1 \otimes y_1 + x_2 \otimes y_2 + \ldots + x_n \otimes y_n$ de E $\times$ E dans une base de e soit nul, il est nécessaire et suffisant que soient les (x_i), soient les (y_i) soient linéairement dépendants.

EX. II.5 :

Soient E et F deux espaces vectoriels sur un même corps K, E_0 un sous-espace de E, F_0 un sous-espace de F. Soit H le sous-espace de E $\otimes$ F engendré par les éléments $x \otimes y$ tels que $x \in \mathrm{E}_0$ *ou bien* $y \in \mathrm{F}_0$.

Construire deux applications linéaires naturelles, réciproques l'une de l'autre, entre $\mathrm{E}/\mathrm{E}_0 \otimes \mathrm{F}/\mathrm{F}_0$ et $(\mathrm{E} \otimes \mathrm{F})/\mathrm{H}$, qui prouvent l'isomorphisme de ces deux espaces vectoriels.

EX. II.6 :

Soient E et F deux espaces vectoriels sur K. Soient $x, x_1, x_2, \ldots, x_p$, $(p+1)$vecteurs de E, $y, y_1, y_2, \ldots, y_p$, $(p+1)$vecteurs de E tels que qu'on ait une égalité :

$$x \otimes y = \sum_{i=1}^{p} x_i \otimes y_i \neq 0.$$

Montrer qu'il en résulte que x est une combinaison linéaire des x_i, et y une combinaison linéaire des y_i.

EX. II.7 :

Soient $a_1, a_2, \dots, a_p$, p applications linéaires, a_j appliquant E_j dans F_j pour $j = 1, 2, \dots, p$.

Les $2p$ espaces vectoriels E_j et F_j étant supposés de dimension finie, montrer que, si le produit tensoriel $a_1 \bar{\otimes} a_2 \dots \bar{\otimes} a_p$ est nul, l'une au moins des applications a_j est nulle.

EX. II.8 :

Soient E un espace vectoriel de dimension finie n sur K et t un tenseur *non nul* de $\otimes^p \mathrm{E}$. Pour chaque $i = 1, 2, \dots, p$, on appelle F_i le sous-espace E^* des formes linéaires α telles que le contracté de α sur la i-ème composantes de t soit nul.

a) montrer que $\dim \mathrm{F}_i \leqslant n - 1$

b) montrer que t est décomposable si et seulement si tous les F_i ont pour dimension $(n-1)$.

c) montrer qu'il suffit que $(p-1)$ des F_i soient de dimension $(n-1)$ pour qu'il en soit de même du p-ème.

EX. II.9 :

Si u et v sont des endomorphismes respectivement des espaces vectoriels de dimension finie pour E et F, montrer que Trace $(u \bar{\otimes} v) = \mathrm{Tr}\, u \cdot \mathrm{Tr}\, v$.

EX. II.10 :

Propriétés du produit tensoriel de matrices. Soient A une matrice $m \times m$ et B une matrice $n \times n$, $A \otimes B$ et $B \otimes A$ leur produits tensoriels (§II.8) dont les lignes et les colonnes sont rangées suivant l'ordre lexicographique sur les couples d'entiers.

a) Montrer qu'il existe une matrice de permutation P ne dépendant pas de A et B mais seulement de m et de n telle que $\mathrm{B} \otimes \mathrm{A} = \mathrm{P}(\mathrm{A} \otimes \mathrm{B}) \overset{-1}{\mathrm{P}}$.

b) Si I_m et I_n désignent les matrices unités $m \times m$ et $n \times n$, en utilisant le fait que $A \otimes B$ est égal au produit de matrices : $(A \otimes \mathrm{I}_n) \cdot (\mathrm{I}_m \otimes \mathrm{B})$ montrer que $\text{dét}(\mathrm{A} \otimes \mathrm{B}) = (\text{dét}\, \mathrm{A})^n (\text{dét}\, \mathrm{B})^m$.

c) Montrer que si A et B sont diagonalisables il en est de même de $A \otimes B$.

d) Montrer que si A et B sont des matrices orthogonales, resp. unitaires, il en est de même de $\mathrm{A} \otimes \mathrm{B}$.

CHAPITRE III

Algèbre tensorielle : les espaces tensoriels, les tenseurs affines et les pseudotenseurs associés à un espace vectoriel de dimension finie

L'instrument construit au chapitre II, le produit tensoriel de deux espaces vectoriels, nous permet maintenant de définir et d'étudier les espaces de tenseurs et de pseudotenseurs *affines* sur un espace vectoriel E de dimension finie. Un tel espace T est lié à E par une correspondance qui à chaque base e de E associe une base de T (d'où les « composantes » de chaque $t \in \mathrm{T}$ que détermine le choix de la base e) et par la présence d'une représentation dans T du groupe linéaire Gl(E) agissant comme « groupe de structure ». Le terme *affine* signifie que E est « nu », sans structure supplémentaire. Le groupe d'automorphismes est donc celui de tous les automorphismes linéaires de E. Les chapitres V, VII et VIII examineront les cas où E est enrichi d'un produit scalaire, qui permet d'identifier E et E*. C'est alors le sous-groupe G de Gl(E) des automorphismes qui le préservent qui sera le groupe de structure.

Les paragraphes 11 et 12 sont consacrés à la structure d'algèbre, le paragraphe 13 aux problèmes universels, le paragraphe 14 aux invariants.

III.1 – Espaces de tenseurs affines sur un espace vectoriel de dimension finie

Dans ce paragraphe et les suivants jusqu'au §III.10 inclus , E est un espace vectoriel *de dimension finie*, car on va faire intervenir systématiquement son dual E^*.

Un *symbole de variance* (v) est une suite de points et d'astéristiques, chaque point représentant l'espace vectoriel E et chaque astérisque son espace dual E^*.

(v) est dit de type (p, q) et d'ordre $\|v\| = p + q$ lorsqu'il contient p points et q astérisques. $\mathrm{X}^{(v)}\mathrm{E}$ représente le produit d'un certain nombre de copies de E et de E^* suivant la formule définie par (v). Par exemple, $\mathrm{X}^{(\cdot *)}\mathrm{E} = \mathrm{E} \times \mathrm{E}^*$.

Un produit tensoriel d'un certain nombre de copies de E ou de E^* suivant la formule définie par (v) est appelé un *espace tensoriel (affine)* de variance (v) sur E et noté $\otimes^{(v)}\mathrm{E}$. On dit aussi que $\otimes^{(v)}\mathrm{E}$ est la puissance tensorielle (v)-ème de E. Les $\otimes^{(v)}\mathrm{E}$, pour toutes les variances (v), forment l'ensemble de tous les espaces tensoriels sur E.

Pour chaque espace tensoriel $\otimes^{(v)}\mathrm{E}$, on a une application $\|v\|$-linéaire structurelle, de $\mathrm{X}^{(v)}\mathrm{E}$ dans $\otimes^{(v)}\mathrm{E}$, dont l'image est l'ensemble des éléments décomposables de $\otimes^{(v)}\mathrm{E}$. Si e est une base de E et e^* sa base duale, $e^{(v)}$ ou $\otimes^{(v)}e$ désigne la base correspondante de $\otimes^{(v)}\mathrm{E}$.

Exemples : On a vu au chapitre II, $\otimes^{(**)}\mathrm{E} = \mathrm{E}^* \otimes \mathrm{E}^*$, $\otimes^{(\cdot *)} = \mathrm{E} \otimes \mathrm{E}^*$. Si $(v) = (..*)$, $e^{(v)}$ est l'ensemble des $(e_i \otimes e_j \otimes e^{*k})$.

Les éléments d'un espace $\otimes^{(v)}\mathrm{E}$ sont appelés *tenseurs affines de variances* (v) *sur* E, d'ordre $\|v\|$, p fois contravariants et q fois covariants si (v) est de type (p,q). Un espace de tenseurs p fois contravariants est noté simplement $\otimes^p\mathrm{E}$, ou puissance tensorielle p-ème de E.

Les composantes d'un tenseur t de $\otimes^{(v)}\mathrm{E}$ dans la base $e^{(v)}$ sont appelées *composantes de t dans e* et s'écrivent avec p indices en haut (contravariants) et q indices en base (covariants) les indices étant placés dans l'ordre de (v).

Exemple : si $(v) = (\cdot\cdot{*}\cdot)$, $t = \sum(e_i \otimes e_j \otimes e^{*k} \otimes e_l)t_k^{ijl}$.

L'ordre des composantes de même nature dans $\otimes^{(v)}\mathrm{E}$ est essentiel... On verra que l'ordre de *toutes* les composantes est essentiel pour les *tenseurs restreints* (cf. Chapitre IV). Par contre, tant qu'il ne s'agit *que de tenseurs affines, et seulement dans ce cas*, on peut négliger l'ordre relatif des composantes de nature différente, et identifier ainsi tous les $\otimes^{(v)}\mathrm{E}$ dont les variances (v) sont de même type (p,q), ce qui revient à ne considérer que des *produits tensoriels notés* $\otimes_{p,q}\mathrm{E} = \underbrace{\mathrm{E}\otimes\ldots\otimes\mathrm{E}}_{p\text{ fois}}\otimes\underbrace{\mathrm{E}^*\otimes\ldots\otimes\mathrm{E}^*}_{q\text{ fois}}$ en plaçant indifféremment les facteurs E devant et les E^* ensuite, ou l'inverse.

Les composantes dans une base e d'un tenseur t de type (p,q) s'écrivent alors : $t_{i_1 i_2\ldots i_p, j_1 j_2\ldots j_q}$. Lors d'un changement de base $e' = e\mathrm{S}$ soit :
$e'_j = \sum e_i s^i_j$; $e'^{*i} = \sum \check{s}^i_k e^{*k}$, les nouvelles composantes de t s'expriment en fonction des anciennes par :

$$(\mathcal{T}) \qquad \boxed{t'^{i_1 i_2\ldots i_p}_{j_1 j_2\ldots j_q} = \sum t^{k_1 k_2\ldots k_p}_{l_1 l_2\ldots l_q} \quad \check{s}^{i_1}_{k_1}\ldots\check{s}^{i_p}_{k_1} s^{l_1}_{j_1}\ldots s^{l_q}_{j_q}}\,.$$

On retrouve la formule du §II.6 qui est maintenant démontrée au lieu d'être prise pour définition.

III.2. – Opérations tensorielles

Les espaces tensoriels sur un espace vectoriel possèdent une structure très riche. Outre les opérations tensorielles, multiplications et contractions par le crochet de dualité entre E et E^* (§II.10), qui les relient entre, eux, le fait que les facteurs d'un $\otimes^{(v)}\mathrm{E}$ soient tous égaux à E ou à E* introduit des opérations nouvelles :

a) le groupe linéaire Gl(E) opère naturellement sur E, et sur E* par la représentation contragrédiente, donc sur tous les facteurs d'un produit tensoriel $\otimes^{(v)}\mathrm{E}$. Le produit tensoriel des opérateurs sur E et sur E* fait opérer Gl(E) sur $\otimes^{(v)}\mathrm{E}$. Les multiplications et contractions commutent avec les opérations de Gl(E) sur les $\otimes^{(v)}\mathrm{E}$.

b) les permutations de facteurs de même nature, E ou E*, opèrent sur les $\otimes^{(v)}\mathrm{E}$. Par exemple, le groupe de permutation $\mathfrak{G}_{(p)}$ de p objets opère sur $\otimes^{p}(\mathrm{E})$. Le sous-espaces de tenseurs qui présentent un caractère de symétrie maximum relativement à $\mathfrak{G}_{(p)}$ sont aussi les sous-espaces irréductibles pour l'action de Gl(E) (H. Weyl).

Nous allons examiner ces opérations en commençant par les multiplications et les contractions.

1) Si (v) et (w) sont deux variances, désignons par (vw) la variance obtenue en plaçant w à la suite de v. La *multiplication tensorielle*, définie au §II.9 applique $\otimes^{(v)}\mathrm{E} \times \otimes^{(w)}\mathrm{E}$ dans $\otimes^{(vw)}\mathrm{E}$. Si l'on identifie les variances de même type, la multiplication tensorielle s'écrit :

$$\otimes^{p}_{q}\mathrm{E} \times \otimes^{r}_{s} \to \otimes^{p+r}_{q+s}\mathrm{E} :$$

Les composantes du produit de deux tenseurs t et u sont les produits, dans l'ordre, des composantes de t et de u :

$$(tu)^{i_1 \cdots i_{p+r}}_{j_1 \cdots j_{q+s}} = t^{i_1 \cdots i_p}_{j_1 \cdots j_q} u^{i_{p+1} \cdots i_{p+r}}_{j_{q+1} \cdots j_{q+s}}.$$

Ce produit n'est évidemment *pas* commutatif.

2) La *contraction* entre le k-ème facteur E et le $\mathfrak{l}$-ème facteur E*, ou, en utilisant le langage des composantes, entre le k-ème indice contravariant et le $\mathfrak{l}$-ème indice covariant est une application linéaire $c^k_{\mathfrak{l}}$:
$\otimes^{(v)}\mathrm{E} \text{ —— } c^k_{\mathfrak{l}} \longrightarrow \otimes^{(w)}\mathrm{E}$, où (w) est (v) moins les $k^{\text{ème}}$ et les $\mathfrak{l}^{\text{ème}}$ indices, en particulier : $\otimes^{p}_{q}\mathrm{E} \text{ —— } c^k_{\mathfrak{l}} \longrightarrow \otimes^{p-1}_{q-1}\mathrm{E}$.

Les composantes du tenseur contracté $(c_l^k t)$ se calculent immédiatement :

$$(c_l^k t)^{i_1 \ldots i_{p-1}}_{j_1 \ldots j_{q-1}} = \sum t^{i_1 \ldots i_{k-1} m i_{k+1} \ldots i_p}_{j_1 \ldots j_{l-1} m j_{l+1} \ldots j_q}.$$

Si $k' \neq k$ et $l' \neq l$, les contractions c_l^k et $c_{l'}^{k'}$ commutent : $c_l^k c_{l'}^{k'} t = c_{l'}^{k'} c_l^k t$.

3) La composée d'une multiplication tensorielle et de contractions sur les indices des facteurs est une *multiplication contractée*. En particulier si l'on multiplie un tenseur t de type (p, q) par p formes $\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(p)}$ et q vecteurs $x_{(1)}, x_{(2)}, \ldots, x_{(q)}$ et que l'on contracte $(p+q)$fois : pour $i = 1, 2, \ldots, p$, le i-ème indice contravariant avec $\alpha_{(i)}$, pour $j = 1, 2, \ldots, q$ le j-ème indice covariant avec $x_{(j)}$, on obtient un scalaire : c'est la *saturation* de t par les $\alpha_{(i)}$ et les $x_{(j)}$.

Si (v^*) désigne la variance duale de (v) obtenue en remplaçant partout E par E^* et E^* par E (par exemple, $(v) = (..*.)$, $(v^*) = (**.*)$, les espaces tensoriels $\otimes^{(v)}\mathrm{E}$ et $\otimes^{(v*)}\mathrm{E}$ sont duaux l'un de l'autre. Le crochet de dualité s'obtient par multiplication tensorielle suivie de la contraction sur toutes les paires d'indices associés de (v) et de (v^*).

Lorsqu'on identifie les variances de même type, on voit ainsi que $\otimes_{p,q}\mathrm{E}$ et $\otimes_{q,p}\mathrm{E}$ sont naturellement duaux l'un de l'autre. En particulier $\otimes_{p,p}\mathrm{E}$ est *auto-dual*, ce qui signifie qu'il existe sur $\otimes_{p,p}\mathrm{E}$ une forme bilinéaire naturelle qui l'identifie à son dual.

Déterminons cette forme. Pour $p = 1$, $\otimes_{1,1}\mathrm{E} = \mathrm{E} \otimes \mathrm{E}^*$ s'identifie à l'espace vectoriel $\mathfrak{L}(\mathrm{E})$ des opérateurs linéaires de E. Le crochet de dualité est la double contraction :

$$(\mathrm{E} \otimes \mathrm{E}^*) \times (\mathrm{E} \otimes \mathrm{E}^*) \to \mathrm{K}.$$

La première contraction est la composition des opérateurs, la seconde est la trace, et le crochet de dualité est donc :

$$\boxed{\forall a, b \in \mathrm{E} \otimes \mathrm{E}^* : \ \langle a, b \rangle = \mathrm{Tr}(a \circ b)}.$$

En particulier pour des éléments décomposables : $a = x \otimes \alpha$ et $b = y \otimes \beta$, $a \circ b = \langle \alpha, y \rangle x \otimes \beta$ et $\langle x \otimes \alpha, y \otimes \beta \rangle = \langle \alpha, y \rangle \langle \beta, x \rangle$.

Dans le cas général, $\otimes_{p,p}\mathrm{E} = (\otimes^p\mathrm{E}) \otimes (\otimes^p\mathrm{E}^*)$ s'identifie à l'algèbre des opérateurs linéaires de $\otimes^p\mathrm{E}$ et la même expression du crochet subsiste.

Application : Autres invariants d'un opérateur linéaire $a \in \mathfrak{L}(\mathrm{E})$. Par l'identification de $\mathfrak{L}(\mathrm{E})$ à $\mathrm{E}\otimes\mathrm{E}^*$, a écrit $a = \sum\limits_j x_j \otimes \alpha_j$, $\bar{\otimes}\,{}^p a$ s'identifie à

$$\begin{aligned}\otimes^p a = \sum_{j_1\cdots j_p} x_{j_1}\otimes x_{j_2}\otimes\ldots x_{j_p}\otimes\alpha_{j_1}\otimes\alpha_{j_2}\otimes\ldots x_{j_p} \in \otimes^p_p\mathrm{E}\\ = (\otimes^p\mathrm{E})\otimes(\otimes^p\mathrm{E}^*).\end{aligned}$$

Si C^r_s désigne la contraction de r-ème indice contravariant et du s-ème indice covariant, $\mathrm{C}^2_1\mathrm{C}^3_2\ldots\mathrm{C}^p_{p-1}(\otimes^p a)$ n'est autre que $a^p \in \mathrm{E}\otimes\mathrm{E}^*$, et si l'on effectue alors la contraction C^1_p on obtient la trace $\mathrm{Tr}(a^p)$ tandis que $\mathrm{C}^1_1\mathrm{C}^2_2\ldots\mathrm{C}^p_p(\otimes^p a) = \mathrm{Tr}(\otimes^p a) = (\mathrm{Tr}\,a)^p$.

L'opérateur identique de E s'écrit dans $\mathrm{E}\otimes\mathrm{E}^*$ $\delta = \sum\limits_i e_i \otimes e^{*i}$ quelle que soit la base e et ses composantes sont les *symboles de Kronecker* $\delta^i_j = 0$ si $i \neq j$, $= 1$ si $i = j$. Mais le dual de $\mathrm{E}\otimes\mathrm{E}^*$ s'identifiant à $\mathrm{E}\otimes\mathrm{E}^*$, les éléments de $\mathrm{E}\otimes\mathrm{E}^*$ peuvent être interprétés comme des formes linéaires sur $\mathrm{E}\otimes\mathrm{E}^*$, donc comme des formes bilinéaires sur $\mathrm{E}\times\mathrm{E}^*$. Si $a \in \mathfrak{L}(\mathrm{E})$ la forme bilinéaire associée est tout simplement : $(x,\alpha) \to \langle\alpha, ax\rangle = \langle{}^t a\alpha, x\rangle$.

En particulier, *l'opérateur identique* δ *de* E *s'identifie de cette façon au crochet de dualité entre* E *et* E^*.

L'espace tensoriel $\otimes^p\mathrm{E}^*$, s'identifie à l'espace vectoriel des formes p-linéaires sur E. En effet, le diagramme universel :

$$\begin{tikzcd} \mathrm{E}\times\mathrm{E}\times\ldots\mathrm{E} \arrow[rr, "f"] \arrow[dr, "u"'] & & \mathrm{K} \\ & \mathrm{E}\otimes\mathrm{E}\otimes\ldots\mathrm{E} \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

permet d'identifier les formes p-linéaires f sur E aux formes linéaires $\widetilde{f}$, sur $\otimes^p\mathrm{E}$, c'est-à-dire aux éléments du dual $(\otimes^p\mathrm{E})^*$ qui lui-même s'identifie comme on vient de le voir à $\otimes^p\mathrm{E}^*$.

Le crochet de dualité entre $\otimes^p\mathrm{E}^*$ et $\otimes^p\mathrm{E}$ s'écrit pour des éléments décomposables de chaque tensoriel :

$$\langle \alpha_{(1)} \otimes \alpha_{(2)} \otimes \dots \alpha_{(p)}, x_{(1)} \otimes x_{(2)} \otimes \dots x_{(p)} \rangle$$
$$= \langle \alpha_{(1)}, x_{(1)} \rangle \langle \alpha_{(2)}, x_{(2)} \rangle \dots \langle \alpha_{(p)}, x_{(p)} \rangle.$$

Si $e = (e_j; j \in \mathrm{J})$ est une base de E, et e^* sa base duale, l'expression d'une forme p-linéaire f sur E :

$$f(x_{(1)}, \dots, x_{(p)}) = \sum f_{i_1 \dots i_p} x_{i_1,(1)} \cdots x_{i_p,(p)}$$
$$= \langle \sum f_{i_1 \dots i_p} e^{*i_1} \otimes \dots e^{*i_p}, x_{(1)} \otimes \dots x_{(p)} \rangle$$

montre bien, comme au §II.2, cette idenfication.

4) Une application : *critère de tensorialité*. On rencontre fréquemment des ensembles de « composantes », donc munies d'indices variant librement de 1 à n : $t_{i_1 \dots i_p, j_1 \dots j_q}$, qui *ne sont pas* les composants d'un tenseur.

Nous nous proposons de donner quelques tests permettant d'établir la nature tensorielle de tels ensembles.

Premier critère : on effectue un changement de base et l'on vérifie que les composantes se transforment selon l'une des formules du § III.1. Les calculs peuvent être pénibles, et il vaut mieux les éviter si l'on peut utiliser l'un des critères suivants.

Second critère dit de saturation : si quels que soient les vecteurs $x_{(1)}, x_{(2)}, \dots x_{(q)}$ de composantes dans une base e : $x^{j_1}_{(1)}, x^{j_2}_{(2)}, \dots$, $x^{j_q}_{(q)}$ et les covecteurs $\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(p)}$ de composantes $\alpha_{(1)i_1}$, $\alpha_{(2)i_2}, \dots, \alpha_{(p)i_p}$ dans e, la somme :

$$\sum t^{i_1 i_2 \dots i_p}_{j_1 j_2 \dots j_q} \alpha_{(1)i_1} \dots \alpha_{(p)i_p} x^{j_1}_{(1)} \dots x^{j_q}_{(q)}$$

est une constante indépendante du choix de la base e, alors les $t^{i_1 \dots i_p}_{j_1 \dots j_q}$ sont celles d'un tenseur p fois contravariant et q fois covariant.

La preuve est immédiate. Choisissons en effet les vecteurs $x_{(k)} = e'_k$ et les covecteurs $\alpha_{(1)} = e'^{*l}$; $k = 1, 2, \dots, q$; $l = 1, 2, \dots, p$; où e' est une autre base. La constance de la somme n'est autre que la formule de transformation des composantes d'un tenseur.

Troisième critère, dit du quotient : si quel que soit le tenseur u de type (r, s), la contraction :

$$v^{i_1 \dots i_p}_{j_1 \dots j_q} = \sum t^{i_1 \dots i_p i_{p+1} \dots i_{p+s}}_{j_1 \dots j_q j_{q+1} \dots j_{q+r}} u^{j_{q+1} \dots j_{q+r}}_{i_{p+1} \dots i_{p+s}},$$

sur $(r + s)$couples d'indices donne les composantes d'un tenseur V, alors les $t^{i_1 \dots i_{p+s}}_{j_1 \dots j_{q+r}}$ sont les composantes d'un tenseur. En effet on peut prendre en particulier pour u le produit tensoriel de s vecteurs et de r formes arbitrairement choisis. En saturant alors les deux membres par contraction avec q vecteurs et p formes arbitrairement choisis, on retombe sur le critère de saturation précédent.

III.3 – Permutations des facteurs d'un espace tensoriel $\otimes^{(v)}\mathrm{E}$ et action du groupe symétrique

Nous allons définir les opérations de permutation des facteurs des produits tensoriels qui généralisent le cas élémentaire de la transposition dans le produits $\mathrm{E} \otimes \mathrm{E}$ ou $\mathrm{E}^* \otimes \mathrm{E}^*$ (§II.3)

Une permutation σ de $(1, 2, \dots, p)$ est une application bijective de cet ensemble sur lui-même que l'on visualise par le tableau :

$$\begin{pmatrix} 1 & 2 \dots p \\ \sigma_1 & \sigma_2 \dots \sigma_p \end{pmatrix} \quad \text{qui se lit : } 1 \to \sigma_1, 2 \to \sigma_2, \text{ etc.}$$

Les permutations de $(1, 2, \dots, p)$ forment un groupe : le groupe *symétrique* d'indice p qu'il est d'usage de noter par la lettre s gothique majuscule : $\mathfrak{S}_p$. Toute permutation σ peut s'écrire d'une infinité de façons comme le produit de N transpositions, la parité de N étant toujours la même, ce qui permet d'assigner à chaque σ le nombre $\varepsilon(\sigma) = +1$ ou -1 suivant que σ est le produit d'un nombre pair (permutation « paire ») ou impair (permutation « impaire ») de transpositions.

ε est la *signature* et $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. La signature est ainsi un homomorphisme du groupe $\mathfrak{S}_p$ sur le groupe multiplicatif à deux éléments $(1, -1)$.

Considérons maintenant des objets rangés en une suite, comme les facteurs d'un produit par exemple. L'objet d'une permutation sur ces objets va être d'en changer l'ordre, et la *permutation va porter* *sur les places (ou rangs) qu'occupent ces objets, indépendamment des indices dont ils peuvent être munis*. La convention universellement adoptée est que σ *envoie l'objet qui occupe* j *au rang* σj.

Il en résulte qu'après l'opération de σ, l'objet qui se trouve au rang j est celui qui se trouvait au rang $\overset{-1}{\sigma} j$. On a donc :

$$(y_1, y_2, \ldots, y_p) = \widetilde{\sigma}(x_1, x_2, \ldots, x_p) = (x_{\overset{-1}{\sigma} 1}, x_{\overset{-1}{\sigma} 2}, \ldots x_{\overset{-1}{\sigma} p}).$$

Si l'on effectue maintenant une seconde permutation τ sur la suite obtenue, l'objet, qui du rang j avait été envoyé au rang σj, est maintenant envoyé au rang $\tau(\sigma j) = (\tau\sigma)j$ et $\widetilde{\tau} \circ \widetilde{\sigma} = (\widetilde{\tau\sigma})$.

En effet, $\widetilde{\tau}(\widetilde{\sigma}(x_1, x_2, \ldots, x_p)) = \widetilde{\tau}(y_1, y_2, \ldots, y_p)$, avec $y_j = x_{\overset{-1}{\sigma} j}$,

$$(y_{\overset{-1}{\tau} 1}, y_{\overset{-1}{\tau} 2}, \ldots, y_{\overset{-1}{\tau} p})$$
$$= (x_{\overset{-1}{\sigma}\overset{-1}{\tau} 1}, \ldots, x_{\overset{-1}{\sigma}\overset{-1}{\tau} p}) = (\widetilde{\tau\sigma})(x_1, x_2, \ldots, x_p).$$

$\mathfrak{S}$ opère également *sur les applications*. Si f est une application qui au p-uple $(x_1, x_2, \ldots, x_p)$ d'objets d'un même ensemble E fait correspondre $f(x_1, x_2, \ldots, x_p) \in$ F, $(\sigma f)(x_1 x_2, \ldots, x_p) = f(\overset{-1}{\widetilde{\sigma}}(x_1, x_2, \ldots, x_p)) = f(x_{\sigma 1}, x_{\sigma 2}, \ldots, x_{\sigma p})$. On a bien $\tau(\sigma f)(x_1 x_2, \ldots, x_p) = (\sigma f)(x_{\tau 1}, x_{\tau 2}, \ldots, x_{\tau p}) = (\sigma f)(y_1, y_2, \ldots, y_p)$, (avec $y_j = x_{\tau j}$) $= f(y_{\sigma 1}, y_{\sigma 2}, \ldots, y_{\sigma p}) = f(x_{\tau\sigma 1}, x_{\tau\sigma 2}, \ldots, x_{\tau\sigma p}) = ((\tau\sigma)f)(x_1, x_2, \ldots, x_p)$. (Voir la fin de ce paragraphe pour une application).

Soient (v) une variance d'ordre $|v| = m$, σ une permutation de $(1, 2, \ldots, m)$ *qui conserve* (v). L'application m-linéaire :

$$\mathrm{X}^{(v)}\mathrm{E} \longrightarrow \otimes^{(v)}\mathrm{E},$$
$$(x_1, x_2, \ldots, x_m) \longrightarrow x_{\overset{-1}{\sigma} 1} \otimes x_{\overset{-1}{\sigma} 2} \otimes \ldots x_{\overset{-1}{\sigma} m})$$

où x_j est un élément du $j^{ème}$ facteur, E ou E*, de $X^{(v)}(\mathrm{E})$; détermine un endomorphisme $\widetilde{\sigma}$ de $\otimes^{(v)}\mathrm{E}$. L'image par $\widetilde{\sigma}$ d'un élément décomposable s'écrit :

$$\boxed{\widetilde{\sigma}(x_1 \otimes x_2 \otimes \ldots x_m) = x_{\overset{-1}{\sigma} 1} \otimes x_{\overset{-1}{\sigma} 2} \otimes \ldots x_{\overset{-1}{\sigma} m}}.$$

Si τ est une autre permutation de $(1, 2, \ldots, m)$ qui conserve (v), on a, d'après ce qui précède, $(\widetilde{\tau\circ\sigma}) = \widetilde{\tau}\circ\widetilde{\sigma}$. En prenant $\tau = \overset{-1}{\sigma}$, on obtient :
$\widetilde{(\overset{-1}{\sigma})} = (\overset{-1}{\widetilde{\sigma}})$ et le fait que $\widetilde{\sigma}$ *est un automorphisme* de l'espace vectoriel $\otimes^{(v)}\mathrm{E}$. Sur les espaces tensoriels $\otimes^p\mathrm{E}$ et $\otimes^p\mathrm{E}^*$ chaque élément σ du groupe symétrique $\mathfrak{S}_p$ opère par automorphisme $\widetilde{\sigma}$, et un élément t de l'un de ces espaces tensoriels est dit *symétrique* si $\widetilde{\sigma}t = t$, $\forall\sigma \in \mathfrak{S}_p$, *antisymétrique* si $\widetilde{\sigma}t = \varepsilon(\sigma)t$ $\forall\sigma \in \mathfrak{S}_p$, où $\varepsilon(\sigma)$ désigne la signature (± 1) de la permutation σ.

L'application $\sigma \in \mathfrak{S}_p \Longrightarrow \widetilde{\sigma} \in \mathrm{Gl}(\otimes^p\mathrm{E})$ (ou $\mathrm{Gl}(\otimes^p\mathrm{E}^*)$) est un homomorphisme de groupes. C'est donc une représentation linéaire du groupe $\mathfrak{S}_p$ (cf. III.3).

Définition III.3 :

Soient F un espace vectoriel sur lequel le groupe symétrique $\mathfrak{S}_p$ opère linéairement par :
$\sigma \in \mathfrak{S}_p \Longrightarrow \widetilde{\sigma} \in \mathrm{Gl}(\mathrm{F})$, V un espace vectoriel sur le même corps, et f une application linéaire de F dans V. f est alors dite *symétrique* si $f(\widetilde{\sigma}u) = f(u)$, *antisymétrique* si $f(\widetilde{\sigma}u) = \varepsilon(\sigma)f(u)$, $\forall\sigma \in \mathfrak{S}_p$ et $u \in \mathrm{F}$.

Remarque : Si $\langle f, t\rangle$ est le crochet de dualité défini ci-dessus entre $f \in \otimes^p\mathrm{E}^*$ et $t \in \otimes^p\mathrm{E}$, on a évidemment $\langle\widetilde{\sigma}f, \widetilde{\sigma}t\rangle = \langle f, t\rangle$ quel que soit $\sigma \in \mathfrak{S}_p$, ce qui écrit $\langle\widetilde{\sigma}f, t\rangle = \langle f, \widetilde{\overset{-1}{\sigma}}t\rangle$. Cette dernière égalité prouve un point important. En effet un élément f de $\otimes^p\mathrm{E}^*$ peut ou bien s'interpréter comme tenseur sur E^*, et $\mathfrak{S}_p$ opère par permutation des facteurs ou bien comme forme p-linéaire sur E, avec, cette fois, une opération de $\mathfrak{S}_p$ sur les arguments (voir ci-dessus). L'égalité précédente montre que ces opérations sont identiques.

III.4 – Représentations tensorielles du groupe linéaire

On a vu aux paragraphes II.8. et II.9. la définition du produit tensoriel de représentations K-linéaires d'un groupe G : quelle que soit l'écriture d'un élément t du produit tensoriel d'espaces

$E_1, E_2, \ldots, E_p$, où opère G, l'action de $g \in G$ sur t s'obtient en distribuant l'action de g dans tous les termes de l'expression de t. Si $\otimes^{(v)}E$ est un espace tensoriel sur E, Gl(E) opère sur E, et sur E^*par la représentation contragrédiente. On obtient ainsi la :

Définition III.4.A :

On appelle représentation tensorielle de variance (v) du groupe linéaire Gl(E) d'un espace vectoriel E le produit tensoriel des représentations, naturelle sur E, contragrédiente sur E^*, de Gl(E) dans les facteurs E ou E^* de l'espace tensoriel $\otimes^{(v)}E$. Si par exemple :

$$t = \sum t^{i_1 \ldots i_p}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots e_{i_p} \otimes e^{*j_1} \otimes \ldots e^{*j_q},$$

$$g \cdot t = \sum t^{i_1 \ldots i_p}_{j_1 \ldots j_q} (ge_{i_1}) \otimes \ldots (ge_{i_p}) \otimes (\check{g}e^{*j_1}) \otimes \ldots (\check{g}e^{*j_q}).$$

Les composantes de $g \cdot t$ dans la base e sont donc celles de t dans la base $\overset{-1}{g}\, e$, ce qui permet de définir directement l'action de Gl(E) en ne se servant que des composantes.

Si $s \in \mathrm{Gl(E)}$ a pour matrice S dans la base e, *les composantes de t dans la base $e' = e\mathrm{S}$ sont celles de $\overset{-1}{s}\, t$ dans la base e : les formules de changement de base ne sont qu'un aspect particulier de la structure de* Gl(E)*-module.*

Remarques :

1) Les éléments de la matrice de $\bar{\otimes}^p g$ relativement à la base $\otimes^p e$ sont des monômes par rapport aux éléments g^i_j de la matrice de g dans e; les éléments de la matrice de $\check{g} = {}^t\overset{-1}{g}$ s'expriment en fonction des g^i_j par une fraction rationnelle avec au dénominateur le déterminant des g^i_j. Il en résulte que les éléments de matrice des représentations tensorielles de Gl(E) sont des fractions rationnelles par rapport aux g^i_j; ce sont des *représentations rationnelles de* Gl(E).

2) g et $\overset{-1}{g}$ sont des éléments de Gl(E), donc de $\mathfrak{L}$(E) qui s'identifie à $E \otimes E^*$. L'action de g sur x ou de ${}^t\overset{-1}{g} = \check{g}$ sur α revient (§II.9) à une multiplication contractée. En particulier, l'action de

Gl(E) sur l'espace vectoriel $\mathrm{E} \otimes \mathrm{E}^*$ des opérateurs linéaires de E s'écrit au moyen de la composition des opérateurs :
si $a = \mathfrak{L}(\mathrm{E}) = \mathrm{E} \otimes \mathrm{E}^*$, soit $a = \sum\limits_i x_{(i)} \otimes \alpha_{(i)}$,

$$(g \otimes \check{g}) \cdot a = \sum_i g \cdot x_{(i)} \otimes g \cdot \alpha_{(i)} = g \circ (\sum_i x_{(i)} \otimes \alpha_{(i)}) \circ \overset{-1}{g} = g \circ a \circ \overset{-1}{g}.$$

3) *Le premier problème* de la théorie des tenseurs affines *est la décomposition des espaces tensoriels* $\otimes^p \mathrm{E}$ *sous l'action du groupe linéaire* Gl(E) *en sous-espaces stables minimaux définissant des sous-représentations irréductibles de* Gl(E).

Ces sous-espaces sont tout simplement formés de tenseurs présentant des propriétés de symétrie maximale, celles-ci formant une suite d'intermédiaires entre la symétrie totale à une extrêmité et l'antisymétrie totale à l'autre. (Leur définition n'a rien d'élémentaire !)

La décomposition a pour point de départ le théorème général suivant :

Théorème III.4.A : Les opérations tensorielles sur les espaces tensoriels $\otimes^{(v)}\mathrm{E}$ sur E : multiplication tensorielle, contraction, permutation de facteurs, commutent avec les opérations du groupe linéaire Gl(E) dans ces espaces. En particulier, les contractions et permutations de facteurs sont des Gl(E)-morphismes.

Preuve : pour la multiplication tensorielle et la permutation de facteurs, c'est une conséquence du fait que l'opération de $g \in \mathrm{Gl(E)}$ sur les éléments des $\otimes^{(v)}\mathrm{E}$ consiste à distribuer g ou $\check{g}$ sur chaque facteur E ou E^*.

Pour la contraction, elle vient du simple fait que, par définition de la contragrédiente, le crochet de dualité est invariant. Si $\alpha \otimes \mathrm{E}^*$ et $x \in \mathrm{E}$:

$$\langle \check{g}\alpha, g\alpha \rangle = \langle {}^t \overset{-1}{g} \alpha, g\alpha \rangle = \langle \alpha, \overset{-1}{g} g\alpha \rangle = \langle \alpha, x \rangle.$$

On a donc, pour tout $t \in \otimes^{(v)}\mathrm{E}$, $u \in \otimes^{(w)}\mathrm{E}$:

$$g \cdot t \otimes g \cdot u = g \cdot (t \otimes u); \widetilde{\sigma}(g \cdot t) = g \cdot \widetilde{\sigma}(t); c_{ij}(g \cdot t) = g \cdot c_{ij}(t).$$

Comme cas particulier de ce théorème, on a le :

Théorème III.4.B : Sur $\otimes^p$E, les opérations du groupe de permutation des facteurs $\mathfrak{S}_p$ commutent avec les opérations du groupe linéaire Gl(E).

Dans les applications du calcul tensoriel, ce sont la plupart du temps des sous-espaces ou des espaces quotients des espaces tensoriels $\otimes^{(v)}$E ainsi des produits tensoriels de ces espaces, qui interviennent.

Définition III.4.B :

Nous appellerons *espaces de tenseurs* (affines) sur un espace vectoriel E tout sous-espace vectoriel d'une puissance tensorielle (v)-ème de E ; $\otimes^{(v)}$E, stable par les opérateurs de permutation de facteurs ainsi que par celles du groupe linéaire Gl(E), ainsi que tout espace quotient de $\otimes^{(v)}$E par un tel sous-espace.

Exemples : Les sous-espaces des tenseurs symétriques, ou antisymétriques de $\otimes^p$E, les puissances extérieures et symétriques de E définies au chapitre suivant.

III.5 – Symétriseurs, antisymétriques et opérateurs de symétrie

L'analyse des opérations du groupe de permutations $\mathfrak{S}_p$ dans $\otimes^p$E, et en particulier la détermination des sous-espaces de tenseurs présentant un caractère de symétrie maximal, rend nécessaire l'introduction d'un nouveau concept, au demeurant fort naturel. Commençons par le cas le plus simple. Dans E$\otimes$E, un tenseur t est symétrique ou antisymétrique suivant qu'il est invariant ou qu'il change de signe par la transposition $\widetilde{\tau}$ des facteurs. Partant d'un tenseur u quelconque de E$\otimes$E, son « symétrisé » est $\mathrm{s}u = \dfrac{1}{2}\,(u+\widetilde{\tau}u)$, son « antisymétrisé » est $\mathbf{a}u = \dfrac{1}{2}\,(u - \widetilde{\tau}u)$ et $u = su + \mathbf{a}u$. On peut considérer que $s = \dfrac{1}{2}$ (Identité $+\widetilde{\tau}$) et $\mathbf{a} = \dfrac{1}{2}$ (Identité $-\widetilde{\tau}$) sont des opérateurs dans l'espace vectoriel E $\otimes$ E. Dans $\mathfrak{L}$(E $\otimes$ E) le sous-espace vectoriel $\{\lambda l + \mu\widetilde{\tau}; \lambda, \mu \in \mathrm{K}\}$ est une sous-algèbre de dimension deux dont s et **a** forment aussi une base, mais telle que $\mathrm{s}^2 = \mathrm{s}$, $\mathbf{a}^2 = \mathbf{a}$, $\mathrm{s}\mathbf{a} = \mathbf{a}\mathrm{s} = 0$.

Dans $\otimes^p\mathrm{E}$ on va de la même façon considérer des opérateurs :

$$\mathbf{s} = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \widetilde{\sigma} \text{ et } \mathbf{a} = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}} \varepsilon(\sigma)\widetilde{\sigma},$$

de symétrisation et d'antisymétrisation (où $\varepsilon(\sigma) \pm 1$ est la parité de σ). Mais lorsque $p > 2$, ils devront être complétés par d'autres opérateurs de symétrie, de définition beaucoup moins intuitives ! Les propriétés algébriques de ces opérateurs ne font intervenir que $\mathfrak{S}_p$ et sont indépendantes des espaces E. Leur clarification impose qu'on les étudie de façon intrinsèque c'est-à-dire sans l'intervention des espaces sur lesquels ils opèrent.

Cela justifie l'introduction de *l'algèbre du groupe symétrique* $\mathfrak{G}_p$ *sur le corps* K. C'est l'espace vectoriel $\mathrm{K}(\mathfrak{S}_p)$ de base $\mathfrak{G}_p$ que forment les combinaisons linéaires d'éléments de $\mathfrak{S}_p$ à coefficients dans K (il s'identifie à l'espace vectoriel des applications de l'ensemble fini $\mathfrak{S}_p$ dans K). Le produit (associatif) du groupe $\mathfrak{S}_p$ détermine une table de multiplication sur la base de $\mathrm{K}(\mathfrak{S}_p)$, qui, par linéarité définit la structure d'algèbre associative unitaire de $\mathrm{K}(\mathfrak{S}_p)$:

$$\text{si} \quad b = \sum_i \lambda_i \sigma_i \text{ et } c = \sum_j \mu_j \sigma_j, \ bc = \sum_{i,j} \lambda_i \mu_j \sigma_i \sigma_j.$$

Une représentation linéaire ρ du groupe $\mathfrak{S}_p$ dans un espace vectoriel F sur le corps K est un homomorphisme de $\mathfrak{S}_p$ dans le groupe linéaire Gl(F), lui-même contenu dans l'algèbre $\mathfrak{L}(\mathrm{F})$ de tous les opérateurs linéaires de F.

L'application ρ de la base $\mathfrak{S}_p$ dans $\mathfrak{L}(\mathrm{F})$ s'étend par linéarité en une application linéaire de $\mathrm{K}(\mathfrak{S}_p)$ dans $\mathfrak{L}(\mathrm{F})$ qui est un homomorphisme d'algèbres puisque ρ est multiplicatif sur la base $\mathfrak{S}_p$.

Il y a donc une équivalence complète entre représentations du *groupe* $\mathfrak{S}_p$ et représentations de *l'algèbre* $\mathrm{K}(\mathfrak{S}_p)$ dans les espaces vectoriels sur le corps K. On peut donc étudier les opérateurs construits à l'aide des éléments de $\mathfrak{S}_p$ directement dans l'algèbre $\mathrm{K}(\mathfrak{S}_p)$. Ces éléments ont souvent été appelés, en raison du rôle qu'ils jouent, opérateurs de symétrie.

Si τ est un élément quelconque du groupe $\mathfrak{S}_p$, l'application $l(\tau)$ $\sigma \in \mathfrak{S}_p \to \tau\sigma \in \mathfrak{S}_p$, ou translation à gauche par τ, est une bijection de $\mathfrak{S}_p$ sur lui-même, dont l'inverse est $l(\overset{-1}{\tau})$. Il en résulte que les éléments :

$$\mathrm{S} = \sum_{\sigma \in \mathfrak{S}_p} \quad \text{et} \quad \mathrm{A} = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma)\sigma,$$

de $\mathrm{K}(\mathfrak{S}_p)$ possèdent les propriétés suivantes :

a) $\tau\mathrm{S} = \mathrm{S} = \mathrm{S}\tau\,;\ \tau\mathrm{A} = \mathrm{A}\tau = \varepsilon(\tau)\mathrm{A},\ \forall\tau \in \mathfrak{S}_p$.
S et A commutent avec *tous* les éléments de l'algèbre $\mathrm{K}(\mathfrak{S}_p)$: ce sont des éléments du *centre* de cette algèbre.

b) $\mathrm{S}^2 = p!\mathrm{S}\,;\ \mathrm{A}^2 = p!\mathrm{A}\,;\ \sum_{\sigma \in \mathfrak{S}_p} \sigma\mathrm{A} = \left(\sum_\sigma \varepsilon(\sigma)\right) \cdot \mathrm{A} = 0$ et de même $\mathrm{AS} = 0$.
Le *symétriseur* $\mathbf{s} = \dfrac{1}{p!}\,\mathrm{S} = \dfrac{1}{p!}\,\sum \sigma$ et *l'antisymétriseur* $\mathbf{a} = \dfrac{1}{p!}\,\mathrm{A} = \dfrac{1}{p!}\sum \varepsilon(\sigma)\sigma$ possèdent donc les propriétés suivantes :

a) $\tau\mathbf{s} = \mathbf{s}\tau = \mathbf{s}\,;\ \tau\mathbf{a} = \mathbf{a}\tau = \varepsilon(\tau)\mathbf{a},\ \forall\tau \in \mathfrak{S}_p$.

b) $\mathbf{s}^2 = \mathbf{s}$ et $\mathbf{a}^2 = \mathbf{a}\,;\ \mathbf{sa} = \mathbf{a}s = 0$

ce que l'on exprime en disant que s et a sont des idempotents appartenant au centre de l'algèbre $\mathrm{K}(\mathfrak{S}_p)$, « algébriquement orthogonaux » ($\mathbf{sa} = \mathbf{as} = 0$) (Attention ! il y en a d'autres dès que $p > 2$!)

En relation avec un espace tensoriel $\otimes^p\mathrm{E}$ sur lequel ils opèrent, on a les propriétés suivantes :

a) Si u est un tenseur décomposable ayant deux facteurs égaux, la transposition τ de ces facteurs laisse u invariant : $\tau u = u$. Puisque $\varepsilon(\tau) = -1$, $\mathbf{a}u = \mathbf{a}\tau u = -\mathbf{a}u$ d'où $\mathbf{a}u = 0$.

Cette propriété de a d'annuler les tenseur décomposables ayant deux facteurs égaux s'exprime en disant que a est *alterné* : a annule tout élément présentant une symétrie partielle.

b) Si t est un élément symétrique de $\otimes^p\mathrm{E}$: $\sigma t = t$, $\forall\sigma \in \mathfrak{S}_p$ d'où : $\mathbf{s}t = t$. Réciproquement, quel que soit $u \in \otimes^p\mathrm{E}$, $\mathbf{s}u$ est symétrique puisque $\sigma\mathbf{s}u = \mathbf{s}u$. Le sous-espace de $\mathrm{S}^p\mathrm{E}$ des tenseurs symétriques de $\otimes^p\mathrm{E}$ est donc l'image du projecteur $\widetilde{\mathbf{s}}$ de $\mathfrak{L}(\otimes^p\mathrm{E})$.

c) D'une façon analogue, si t est un élément antisymétrique de $\otimes^p\mathrm{E} : \sigma t = \varepsilon(\sigma)t,\ \forall\sigma \in \mathfrak{S}_p,\ \mathbf{a}t = \dfrac{1}{p!}\ \sum \varepsilon(\sigma)^2 t = t$.

Réciproquement, quel que soit $u \in \otimes^p\mathrm{E}$, $\mathbf{a}u$ est antisymétrique puisque $\sigma\mathbf{a}u = \varepsilon(\sigma)\mathbf{a}u$. Le sous-espace $\mathrm{A}^p\mathrm{E}$ des tenseurs antisymétriques de $\otimes^p\mathrm{E}$ est donc l'image du projecteur $\widetilde{\mathbf{a}}$ de $\mathfrak{L}(\otimes^p\mathrm{E})$.

III.6 – Pseudoscalaires

Nous poursuivons ce chapitre par deux paragraphes consacrés respectivement aux pseudoscalaires et aux pseudotenseurs. C'est délibérement que nous les décrivons indépendamment de l'algèbre extérieure : il est plus clair, en effet d'utiliser les tenseurs antisymétriques, en faisant l'économie d'un passage au quotient. Le raccord se fera au chapitre IV.

Nous avons vu que l'antisymétriseur A est alterné. Si $m > n = \dim \mathrm{E}$, tout élément de la base de $\otimes^m \mathrm{E}$ associée à une base donnée de E contient au moins deux facteurs égaux et est donc annulé par A. Il n'y a donc aucun tenseur antisymétrique différent de zéro dans $\otimes^p \mathrm{E}$ si $m > n$. Dans $\otimes^m \mathrm{E}$ le sous-espace vectoriel D des éléments antisymétriques, image de A, est de dimension un. En effet A annule tous les éléments $e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_n}$ d'une base de $\otimes^n \mathrm{E}$ associée à la base e de E sauf ceux dont tous les éléments sont distincts et qui peuvent s'écrire : $e_{\sigma 1} \otimes e_{\sigma 2} \otimes \ldots e_{\sigma n} = \overset{-1}{\sigma} \cdot (e_1 \otimes e_2 \otimes \ldots e_n)$ avec $\sigma \in \mathfrak{S}_n$.

L'image par A d'un tel élément est donc, puisque $\varepsilon(\overset{-1}{\sigma}) = \varepsilon(\sigma)$ et $\mathrm{A}\,\overset{-1}{\sigma} = \varepsilon(\sigma)\mathrm{A}$, égale à $\varepsilon(\sigma)\mathrm{A} \cdot (e_1 \otimes e_2 \otimes \ldots e_n)$:

Ainsi, l'élément $\overline{e} = \mathrm{A} \cdot (e_1 \otimes e_2 \otimes \ldots e_n) = \sum\limits_{\sigma} \varepsilon(\sigma) e_{\sigma 1} \otimes e_{\sigma 2} \otimes \ldots e_{\sigma n}$ non nul puisque tous les termes de la base qui apparaissent dans son expression sont linéairement indépendants, *forme la base associée à* e *de la droite* D *des tenseurs antisymétriques de* $\otimes^n \mathrm{E}$.

Un tenseur v appartenant à D est un multiple scalaire de l'élément de base $v = \lambda \overline{e} = \lambda \mathrm{A}(e_1 \otimes e_2 \otimes \ldots e_n)$ et ses composantes dans e s'écrivent :

$$v^{i_1 i_2 \ldots i_n} = \lambda \varepsilon^{i_1 i_2 \ldots i_n},$$

où $\varepsilon^{i_1 i_2 \ldots i_n}$ est nul si deux indices sont égaux et sinon, est égal à la signature de la permutation qui fait passer de $(1, 2, \ldots, n)$ à $(i_1 i_2 \ldots i_n)$.

Les composantes de v dans e se réduisent donc à la donnée du seul nombre λ.

A chaque tenseur t de $\otimes^n \mathrm{E}$ on peut faire correspondre relativement à la base e de E, le nombre $v_e(t)$ défini par :

$$\boxed{\mathrm{A}t = v_e(t)\mathrm{A}(e_1 \otimes e_2 \otimes \ldots e_n) = v_e(t)\overline{e}}\,.$$

Nous appellerons $v_e(t)$ *le volume algébrique de t relativement à la base ordonnée* $e = (e_1, e_2, \dots, e_n)$.

Le déterminant de n ($n = \dim \mathrm{E}$) *vecteurs* $x_1, x_2, \dots, x_n$ *relativement à* une base ordonnée de E, noté $\mathrm{dét}_e(x_1, x_2, \dots, x_n)$ est le volume algébrique de leur produit tensoriel. Celui-ci s'écrit dans la base e :

$$x_1 \otimes x_2 \otimes \dots x_n = \sum x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} e_{i_1} \otimes e_{i_2} \otimes \dots e i_n.$$

Calculons ce déterminant. A annule les éléments de la base ayant deux facteurs égaux et il reste :

$$\mathrm{A} \cdot x_1 \otimes x_2 \otimes \dots x_n = \sum_{\sigma \in \mathfrak{S}_n} x_1^{\sigma 1} x_2^{\sigma 2} \dots x_n^{\sigma n} \mathrm{A}(e_{\sigma 1} \otimes \dots e_{\sigma n}),$$

d'où : $$\mathrm{dét}_e(x_1, x_2, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_1^{\sigma 1} x_2^{\sigma 2} \dots x_n^{\sigma n}.$$

Si X est la matrice dont la j-ème colonne est formée des composantes du vecteurs x_j dans e, on a donc :

$$\boxed{\mathrm{dét}_e(x_1, x_2, \dots, x_n) = \mathrm{dét}\, \mathrm{X}}$$

La droite D de $\otimes^n \mathrm{E}$ est stable par le groupe symétrique ainsi que par le groupe linéaire Gl(E). C'est donc un espace de tenseurs. Nous allons déterminer la représentation linéaire de Gl(E) dans D. Chaque élément g de Gl(E) opère dans D par une homothétie :

$$\begin{aligned} g \cdot \overline{e} &= g \cdot \mathrm{A}(e_1 \otimes e_2 \dots \otimes e_n) \\ &= \mathrm{A} \cdot g e_1 \otimes g e_2 \dots \otimes g e_n \\ &= \mathrm{A} \sum g_1^{i_1} g_2^{i_2} \dots g_n^{i_n} e_{i_1} \otimes e_{i_1} \otimes e_{i_2} \dots \otimes e_{i_n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) g_1^{\sigma_1} g_2^{\sigma_2} \dots g_n^{\sigma_n} \mathrm{A}(e_1 \otimes e_2 \otimes \dots e_n), \end{aligned}$$

d'où : $$g\overline{e} = (\mathrm{dét}\ g) \cdot \overline{e}.$$

En particulier, le changement de base : $e' = e \cdot \mathrm{S}$ dans E détermine le changement de base dans D : $\overline{e}' = \mathrm{dét}\ \mathrm{S} \cdot \overline{e}$

L'unique composante λ d'un tenseur v de D devient lors de ce changement de base : $\boxed{\lambda' = (\overset{-1}{\mathrm{dét}}\ S)\lambda}$, en particulier : $\mathrm{dét}_{e'}(x_1, x_2, \dots, x_n) = (\overset{-1}{\mathrm{dét}}\ S) \mathrm{dét}_e(x_1, x_2, \dots, x_n)$.

On a aussi :

$$\boxed{\text{dét}_e(gx_1, gx_2, \ldots, gx_n) = (\text{dét}\, g)\text{dét}_e(x_1, x_2, \ldots, x_n)}.$$

Une opération de g multiplie les volumes par dét g.

Des considérations analogues valent dans le dual $\otimes^n \mathrm{E}^*$ où une base e de E détermine une base $\overline{e}_* = \mathrm{A}(e^{*1}e^{*2}\ldots e^{*n})$ de la droite D_* des tenseurs covariants antisymétriques d'ordre n. Mais puisque le changement de base $e' = e\mathrm{S}$ donne dans E^* le changement de base ${}^t e'^* = {}^t e^* \check{\mathrm{S}}$ (ou $e'^* = \overset{-1}{\mathrm{S}}\, e^*$), le changement de base dans D_* devient : $\overline{e}'_* = (\overset{-1}{\text{dét}}\ \mathrm{S})\overline{e}_*$ et Gl(E) opère par les homothéties inverses : $g \cdot \overline{e}_* = (\overset{-1}{\text{dét}}\ g) \cdot \overline{e}_*$ sur D_*.

Un tenseur w de D_* a une seule composante relativement à la base $\overline{e}_*$ associée à la base e : $w = \mu\overline{e}_*$ et lors d'un changement de base $e' = e\mathrm{S}$,

$$\boxed{\mu' = (\text{dét}\ S)\mu}.$$

Définition III.6 :

Soient E un espace vectoriel de dimension finie n, D la droite $\mathrm{A} \cdot (\otimes^n \mathrm{E})$ des tenseurs antisymétriques de $\otimes^n \mathrm{E}$, D_* la droite $\mathrm{A} \cdot (\otimes^n \mathrm{E}^*)$ des tenseurs antisymétriques de $\otimes^n \mathrm{E}^*$, munies des représentations naturelles du groupe linéaire Gl(E) : si $g \in \mathrm{G}$ et $x \in \mathrm{D}$, $\mathrm{D}(g)x = (\text{dét}\ g)x$, si $\varphi \in \mathrm{D}_*$, $\mathrm{D}_*(g)\varphi = (\overset{-1}{\text{dét}}\ g)\varphi$.

D est la droite canonique des *pseudoscalaires contravariants de* E, D_* celle des *pseudoscalaires covariants*. Si r est un entiers positif les puissances tensorielles $\mathrm{D}^r = \otimes^r \mathrm{D}$ et $\mathrm{D}^r_* = \otimes^r \mathrm{D}_*$, munies du produit tensoriel des représentations de Gl(E) dans chaque facteur, sont des droites (§II.9).

Si $g \in \mathrm{Gl(E)}$ et $u \in \mathrm{D}^r$, $\mathrm{D}^r(g) \cdot u = (\text{dét}\ g)^r \cdot u$, si $v \in \mathrm{D}^r_*$, $\mathrm{D}^r_*(g) \cdot v = (\text{dét}\, g)^{-r} v$. D^r et D^r_* sont les droites canoniques des *pseudo-scalaires de* E *contravariants d'ordre* r (ou de variance scalaire $(-r)$, resp. *covariants d'ordre* r (ou de variance scalaire $+r$).

Si e est une base de E, $\overline{e} = \mathrm{A} \cdot (e_1 \otimes e_2 \otimes \ldots e_n)$, $\overline{e}_* = \mathrm{A}(e^{*1} \otimes e^{*2} \otimes \ldots e^{*n})$ les bases associées de D et D_*, $\overline{e}^r = \overline{e} \otimes \overline{e} \ldots \overline{e}$,

$\overline{e}_{r,*} = \overline{e}_* \otimes \overline{e}_* \otimes \ldots \overline{e}_*$ les bases associées de D^r et D^r_*, on voit qu'un pseudoscalaire p de D^r (resp. p_* de D^r_*) est la donnée pour chaque base e de E d'un nombre, son unique coordonnée $p(e)$ par rapport à $\overline{e}^r$ (resp. $p_*(e)$ par rapport à $\overline{e}^r_*$), qui se transformera lors d'un changement de base $e' = e\mathrm{S}$ par : $p(e') = (\text{dét } S)^{-r}p(e)$ (contravariance scalaire) resp. $p_*(e') = (\text{dét S})^r p_*(e)$ (covariance scalaire).

Remarque : Il faut bien noter que dans la structure d'une droite de pseudoscalaires intervient non seulement la représentation du groupe linéaire Gl(E) par une puissance entière du déterminant mais également la correspondance : (base de e) $\longrightarrow$ (base $\overline{e}^r$ ou $\overline{e}^r_*$ de D^r ou D^r_*) ces dernières jouant le rôle d'« unités » des pseudoscalaires correspondants, et permettant de leur attribuer une composante dans chaque base.

2) Le produit tensoriel de la droite D par la droite D_* est une droite $\mathrm{D} \otimes \mathrm{D}_*$ sur laquelle le produit tensoriel des représentations de Gl(E) est la représentation identique. Du point de vue tensoriel, $\mathrm{D} \otimes \mathrm{D}_*$ s'identifie donc aux scalaires. Quelle que soit la base e de E, $\overline{e} \otimes \overline{e}_*$ est le même élément de $\mathrm{D} \otimes \mathrm{D}_*$, unité des scalaires. De même $\mathrm{D}^r \otimes \mathrm{D}^r_*$ est canoniquement isomorphe, en tant que Gl(E)-module, au corps des scalaires muni de la représentation identique.

Exemples :

1) $u(u_1, u_2, \ldots, u_n)$ étant une base fixée servant de définition de l'unité de volume dans E, associons à chaque base e le nombre $v_u(e)$, volume de e relativement à u. Si $e' = e\mathrm{S}$, on a $v_u(e') = (\text{dét S}) \cdot v_u(e)$. La fonction : $e \to v_u(e)$ est un pseudoscalaire de variance 1 tandis que la fonction : $e \to v_u(e^*)$ est un pseudoscalaire de variance -1.

2) Si $x_1, x_2, \ldots, x_n$ sont des vecteurs fixes, la donnée pour chaque base e du nombre $\text{dét}_e(x_1, x_2, \ldots, x_n)$ est un pseudoscalaire, de variance (-1).

3) Les pseudoscalaires de variance $+2$ et -2 apparaissent habituellement sous la forme suivante :

Si f est une forme bilinéaire sur E, F et F$'$ les matrices des composantes de f dans les bases e et $e' = e \cdot \mathrm{S}$, on passe de F à F$'$ par : $\mathrm{F}' = {}^t\mathrm{SFS}$ et si à chaque base e on associe le nombre

$\varphi_e = \text{dét}\ F$, on a donc

$$\boxed{\varphi_{e'} = (\text{dét S})^2 \varphi_e}.$$

Si au lieu de la matrice F des composantes de f on prend la matrice H des composantes dans e d'un tenseur deux fois contravariant les nombres $\psi_e = \text{dét H}$ satisfont à :

$$\boxed{\psi_{e'} = (\text{dét S})^{-2} \psi_e}.$$

Remarque : Si l'on *restreint* le choix des bases de E à un ensemble de bases qui se déduisent l'une de l'autre par des matrices de changement de base de déterminant $+1$, ce qui revient à *restreindre* le groupe opérant sur E au sous-groupe spécial linéaire Sl(E) que forment les opérateurs linéaires de déterminant $+1$, le volume devient invariant, les *pseudoscalaires deviennent invariants et se comportent comme des scalaires.*

III.7 – Les pseudotenseurs. Exemple : les pseudotenseurs ε de Levi-Civita

On rencontre fréquemment dans les applications des objets caractérisés, comme les tenseurs, par des ensembles de composantes garnies d'indices, associées à chaque base de l'espace vectoriel de référence E. Cependant les formules de changements de base révèlent que ces objets ne sont des tenseurs « qu'à un facteur scalaire près ».

Ils appartiennent à des Gl(E)-modules un peu plus généraux que les représentations tensorielles du groupe linéaire de E : les *représentations pseudotensorielles de* Gl(E).

Définition III.7 :

L'espace pseudotensoriel de variance (v) et de variance scalaire r, sur l'espace vectoriel E est le produit vectoriel E est le produit tensoriel de la droite D^r des pseudoscalaires de variance r, entier $\rangle$ ou $\langle 0$ par l'espace tensoriel $\otimes^{(v)}\mathrm{E}$. En tant qu'espace vectoriel, il s'identifie donc à $\otimes^{(v)}\mathrm{E}$ (cf. I.8) et un pseudotenseur possède dans chaque base e de E un ensemble de composantes analogues à celles d'un élément de $\otimes^{(v)}\mathrm{E}$.

Le groupe Gl(E) opère sur $(\otimes^{(v)}\mathrm{E})\otimes\mathrm{D}^r$ par le produit tensoriel de ses représentations ce qui fournit les formules de transformations des composantes par changement de base.

Si $s \in \mathrm{Gl(E)}$ a pour matrice S dans la base e, les composantes du pseudotenseur t dans la base $e' = e\mathrm{S} = se$ sont celles de $\overset{-1}{s}\, t$ dans e soient :

$$\boxed{\sum t'^{i_1\ldots i_p}_{j_1\ldots j_q} e'_{i_1} \otimes \ldots \otimes e'^{*j_1} \ldots = (\text{dét S})^r \sum t^{k_1\ldots k_p}_{l_1\ldots l_q} e_{k_1} \otimes \ldots \otimes e^{*l_1}},$$

$$\boxed{t'^{i_1\ldots i_p}_{j_1\ldots j_q} = (\text{dét S})^r \sum t^{k\ldots k_p}_{l_1\ldots l_q} \quad \check{s}^{i_1}_{k_1} \ldots \check{s}^{i_p}_{j_p} s^{l_1}_{j_1} \ldots s^{j_q}_{j_q}},$$

et ces formules caractérisent les pseudotenseurs p fois contravariants, q fois contravariants et de variance scalaire r.

Exemple : les pseudotenseurs de Levi-Civita

Soient E un espace vectoriel de dimension n, D et D* des droites de pseudoscalaires de variance scalaire (-1) et $(+1)$. Le pseudotenseur, n fois contravariant, de variance scalaire $(+1)$ défini dans $(\otimes^n\mathrm{E}) \otimes \mathrm{D}^*$ quelle que soit la base e de E par :

$$\varepsilon = \mathrm{A}(e_1 \otimes e_2 \otimes \ldots e_n) \otimes \overline{e}_*$$

est un pseudotenseur invariant (pour les opérateurs de Gl(E)) appelé *peudotenseur ε contravariant de Levi-Civita.* Ses composantes sont, dans n'importe quelle base :

$$\varepsilon^{i}1^{i}2^{\cdots i}n = 0 \text{ si deux des indices sont égaux}$$

$$= \pm 1, \text{ signature de la permutation}$$

$$(1, 2, \ldots, n) \longrightarrow (i_1 i_2 \ldots i_n) \text{ si les indices sont distincts.}$$

De la même façon, le pseudotenseur n fois covariant, de variance scalaire (-1), défini dans $(\otimes^n\mathrm{E}^*) \otimes \mathrm{D}$ quelle que soit la base e de E par :

$$\varepsilon_* = \mathrm{A}(e^{*1} \otimes e^{*2} \otimes \ldots e^{*n}) \otimes \overline{e}$$

est un pseudotenseurs appelé *pseudotenseur ε_* covariant de Levi-Civita* dont les composantes de l'ε contravariant!

Si l'on s'en tenait à la valeur des composantes rien ne distinguerait les pseudotenseurs ε et ε_*. La différence réside évidemment dans la variance, et il est intéressant d'écrire les formules de changement de base pour ces deux pseudotenseurs. Si $e' = e\mathrm{S}$:

$$\boxed{\begin{aligned} \varepsilon_{j'_1 j'_2 \dots j'_n} &= \overset{-1}{(\text{dét S})} \sum \varepsilon_{j_1 j_2 \dots j_n} s^{j_1}_{j'_1} s^{j_2}_{j'_2} \dots s^{j_n}_{j'_n} \\ \varepsilon^{i'_1 i'_2 \dots i'_n} &= (\text{dét S}) \sum \varepsilon^{i_1 i_2 \dots i_n} \check{s}^{i'_1}_{l_1} \check{s}^{i'_2}_{i_2} \dots \check{s}^{i'_n}_{i_n} \end{aligned}}\,.$$

Les sommes des membres de droites sont les développements de dét S et de dét Š dans lesquels on a permuté l'ordre des colonnes en effectuant sur celles-ci les permutations : $(1, 2, \dots, n) \to (j'_1 j'_2 \dots j'_n)$ et $\to (i'^1 i'^2 \dots i'^n)$ respectivement, d'où le résultat.

Toutes les opérations tensorielles : multiplication, contraction, permutation de facteurs, peuvent être effectuées sur les pseudotenseurs. En particulier le produit de deux pseudotenseurs de variances scalaires opposées est un tenseur.

Exemples :

1) Le produit des deux pseudotenseurs de Levi-Cevita est le *tenseur de Kronecker*, élément de $(\otimes^n \mathrm{E}) \otimes (\otimes^n \mathrm{E}^*) = \otimes^n_n \mathrm{E}$ qui s'écrit :

$$\begin{aligned} \delta &= \mathrm{A}(e_1 \otimes e_2 \dots \otimes e_n) \otimes \mathrm{A}(e^{*1} \otimes e^{*2} \otimes \dots e^{*n} \\ &= \sum \varepsilon^{i_1 i_2 \dots i_n} \varepsilon_{j_1 j_2 \dots j_n} (e_{i_1 \otimes i_2 \otimes \dots i_n}) \otimes (e^{*i_1} \otimes e^{*i_1} \otimes \dots e^{*i_n}), \end{aligned}$$

δ est un invariant, comme produit de pseudotenseurs de variances opposées.

$\otimes^n_n \mathrm{E}$ représente l'algèbre des opérateurs linéaires aussi bien de l'espace vectoriel $\otimes^n \mathrm{E}$ que de son dual $\otimes^n \mathrm{E}^*$. δ est ainsi un opérateur de rang un de $\otimes^n \mathrm{E}$ comme de $\otimes^n \mathrm{E}^*$: c'est, dans chacun de ces espaces, l'opérateur d'antisymétrisation A dont δ est donc la représentation tensorielle.

Les composantes de δ sont *les mêmes dans toutes les bases.* Ce sont les *symboles de Kronecker :*

$\delta^{i_1 \dots i_n}_{j_1 \dots j_n} = 0$ si deux indices en haut ou deux indices en bas sont égaux.

$\delta^{i_1 \dots i_n}_{j_1 \dots j_n} \pm 1$; signature de la permutation $(i_1 i_2 \dots i_n) \to (j_1 j_2 \dots j_n)$ si tous les indices en haut et en bas sont distincts.

2) Le contracté de $(n-1)$ vecteurs $x_1, x_2, \dots, x_{n-1}$ avec les $(n-1)$ derniers indices du pseudotenseur covariant ε_* est un pseudo-covecteur de variance scalaire (-1) : c'est le *produit vectoriel* $x_1 \times x_2 \times \dots x_{n-1}$ que l'on retrouvera au §IV.10., appelé parfois *pseudoproduit vectoriel*.

3) Le contracté sur tous les indices du pseudotenseur covariant ε_* avec n vecteurs $x_1 x_2 \dots x_n$ est le pseudoscalaire contravariant (variance scalaire -1) :

$$e \to \text{dét}_e(x_1 x_2 \dots x_n) = \sum \varepsilon_{i_1 i_2 \dots i_n} x^{i_1}_{,1} x^{i_2}_{,2} \dots x^{i_n}_{,n} = \text{dét X}.$$

Si l'on dispose d'un pseudoscalaire de variance $+1$ ou (-1), on peut, en, élevant à la puissance $(-r)$, extraire par multiplication la partie purement tensorielle d'un pseudotenseur de variance scalaire r. C'est ce que l'on fait systématiquement dans les espaces munis d'une géométrie. Dans le cas des espaces euclidiens, le déterminant G des composantes du produit scalaire est de variance 2. On utilise donc $\sqrt{\mathrm{G}}$, tandis que dans l'espace de Minkowski, ce déterminant G est toujours négatif, et on utilise $\sqrt{-\mathrm{G}}$, en prenant bien garde toutefois lors d'un changement de base $e' = e\mathrm{S}$ que l'on a : $\sqrt{\mathrm{G}'} = |\text{dét S}|\sqrt{\mathrm{G}}$ avec $|\text{dét S}|$ au lieu de dét S. Cela conduit à la nouvelle espèce de pseudotenseurs du § suivant.

III.8 – Espaces liés à un espace vectoriel E. Scalaires orientés d'un espace vectoriel réel. Tenseurs et pseudotenseurs affines orientés sur un espace vectoriel réel

Nous avons vu successivement la définition des espaces tensoriels $(\otimes^{(v)}\mathrm{E})$, des espace pseudotensoriels (§III.6 et III.7) sur E, ces derniers démontrant que le produit tensoriel d'espaces de tenseurs *n'est pas* en général un espace de tenseurs suivant la définition donnée au §III.4. Nous allons donc devoir définir une structure plus générale, englobant les précédentes :

Définition III.8 :

Soit E un espace vectoriel. On appelle *espace lié à* E pour sa seule structure linéaire tout espace vectoriel T muni des structures suivantes :

1) une représentation linéaire ρ de Gl(E) dans T. Gl(E) est le *groupe de structure* de T

2) une application τ qui à chaque base e de E fait correspondre une base τ_e de T, dite base associée à e de telle sorte que $\forall e$, $\forall g \in \mathrm{Gl(E)}$, $\tau_{g \cdot e} = \rho(g)\tau_e$ (il suffit donc que τ_e soit définie pour une base particulière e, $\tau_{g \cdot e}$ s'en déduit alors par 2)).

A chaque élément t de T correspond ainsi pour chaque base de e un *ensemble de composantes*, ses composantes dans τ_e.

Exemple : scalaires orientés d'un espace vectoriel réel E. Les bases de E se divisent en deux classes par la relation d'équivalence : $\{$ si $e' = e\mathrm{S}$, $e' \sim e \Leftrightarrow$ dét $\mathrm{S} > 0\}$ Orienter E, c'est, par définition, choisir l'une de ces classes, dont les bases seront alors dites positives. Une orientation de E est équivalente à une orientation de la droite D des pseudoscalaires contravariants, par les bases $\overline{e}$ associées, puisque si $e' = e\mathrm{S}$, $\overline{e}' =$ dét $\mathrm{S} \cdot \overline{e}$ (définition III.6). Elle est aussi équivalente à l'orientation de D_* par les bases $\overline{e}_*$.

Soient R^0 une droite réelle, ω un élément non nul de R^0, τ l'application qui à toutes les bases d'une classe associe ω, à toutes les bases de l'autre classe, associe $(-\omega)$. Si ρ est la représentation linéaire de Gl(E) dans R^0 qui associe à $g \in \mathrm{Gl(E)}$ l'homothétie $\rho(g) =$ Signe (dét g), on a bien $\tau_{g \cdot e} = \rho(g)\tau_e$.

Munie de ces structures R^0 est un espace lié à E appelé droite des scalaires orientés de E. Par un automorphisme positif de E (dét $g > 0$) les scalaires orientés sont invariants, tandis qu'un automorphisme négatif les fait changer de signe.

Si T est un espace de tenseurs ou de pseudotenseurs affines sur E le produit tensoriel $\mathrm{T}^0 = \mathrm{T} \otimes \mathrm{R}^0$, muni du produit tensoriel des représentations de Gl(E), est un espace lié à E appelé espace des tenseurs ou pseudotenseurs orientés associés à T.

En tant qu'espaces vectoriels, T^0 et T sont isomorphes. Par contre les représentations de Gl(E) dans T et T^0, donc les formules de transformation des composantes par changement de base diffèrent. Ce sont les mêmes que celles de T lorsque l'automorphisme (ou le changement de base) est positif. Il s'y ajoute un changement de signe lorsque l'automorphisme est négatif.

La géométrie et la physique offrent de nombreux exemples de tenseurs et pseudotenseurs orientés (cf. chapitre V).

III.9 – Tenseurs associés aux espaces vectoriels complexes

On a vu à la fin du §II.8 la définition du complexifié $V_{\mathbb{C}} = V \otimes \mathbb{C}$ d'un espace vectoriel réel V.

Les questions concernant les espaces vectoriels complexes font fréquemment intervenir des produits tensoriels sur $\mathbb{R}$ ou sur $\mathbb{C}$. Lorsqu'une ambigüité est possible on indique le corps de base en indice du signe $\otimes$, soit $\otimes_{\mathbb{R}}$ ou $\otimes_{\mathbb{C}}$. Si V et W sont deux espaces vectoriels réels on vérifie par exemple à l'aide de bases, l'isomorphisme d'espaces vectoriels complexes :

$$(V \otimes_{\mathbb{R}} \mathbb{C}) \otimes_{\mathbb{C}} (W \otimes_{\mathbb{R}} \mathbb{C}) \equiv (V \otimes_{\mathbb{R}} W) \otimes_{\mathbb{R}} \mathbb{C}.$$

Le complexifié d'une puissance tensorielle p-ème de V s'identifie ainsi à la puissance tensorielle p-ème (sur $\mathbb{C}$) du complexifié de V : $(\otimes^p V) \otimes \mathbb{C} \equiv \otimes^p_{\mathbb{C}}(V \otimes \mathbb{C})$.

Si E est un espace vectoriel complexe, on note ${}_{\mathbb{R}}E$ l'espace vectoriel réel sous-jacent obtenu par restriction à $\mathbb{R}$ des scalaires. $\dim_{\mathbb{R}} E = 2\dim_{\mathbb{C}}(E)$.

Il faut insister sur le fait que l'expression de l'espace vectoriel complexe E comme complexifié d'un sous-espace vectoriel V de ${}_{\mathbb{R}}E$ doit être considérée comme une structure additionnelle, dite « structure réelle » sur E. Elle est définie par la donnée de V appelé « forme réelle de E » dont les éléments sont dits « éléments réels ».

V doit être tel que ${}_{\mathbb{R}}E = V \oplus iV$ ce qui est équivalent aux *deux* conditions : $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} E$ et $V \cap iV = 0$.

Il faut et il suffit pour cela qu'une base (réelle) de V soit une base (complexe) de E. Une application linéaire complexe de $E = V_{\mathbb{C}}$ est donc complètement déterminée par sa restriction à V. En particulier le dual de E s'identifie à l'espace vectoriel des applications linéaires réelles de V dans $\mathbb{C}$: $(V_{\mathbb{C}})^* = \mathfrak{L}_{\mathbb{R}}(V; \mathbb{C})$.

Soit u un vecteur de $E = V \oplus iV$. Il s'écrit $u = x + iy$ avec $x, y \in V$. Le vecteur $Cu = \overline{u} = x - iy$ est le *conjugué de* u, relativement à la structure réelle de E.

L'application C est la *conjugaison*. C'est une involution antilinéaire de E : $C^2 =$ identité, $C(\lambda u) = \overline{\lambda} C(u)$.

Réciproquement, une involution antilinéaire C de E y détermine une forme réelle V sous-espace propre de C pour la valeur propre (+1).

Un espace vectoriel complexe de E n'intervient pas seulement dans les applications par les constructions linéaires ou multilinéaires sur $\mathbb{C}$ qu'on peut lui associer. Par exemple, les produits scalaires hermitiens ou pseudohermitiens, extensions des produits scalaires euclidiens ou pseudoeuclidiens s'expriment à l'aide de formes ou d'applications antilinéaires. Ces dernières doivent donc faire partie des constructions tensorielles sur E. La multiplication par i définissant la structure complexe de E est, pour cela, représentée par un opérateur linéaire réel J de ${}_{\mathbb{R}}\mathrm{E}$ de carré égal à moins l'identité de E, et $\mathrm{E} \equiv ({}_{\mathbb{R}}\mathrm{E}; \mathrm{J})$.

Avec E, on considère l'espace vectoriel complexe *conjugué* $\overline{\mathrm{E}}$ défini comme le couple $\overline{\mathrm{E}} = ({}_{\mathbb{R}}\mathrm{E}; -\mathrm{J})$. L'application identique de ${}_{\mathbb{R}}\mathrm{E}$ est ainsi une application antilinéaire de E sur $\overline{\mathrm{E}}$ ou réciproquement. Une application linéaire complexe $a \in \mathfrak{L}(\mathrm{E})$ est une application linéaire réelle $a \in \mathfrak{L}({}_{\mathbb{R}}\mathrm{E})$ qui commute avec J. Mais si $a\mathrm{J} = \mathrm{J}a$, on a aussi $a(-\mathrm{J}) = (-\mathrm{J})a$ et $\mathfrak{L}(\mathrm{E}) = \mathfrak{L}(\overline{\mathrm{E}})$.

En particulier les groupes linéaires $\mathrm{Gl}(\mathrm{E})$ et $\mathrm{Gl}(\overline{\mathrm{E}})$ forment le même sous-groupe de $\mathrm{Gl}({}_{\mathbb{R}}\mathrm{E})$: $\mathrm{Gl}(\mathrm{E}) = \mathrm{Gl}(\overline{\mathrm{E}})$.

Une base e de E est aussi une base de $\overline{\mathrm{E}}$. Si $a \in \mathfrak{L}(\mathrm{E}) = \mathfrak{L}(\overline{\mathrm{E}})$, $ae_j = \sum e_k(\alpha_j^k + \mathrm{J}\beta_j^k) = \sum e_k(\alpha_j^k - (-\mathrm{J})\beta_{k,j})$. Il résulte que si A est la matrice de a relative à e dans l'espace vectoriel complexe E, la matrice de a relative à la même base e dans l'espace $\overline{\mathrm{E}}$ est $\overline{\mathrm{A}}$. Le groupe $\mathrm{Gl}(\mathrm{E}) = \mathrm{Gl}(\overline{\mathrm{E}})$ et n'importe lequel de ses sous-groupes G se trouve ainsi représenté comme groupe d'opérateurs linéaires complexes de *deux* espaces complexes distincts : E et $\overline{\mathrm{E}}$. Chacune de ces représentations est appelée la *conjuguée* de l'autre. Concrètement, par le choix d'une base $e = (e_1, e_2, \ldots, e_n)$, G se trouve représenté dans E comme un groupe de matrices complexes $\mathrm{G} \subset \mathrm{Gl}(n; \mathbb{C})$. A chaque $g \in \mathrm{G}$ correspond une matrice $\mathrm{A}(g)$ dans E et la matrice $\overline{\mathrm{A}(g)}$ dans $\overline{\mathrm{E}}$. Pour un sous-groupe G donné de $\mathrm{Gl}(\mathrm{E})$, on peut se demander si ces représentations sont équivalentes, c'est-à-dire s'il existe un isomorphisme $\mathbb{C}$-linéaire s de E sur $\overline{\mathrm{E}}$ tel que $\forall g \in \mathrm{G} : \overline{g}s = sg$, ce qui revient à l'existence d'une matrice inversible S telle que $\overline{\mathrm{A}(g)} = \mathrm{S} \cdot \mathrm{A}(g) \overset{-1}{\mathrm{S}}$, $\forall g \in \mathrm{G}$. Une condition nécessaire est que $\mathrm{Trace}\, \overline{\mathrm{A}}(g) = \mathrm{Trace}\, \mathrm{A}(g)$, $\forall g \in \mathrm{G}$, autrement dit que le caractère de la représentation de G dans E :

$g \to \chi(g) = \text{Trace A}(g)$ soit à valeurs réelles. Cela montre qu'en général, deux représentations complexes conjuguées d'un groupe G ne sont pas équivalentes.

On voit ainsi que, alors que dans d'un corps quelconque K, un espace vectoriel E entraînait avec lui son espace dual E^* et tous les espaces de tenseurs bâtis sur E et E^*, un espace vectoriel complexe de E va entraîner avec lui son conjugué $\overline{E}$, son dual E^*, et le dual de $\overline{E}$: $\overline{E}^*$ ou antidual de E, et tous les espaces de tenseurs bâtis sur eux.

Les valeurs propres de l'opérateur J de ${}_{\mathbb{R}}E$ définissant la structure complexes sont i et $(-i)$.

J n'a donc aucun vecteur propre (réel) dans ${}_{\mathbb{R}}E$ alors que tous les vecteurs (complexes) de E sont des vecteurs propres de J pour la valeur propre i. Cela s'explique par le fait que si u est un vecteur non nul de E, la droite complexe définie par u est dans ${}_{\mathbb{R}}E$ un plan réel U, sous-tendu par u et iu, qui sont linéairement indépendants sur $\mathbb{R}$. La droite complexe sous-tendue par u et le plan réel U sont stables par J, la restriction de J à U étant une rotation de $+\frac{\pi}{2}$. La nécessité d'exprimer la structure définie par l'opérateur J impose la complexification de ${}_{\mathbb{R}}E$. Cela donne un espace vectoriel complexe $({}_{\mathbb{R}}E)_{\mathbb{C}} = {}_{\mathbb{R}} E \otimes \mathbb{C}$ de dimension $2n$ (sur $\mathbb{C}$) que l'on note simplement $E_{\mathbb{C}}$ (c'est donc le « complexifié de l'espace vectoriel complexe E »). L'opérateur réel J de ${}_{\mathbb{R}}E$ se prolonge en un opérateur linéaire complexe $J_{\mathbb{C}}$ de $E_{\mathbb{C}}$: si x et $y \in {}_{\mathbb{R}}E$, $J_{\mathbb{C}}(x + iy) = Jx + iJy$. On a $(J_{\mathbb{C}})^2 = -(\text{Identité de } E_{\mathbb{C}})$ et les valeurs propres de $J_{\mathbb{C}}$, qui sont celles de J, sont i et $(-i)$. Le sous-espace propre de $J_{\mathbb{C}}$ pour la valeur propre i est noté $E_{(1,0)}$. Ses éléments sont appelés *vecteurs de type* $(1, 0)$. Le sous-espace propre de $(-i)$ est noté $E_{(0,1)}$. Ses élément sont les *vecteurs de type* $(0, 1)$. $E_{(1,0)}$ et $E_{(0,1)}$ sont des sous-espaces vectoriels complexes de $E_{\mathbb{C}}$, qui en est la somme directe. La conjugaison associée à la structure de complexification les échange : $E_{(0,1)} = \overline{E}_{(1,0)}$ et $E_{(1,0)} = \overline{E}_{(0,1)}$ et ils ont donc même dimension n. $J_{\mathbb{C}}i = iJ_{\mathbb{C}}$ est une involution linéaire complexe de $E_{\mathbb{C}}$ dont $E_{(1,0)}$ et $E_{(0,1)}$ sont les sous-espaces propres pour les valeurs propres (-1) et $(+1)$. Les projecteurs sur ces sous-espaces sont $p_{(1,0)} = \frac{1}{2}\,\{I - iJ_{\mathbb{C}}\}$ soit $E_{10} = \text{Im}(I - iJ_{\mathbb{C}}\}$ et $p_{(0,1)} = \frac{1}{2}\,\{I + iJ_{\mathbb{C}}\}$ soit $E_{(0,1)} = \text{Im}(I + iJ_{\mathbb{C}})$. D'après la propriété universelle de la complexification (fin du §II.8), l'application identique de ${}_{\mathbb{R}}E$ sur E se prolonge en une application linéaire complexe P de $E_{\mathbb{C}}$ sur E par $P(x+iy) = x + Jy$ et une application linéaire complexe Q de $E_{\mathbb{C}}$ sur $\overline{E}$ par $Q(x+iy) = x - Jy$.

Soient Φ et Ψ les applications de E dans $E_{\mathbb{C}}$ définies par :

$$\Phi = p_{(1,0)}u = \frac{1}{2}\,(u - iJu) \text{ d'où } \Phi(Ju) = \frac{1}{2}\,(Ju + iu) = i\Phi(u),$$
$$\Psi = p_{(0,1)}u = \frac{1}{2}\,(u + iJu) \text{ d'où } \Psi(Ju) = \frac{1}{2}\,(Ju - iu) = -i\Psi(u).$$

On a :

1) $P \circ \Phi =$ Identité de E, ce qui prouve *l'isomorphisme des espaces vectoriels complexes* E *et* $E_{(1,0)}$. (Sur ce dernier les structures complexes induites par $E_{\mathbb{C}}$ et par $J_{\mathbb{C}}$ sont les mêmes : $J_{\mathbb{C}} = i$ sur $E_{(1,0)}$. Puisque $P \circ p_{(0,1)} = 0$: le noyau de P est $E_{(0,1)}$.

2) Ψ est antilinéaire de E sur $E_{(0,1)}$ donc linéaire de $\overline{E}$ sur $E_{(0,1)}$. On a : $Q \circ \Psi =$ Identité de E, ce qui prouve *l'isomorphisme des espaces vectoriels complexes* $\overline{E}$ *et* $E_{(0,1)}$ (sur ce dernier les structures complexes définies par $E_{\mathbb{C}}$ et par $J_{\mathbb{C}}$ sont opposées : $J_{\mathbb{C}} = -i$ sur $E_{(0,1)}$). Le noyau de Q est $E_{(1,0)}$.

Exemple d'application de la complexification $E_{\mathbb{C}}$ *de* E *:* Soit $a \in \mathfrak{L}(E)$. C'est un opérateur ${}_{\mathbb{R}}a \in \mathfrak{L}({}_{\mathbb{R}}E)$ qui commute avec J et dont le complexifié $({}_{\mathbb{R}}a)_{\mathbb{C}} = \widetilde{a}$ est donc un opérateur linéaire complexe de $E_{\mathbb{C}}$ qui laisse stables les sous-espaces propres de $J_{\mathbb{C}}$. Une base e de E détermine, par l'isomorphisme Φ et l'antisomorphisme Ψ, des bases sur $E_{(1,0)}$ et $E_{(0,1)}$. Si A est la matrice de a dans e, c'est aussi la matrice de $\widetilde{a}$ dans la base correspondante sur $E_{(1,0)} \equiv E$, tandis que $\overline{A}$ est la matrice de la restriction de $\widetilde{a}$ à $E_{(0,1)} \equiv \overline{E}$.

Il en résulte que $\text{dét}_{\mathbb{R}} a = \text{dét}\, \widetilde{a} = \text{dét}\, A \cdot \text{dét}\, \overline{A} = |\text{dét}\, A|^2$.

Une conséquence de ce résultat est que les changements de bases complexes de E conservent l'orientation de ${}_{\mathbb{R}}E$.

Une forme linéaire $\varphi \in E_{\mathbb{C}}^*$ se décompose en : $\varphi = \varphi \circ p_{(1,0)} + \varphi \circ p_{(0,1)}$. Notons :

$$p^{(1,0)}\varphi = \varphi \circ p_{(1,0)} = \frac{1}{2}\,(\varphi - i\varphi J_{\mathbb{C}})$$
$$\text{et } p^{(0,1)}\varphi = \varphi \circ p_{(0,1)} = \frac{1}{2}\,(\varphi + i\varphi J_{\mathbb{C}})$$

$p^{(1,0)}$ et $p^{(0,1)}$ sont des projecteurs supplémentaires de $E^*_{\mathbb{C}}$.

Soit $E^{(1,0)} = \operatorname{Im} p^{(1,0)}$. Ses éléments sont appelés *formes de type* $(1,0)$. Une forme $\varphi \in E^*_{\mathbb{C}}$ appartient à $E^{(1,0)}$ si et seulement si elle est nulle sur $E_{(0,1)}$, ou si et seulement si sa restriction à $E \subset E_{\mathbb{C}}$ est une forme linéaire complexe sur $E = ({}_{\mathbb{R}}E; J) : E^{(1,0)} \equiv E^*_{(1,0)} \equiv E^*$.

De même, on note $E^{(0,1)} = \operatorname{Im} p^{(0,1)}$. Ses éléments sont appelés *formes de type* $(0,1)$. $\varphi \in E^*_{\mathbb{C}}$ appartient à $E^{(0,1)}$ si et seulement si elle est nulle sur $E_{(1,0)}$, ou si et seulement si sa restriction à $E \subset E_{\mathbb{C}}$ est antilinéaire sur $E = ({}_{\mathbb{R}}E; J)$.

On a : $E^{(0,1)} \equiv E^*_{(0,1)} \equiv \overline{E}^*$, antidual de E de $\overline{E}$.

Si $e = \{e_1, e_2, \dots, e_n\}$ est une base (complexe) de E, on peut lui associer la base ${}_{\mathbb{R}}e = \{e_1, e_2, \dots, e_n, Je_1, Je_2, \dots, Je_n\}$ de ${}_{\mathbb{R}}E$ et les bases $\varepsilon = \{\varepsilon_j = \Phi(e_j) = \frac{1}{2}\,(e_j - iJe_j)\}$ de $E_{(1,0)}$ et $\overline{\varepsilon} = \{\overline{\varepsilon}_j = \overline{\Psi}(e_j) = \frac{1}{2}\,(e_j + iJe_j)\}$ de $E_{(0,1)}$. La réunion de ε et $\overline{\varepsilon}$ est une base complexe de $E_{\mathbb{C}}$. Dans cette base, les vecteurs de ${}_{\mathbb{R}}e$ s'écrivent : $e_j = \varepsilon_j + \overline{\varepsilon}_j$, $Je_j = i(\varepsilon_j - \overline{\varepsilon}_j)$. Un vecteur u de E s'écrit dans la base ${}_{\mathbb{R}}e$: $u = \sum e_j x^j + \sum Je_k y^k$ où x^j et y^k sont les formes coordonnées (réelles) associées à ${}_{\mathbb{R}}e$.

Dans la base $(\varepsilon, \overline{\varepsilon})$, u s'écrit : $u = \sum(\varepsilon_j + \overline{\varepsilon}_j)x^j + i\sum(\varepsilon_k - \overline{\varepsilon}_j k)y^k$ soit $u = \sum \varepsilon_j(x^j + iy^j) + \sum \overline{\varepsilon}_k(x^k - iy^k) = \sum \varepsilon_j z^j + \sum \overline{\varepsilon}_k \overline{z}_k$.

Chaque fonction z^j est une application $\mathbb{R}$-linéaire de E dans $\mathbb{C}$ déterminée par $z^j(e_k) = \delta^j_k$ et $z^j(Je_k) = i\delta^j_k$. Elle se prolonge en une forme $\mathbb{C}$-linéaire sur $E_{\mathbb{C}}$ par $z^j_{\mathbb{C}}(u + iv) = z^j(u) + iz^j(v)$. On a $z^j_{\mathbb{C}}(\varepsilon_k) = \delta^j_k$ et $z_{j,\mathbb{C}}(\overline{\varepsilon}_k) = 0$. Les fonctions $z^j_{\mathbb{C}}$ sont donc les formes coordonnées ε^{*j} base de $E^{(1,0)}$. Comme au §II.2., la géométrie différentielle préfère, par souci de clarté, utiliser pour désigner ces formes coordonnées, la notation fonctionnelle dz^j.

On a donc :

$$dz^j \equiv \varepsilon^{*j}.$$

De la même façon, la fonction $\overline{z}^j$ est une application $\mathbb{R}$-linéaire de E dans $\mathbb{C}$ déterminée par $\overline{z}^j(e_k) = \delta^i_k$ et $\overline{z}^j(Je_k) = -i\delta^j_k$. Elle est antilinéaire sur (E; J). Elle se prolonge en une forme $\mathbb{C}$-linéaire sur $E_{\mathbb{C}}$ par $\overline{z}^j_{\mathbb{C}}(u + iv) = \overline{z}^j(u) + i\overline{z}^j(v)$. Observons qu'alors que $\overline{z}^j$

est complexe conjugué de z^j sur (E; J), $\overline{z}^j_{\mathbb{C}}$ est indépendant de $z^j_{\mathbb{C}}$ sur $\mathrm{E}_{\mathbb{C}}$.

En effet : $\overline{z}^j_{\mathbb{C}}(\varepsilon_k) = 0$ et $\overline{z}^j_{\mathbb{C}}(\overline{\varepsilon}^j) = \dfrac{1}{2}\ \overline{z}^j(e_j + \mathrm{J}e_j) = 1.$ Les fonctions $\overline{z}^j_{\mathbb{C}}$ sur $\mathrm{E}_{\mathbb{C}}$ sont les formes coordonnées $\overline{\varepsilon}^{*j}$ base de $\mathrm{E}^{(0,1)}$. La géométrie différentielle préfère les noter $d\overline{z}^j$. On a donc : $d\overline{z}^j = \overline{\varepsilon}^{*j}$.

Définition III.9 :

Soit E un espace vectoriel complexe de dimension finie. On appelle *espace de tenseurs associés à* E tout produit tensoriel sur $\mathbb{C}$ d'espace vectoriels complexes $\mathrm{E}_{(1,0)}$, $\mathrm{E}_{(0,1)}$, $\mathrm{E}^{(1,0)}$, $\mathrm{E}^{(0,1)}$, qui s'identifient respectivement à E, $\overline{\mathrm{E}}$, E^*, $\overline{\mathrm{E}}^*$. La structure tensorielle de ces espaces est déterminées par la correspondance qui à chaque base e de E associe les bases ε, $\overline{\varepsilon}$, ε^*, $\overline{\varepsilon}^*$ de $\mathrm{E}_{(1,0)}$, $\mathrm{E}_{(0,1)}$, $\mathrm{E}^{(1,0)}$, $\mathrm{E}^{(0,1)}$ et les produits tensoriels de ces bases, ainsi que par les représentations linéaires associées du groupe linéaire Gl(E) qui en est le groupe de structure.

Il est d'usage de surligner les indices correspondant aux facteurs $\overline{\mathrm{E}} = \mathrm{E}_{(0,1)}$ et $\overline{\mathrm{E}}^* = \mathrm{E}^{(0,1)}$ des produits tensoriels.

Un élément de $\mathrm{E} \otimes \overline{\mathrm{E}} \otimes \overline{\mathrm{E}}^* \otimes \mathrm{E}$ s'écrit par exemple :

$$\sum t^{i\overline{j}l}_{\overline{k}} \varepsilon_i \otimes \overline{\varepsilon}_j \otimes \overline{\varepsilon}^{*k} \otimes \varepsilon_l.$$

Remarque : On peut s'interroger sur la nécessité d'introduire ces espaces ornés d'indices : $\mathrm{E}_{(1,0)}$, $\mathrm{E}_{(0,1)}$, $\mathrm{E}^{(1,0)}$, $\mathrm{E}^{(0,1)}$, alors qu'ils représentent simplement E, $\overline{\mathrm{E}}$, E^*, $\overline{\mathrm{E}}^*$. La raison en est simple : la manipulation de deux structures complexes J et (−J) sur le même espace ${}_{\mathbb{R}}\mathrm{E}$ conduirait inévitablement à de dangereuses acrobaties. La sagesse est d'avoir *deux* espaces distincts et de manifester leur différence par des indices bien visibles.

Exemples : Formes sesquilinéaires et hermitiennes sur un espace vectoriel complexe E.

Une forme sesquilinéaire gauche sur E est une application f de $\mathrm{E}\times\mathrm{E}$ dans $\mathbb{C}$: $(x, y) \rightarrow f(x, y)$, linéaire dans son premier argument x, antilinéaire dans le second y. C'est donc une forme bilinéaire sur

$E \times \overline{E} = E_{(1,0)} \times E_{(0,1)}$ et elle peut être représentée par un élement $\varphi = \sum_j \alpha_j \otimes \overline{\beta}_j = \sum f_{i\overline{j}} dz^i \otimes d\overline{z}^j$ du produit tensoriel $E^* \otimes \overline{E}^*$.

A chaque vecteur y, f associe une forme linéaire $\rho(y)$ sur E par $x \to f(x,y) = \langle x, \rho(y) \rangle$. ρ est une application antilinéaire de E dans E^* (ou linéaire de $\overline{E}$ dans E^*) dite associée à droite de $f \cdot \rho(y)$ est tout simplement la contraction à droite de φ par y. De même, chaque vecteur x détermine une forme $\gamma(x)$ antilinéaire sur E par : $y \to f(x,y) = \langle \gamma(x), y \rangle$. γ est une application antilinéaire de E dans $\overline{E}^*$, transposée de ρ et $\gamma(x)$ est la contraction à gauche de φ par x. On a : rang f = rang $|f_{ij}|$ = rang ρ = dim $\mathrm{Im}\,\rho$ = rang γ = dim $\mathrm{Im}\,\gamma$.

D'une façon analogue, une forme sesquilinéaire droite sur E est une application f de $E \times E$ dans $\mathbb{C}$: $(x,y) \to f(x,y)$ linéaire dans le second argument y, antilinéaire dans le premier x, qui peut être représentée par un élément $\varphi = \sum \overline{\alpha}_j \otimes \beta_j = \sum f_{\overline{i}j} \overline{dz}^i \otimes dz^j$ du produit tensoriel $\overline{E}^* \otimes E^*$.

Une forme sesquilinéaire f est dite *hermitienne* si ${}^t\overline{f} = f$, soit : $f(y,x) = \overline{f(x,y)}$ *antihermitienne* si ${}^t\overline{f} = -f$.

L'espace vectoriel complexe S(E) des formes sesquilinéaires gauches (par exemple) sur E possède une structure naturelle de complexifié avec pour forme réelle le sous-espace $\mathfrak{H}$ des formes hermitiennes, le sous-espace réel : $i\mathfrak{H}$ étant celui des formes antihermitiennes : $\delta(E) = \mathfrak{H} \oplus i\mathfrak{H}$. La conjugaison associée est l'involution antilinéaire : $f \to^t \overline{f}$.

$$f(x,y) = \frac{1}{2}\,\{f(x,y) + \overline{f(y,x)}\} + \frac{1}{2}\,\{f(x,y) - \overline{f(y,x)}\}.$$

On appelle forme *semiquadratique* sur E toute fonction à valeurs complexes r obtenue en égalant les deux arguments d'une forme sesquilinéaire f : $r(x) = f(x,x)$. On a : $r(\lambda x) = |\lambda|^2 r(x)$. Contrairement à ce qui se passe pour les formes bilinéaires, on a une bijection entre formes sesquilinéaires (gauches par exemple) et formes semiquadratiques. En effet, f s'exprime en fonction de r par : $f(x,y) = \dfrac{1}{4}\,\{r(x+y) - r(x-y) + ir(x+iy) - ir(x-iy)\}$. f est hermitienne si et seulement si $r(x) = f(x,x)$ est à valeurs réelles.

III.10 – Vecteurs tangents complexes

Soient E un espace vectoriel complexe de dimension finie, $E_{\mathbb{C}}$ son complexifié décomposé en $E_{\mathbb{C}} = E_{(1,0)} \oplus E_{(0,1)}$. Soient $e = \{e_j; j = 1, 2, \ldots\}$ une base de E, $\varepsilon = \{\varepsilon_j\}$ et $\overline{\varepsilon} = \{\overline{\varepsilon}_j\}$ les bases images de e dans $E_{(1,0)}$ et $E_{(0,1)}$. Un vecteur u de E s'écrit dans ${}_{\mathbb{R}}E$ et dans $E_{\mathbb{C}}$: $u = \sum x^j e_j + \sum y^k \mathrm{J}e_k = \sum z^j \varepsilon_j + \sum \overline{z}^k \overline{\varepsilon}_k$. Suivant les définitions du §II.2. les $2n$ vecteurs tangents $\left\{\dfrac{\partial}{\partial x^j}, \dfrac{\partial}{\partial y^k}\right\}$ forment une base de l'espace tangent T_u en un point quelconque u de ${}_{\mathbb{R}}E$, tandis que les formes coordonnées $\{dx^j, dy^k\}$ forment la base duale de l'espace dual T_u^* ou cotangent. Les complexifiés $\mathrm{T}_u \otimes \mathbb{C}$ et $\mathrm{T}_u^* \otimes \mathbb{C}$ sont appelés espaces tangents et cotangents complexes en u. Les éléments de $(\mathrm{T}_u)_{\mathbb{C}}$ sont les vecteurs tangents complexes, ceux de $(\mathrm{T}_u^*)_{\mathbb{C}}$ les différentielles complexes.

Soit maintenant f une application d'un voisinage de $u \in {}_{\mathbb{R}}E$ dans l'espace vectoriel complexe F, différentiable en u. Sa différentielle df_u est une application $\mathbb{R}$-linéaire de ${}_{\mathbb{R}}E$ dans F que l'on peut plus précisément considérer comme une application $\mathbb{R}$-linéaire de T_u dans F, et $df_u(e_j) = \left(\dfrac{\partial f}{\partial x^j}\right)_u$, $df_u(\mathrm{J}e_j) = \left(\dfrac{\partial f}{\partial y^j}\right)_u$.

L'extension complexe de df_u au complexifié $(\mathrm{T}_u)_{\mathbb{C}}$ est une application $\mathbb{C}$-linéaire $\widetilde{df_u}$ de $(T_u)_{\mathbb{C}}$ dans F, ou différentielle complexe de f en u. Si $a, b \in \mathrm{T}_u$, $\widetilde{df_u}(a + ib) = df_u(a) + idf_u(b)$. On a :

$$\begin{aligned} df_u(\varepsilon_j) &= \widetilde{df_u}\Big\{\frac{1}{2}\,(e_j - i\mathrm{J}e_j)\Big\} = \frac{1}{2}\,\{df_u(e_j) - idf_u(\mathrm{J}e_j)\} \\ &= \frac{1}{2}\left\{\Big(\frac{\partial f}{\partial x^j}\Big)_u - i\Big(\frac{\partial f}{\partial y^j}\Big)_u\right\} = \frac{1}{2}\,\Big(\frac{\partial}{\partial x^j} - i\,\frac{\partial}{\partial y^j}\Big)_u \cdot f \end{aligned}$$

et de même $\widetilde{df_u}(\overline{\varepsilon_j}) = \dfrac{1}{2}\,\Big(\dfrac{\partial}{\partial x^j} + i\,\dfrac{\partial}{\partial y^j}\Big)_u \cdot f$ qui définissent les vecteurs (tangents complexes : $\dfrac{\partial}{\partial z^j} = \dfrac{1}{2}\,\Big(\dfrac{\partial}{\partial x^j} - i\,\dfrac{\partial}{\partial y^j}\Big)$ et $\dfrac{\partial}{\partial \overline{z}^j} = \dfrac{1}{2}\Big(\dfrac{\partial}{\partial x^j} + i\,\dfrac{\partial}{\partial y^j}\Big)$ formant les bases associées de $\mathrm{T}_{(1,0)}$ et de $\mathrm{T}_{(0,1)}$. Les bases duales de $\mathrm{T}^{(1,0)}$ et $\mathrm{T}^{(0,1)}$ sont constituées par les différentielles dz^j et $d\overline{z}^k$.

Proposition III.10 :

La condition nécessaire et suffisante pour qu'une application $\mathbb{R}$-linéaire f de ${}_{\mathbb{R}}\mathrm{E}$ dans F soit $\mathbb{C}$-linéaire est que son extension $\widetilde{f}$ à $\mathrm{E}_{\mathbb{C}}$ soit nulle sur $\mathrm{E}_{(0,1)}$.

Preuve : $\widetilde{f}(u + i\mathrm{J}u) = f(u) + if(\mathrm{J}u) = 0$ si et seulement si $f(\mathrm{J}u) = if(u)$, cqfd.

Corollaire III.10 : Soit f une application continûment différentiable d'un ouvert de E dans F. La condition nécessaire et suffisante pour qu'en tout point la différentiable de f soit $\mathbb{C}$-linéaire est que f vérifie le système d'équations : $\dfrac{\partial f}{\partial \bar{z}^j} = \dfrac{1}{2}\left(\dfrac{\partial f}{\partial x^j} + i\,\dfrac{\partial f}{\partial y^j}\right) = 0$ pour $j = 1, 2, \ldots, n$.

Remarque : f est alors analytique complexe. Dans le cas le plus simple : $\mathrm{E} = \mathrm{F} = \mathbb{C}$, l'unique équation $\dfrac{\partial f}{\partial \bar{z}} = 0$ est la condition de Cauchy qui s'écrit à l'aide des parties réelle et imaginaire de $f = \mathrm{U} + i\mathrm{V}$: $\dfrac{\partial \mathrm{U}}{\partial x} = \dfrac{\partial \mathrm{V}}{\partial y}$ et $\dfrac{\partial \mathrm{V}}{\partial x} = -\dfrac{\partial \mathrm{U}}{\partial y}$.

III.11 – Algèbre. Graduations. Produits tensoriels d'algèbres

Une algèbre sur un corps K est un espace vectoriel sur K de dimension finie A muni d'une *application linéaire* $\mu : \mathrm{A} \times \mathrm{A} \to \mathrm{A}$. Cette application μ prend, suivant les cas, le nom de multiplication, produit, composition, crochet, etc... On note génériquement l'image par μ d'un couple $(a, b) \in \mathrm{A} \times \mathrm{A}$ par une simple juxtaposition des éléments : $\mu(a, b) = ab$. Mais on utilise des notations spécifiques dans de nombreux cas particuliers : $\otimes$, $\circ$, $\wedge$, $\vee$, $[,]$ etc... μ peut aussi se définir par une application $\mathrm{L} : a \to \mathrm{L}_a$ de A dans $\mathfrak{L}(\mathrm{A})$, soit $\mathfrak{L}_a b = ab$: L_a est la translation à gauche par a, ou par une application $\mathbb{R} : \mathrm{A} \to \mathfrak{L}(\mathrm{A})$, avec cette fois $\mathrm{R}_b a = ab$: R_b est la translation à droite par b. Une unité est un élément 1 tel que $\mathrm{L}_1 = \mathrm{R}_1 =$ Identité de A. Si A possède une unité, elle est unique, puisque si 1 et $1'$ sont deux unités, $1 \cdot 1' = 1 = 1'$. A est alors dite unitaire ou avec unité. Une algèbre A est dite *associative* si les translations à gauche commutent avec les translations à droite : $\forall a, b, x \in \mathrm{A} : \mathrm{L}_a \cdot \mathrm{R}_b \cdot x = a(xb) = \mathrm{R}_b \cdot \mathrm{L}_a x = (ax)b$.

A est dite *commutative* si $\forall a, b \in \mathrm{A} : ab = ba$.

Une *dérivation* d'une algèbre A est un endomorphisme linéaire D de A qui vis-à-vis de la multiplication possède la propriété : $\mathrm{D}(ab) = \mathrm{D}a \cdot b + a \cdot \mathrm{D}b$. Cette notion sera largement généralisée au Chapitre IV.

Une sous-algèbre, un idéal à gauche, un idéal à droite, un idéal de A, sont des sous-espaces vectoriels stables respectivement par μ, toutes les L_a, toutes les L_a et R_b.

Si $e = \{e_i ; i \in \mathrm{I}\}$ est une base de A, on a, pour tout couple ordonné $(i, j) \in \mathrm{I} \times \mathrm{I}$ une égalité : $e_i \cdot e_j = \sum\limits_k \mathrm{C}_{ij}^k e_k$. L'ensemble de ces égalités, ou table de multiplication de A dans e, détermine μ. Le cas le plus simple d'une table de multiplication est fourni par un ensemble quelconque S muni d'une multiplication interne $\mathrm{S} \times \mathrm{S} \to \mathrm{S}$, soit $(s, t) \to st \in \mathrm{S}$.

L'espace vectoriel sur K de base S : $\mathrm{K}^{(\mathrm{S})}$, dont les éléments sont les combinaisons linéaires finies d'éléments de S à coefficients dans K, est alors une algèbre grâce à la table de multiplication qui applique la base S en elle-même.

Exemples :

1) S est un groupe fini; $\mathrm{K}^{(\mathrm{S})}$ est alors l'algèbre du groupe S à coefficients dans K.

2) Soit X un ensemble quelconque et S l'ensemble des suites finies $(x_1 x_2 \ldots x_m)$ d'éléments de X auxquelles on ajoute la suite vide ( ). On définit dans S un produit : $\mathrm{S} \times \mathrm{S} \to \mathrm{S}$ obtenu par juxtaposition des suites : $(x_1 x_2 \ldots x_m)(x'_1 x'_2 \ldots x'_p) = (x_1, \ldots, x_m x'_1 \ldots x_p)$. Ce produit est associatif et ( ) est un élément neutre. L'algèbre $\mathrm{K}^{(S)}$ est une algèbre associative avec unité appelée *algèbre libre de l'ensemble* X *sur* K. Lorsque X est réduit à un seul élément (x), l'algèbre libre de X sur K est l'algèbre des polynômes d'une variable x à coefficients dans K.

Remarque : La donnée d'une multiplication μ sur A est équivalente à la donnée d'une application linéaire de $\mathrm{A} \otimes \mathrm{A}$ dans A. Lorsque A est de dimension finie, $\mathfrak{L}(\mathrm{A} \otimes \mathrm{A}; \mathrm{A})$ s'identifie à $\mathrm{A} \otimes (\mathrm{A} \otimes \mathrm{A})^* = \mathrm{A} \otimes \mathrm{A}^* \otimes \mathrm{A}^*$. Une structure multiplicative μ sur A peut ainsi être considérée comme un tenseur C une fois contravariant et deux fois covariant sur A. Dans une base e de A, ses composantes sont les « constantes des structures » C^k_{ij} de la table de multiplication de μ dans e tandis que les translations à gauche et à droite sont les contractions sur chacun des indices covariants.

Un homomorphisme f d'une algèbre A dans une algèbre A$'$ sur le même corps est une application linéaire qui respecte les structures multiplicatives : $f(a) = f(a) \cdot f(b)$. Si f est bijective, c'est un isomorphisme car alors $\overset{-1}{f}$ est également linéaire, et l'égalité précédente s'écrit :

$$\overset{-1}{f}(f(a) \cdot f(b)) = ab = \overset{-1}{f}(f(a)) \cdot \overset{-1}{f}(f(b)).$$

Ne sont intéressantes que les algèbres qui satisfont à une propriété supplémentaire. Elles se répartissent en deux classes :

1) les algèbres *alternées* pour lesquelles $aa = 0$, $\forall a \in$ A, dont les principales sont les algèbres de Lie. Elles ne peuvent avoir ni unité, ni élément idempotent non nul.

2) les algèbres *associatives* et les algèbres «*presque associatives*», ces dernières satisfaisant à une forme plus faible d'associativité. Les plus importantes en sont les algèbres de Jordan et les algèbres alternatives.

Les éléments idempotents jouent un rôle essentiel dans l'étude de leur structure.

Si A est une algèbre associative, on peut définir sur l'espace vectoriel A deux autres structures d'algèbres A_- et A_+ à l'aide des formes affaiblies de la multiplication suivantes :

A_- avec le produit antisymétrisé ou crochet $[a, b] = ab - ba$

A_+ avec le produit symétrisé : $a \circ b = \dfrac{1}{2}(ab + ba)$

A_- et A_+ ne sont pas associatives mais l'associativité de A s'y manifeste par des propriétés simples de la multiplication qui servent d'axiomes à deux catégories importantes d'algèbres :

Définition III.11.A :

Une algèbre $\mathfrak{g}$ est une *algèbre de Lie* si sa multiplication, notée par un crochet $[,]$ satisfait aux deux axiomes :

1) $[x, x] = 0$, $\forall x \in \mathfrak{g}$ d'où $[x, y] = -[y, x]$ $\forall x, y \in \mathfrak{g}$ (caractère alterné ou antisymétrique du crochet);

2) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$, $\forall x,y,z\in\mathfrak{g}$ *identité de Jacobi*, qui exprime que les translations à gauche sont des dérivations de l'algèbre, cf. §IV.

Une algèbre J est une *algèbre de Jordan* si sa multiplication satisfait aux deux axiomes :

1) $ab=ba\ \forall a,b\in \mathrm{J}$ (commutativité);

2) $a^2(ab)=a(a^2b)\ \forall a,b\in\mathrm{J}$, *identité de Jordan* qui s'écrit aussi $\mathrm{L}_a\mathrm{L}_{a^2}=\mathrm{L}_{a^2}\mathrm{L}_a$, $\forall a\in\mathrm{J}$.

Un homomorphisme d'une algèbre A, associative, de Lie ou de Jordan, respectivement, dans l'algèbre $\mathfrak{L}(\mathrm{E})$, $\mathfrak{L}(\mathrm{E})_-$ ou $\mathfrak{L}(\mathrm{E})_+$ des endomorphismes d'un espace vectoriel E sur le même corps, qui fait ainsi opérer A dans E, par une *représentation linéaire* détermine sur E une *structure de* A-*module*.

Si E et F sont des A-modules, une application linéaire f de E dans F qui respecte les opérations de A : $f(ax)=af(x)\ \forall a\in\mathrm{A}$, $x\in\mathrm{E}$ est appelée un *morphisme de* A-*modules*.

Le noyau et l'image d'un A-morphisme sont des sous-modules.

Exemples : l'algèbre associative $\mathrm{K}(n)$ des matrices $n\times n$ sur K et les algèbres de Lie : $\mathrm{K}(n)_-=\mathfrak{gl}(n;\mathrm{K})$ et de Jordan $\mathrm{K}(n)_+=\mathrm{J}(n;\mathrm{K})$ associées.

Une structure supplémentaire fréquente est celle de graduation. Sur un espace vectoriel E une *graduation* est simplement une décomposition en somme directe : $\mathrm{E}=\oplus_{,i\in\mathrm{I}}\mathrm{E}_i$ où l'ensemble d'indices I est un ensemble quelconque. Un élément x d'un E_i est *homogène de degré* i. Tout élément x non nul E s'écrit d'une façon et d'une seule comme une somme finie : $x=\sum\limits_{i\in\mathrm{I}}x_i$, où x_i est la composante de degré i de x.

Un sous-espace F de E est dit gradué s'il est égal à la somme de ses intersections avec les E_i : $\mathrm{F}=\oplus_{,i\in\mathrm{I}}(\mathrm{F}\cap\mathrm{E}_i)$ ce qui revient à dire que si $x\in\mathrm{F}$ toutes les composantes x_i de x appartiennent à F.

Sur une algèbre A, une *graduation* est d'abord une graduation de l'espace vectoriel A : $\mathrm{A}=\oplus_{,i\in\mathrm{I}}\mathrm{A}_i$. Mais cette fois on impose à l'ensemble d'indices I de posséder une addition de telle sorte que le degré du produit x_i, y_j de deux éléments homogènes soit la somme des degrés des facteurs, ce qui revient à dire que $\mathrm{A}_i\mathrm{A}_j$ et $\mathrm{A}_j\mathrm{A}_i\subset\mathrm{A}_{i+j}$. Les ensembles d'indices les plus courants sont : $\mathbb{N}$, entiers $\geqslant 0$, $\mathbb{Z}$, entiers quelconques, $\mathbb{Z}_2$, graduation « pair-impair », et leurs produits comme $\mathbb{N}^m=\mathbb{N}\times\mathbb{N}\times\ldots\mathbb{N}$ qu'on appelle souvent multidegrés. Ce sont donc soit des groupes additifs, soit des parties de groupes additifs stables par l'addition et contenant 0, comme $\mathbb{N}$.

Si A est une algèbre graduée avec unité, le degré de l'unité est nécessairement nul.

Suivant le même principe, si deux espaces vectoriels E et F sur le même corps sont gradués par le même ensemble additif d'indices l : $\mathrm{E} = \bigoplus_{i\in \mathrm{I}}$, $\mathrm{F} \bigoplus_{j\in \mathrm{I}} \mathrm{F}_j$, le produit tensoriel $\mathrm{E}\otimes\mathrm{F}$ peut être gradué de deux façons, en attribuant au produit $x_i \otimes y_j$ d'éléments homogènes de degrés i et j, le bidegré $(i, j) \in \mathrm{I}\times\mathrm{I}$ et le degré $(i + j) \in \mathrm{I}$.

Dans le dernier cas : $(\mathrm{E}\otimes\mathrm{F})_r = \bigoplus_{i+j=r} \mathrm{E}_i \otimes \mathrm{F}_j$.

Définition III.11.B

Une algèbre A graduée par $\mathbb{N}, \mathbb{Z}$, ou simplement par $\mathbb{Z}_2$ est dite *anticommutative* si quels que soient les éléments homogènes $x_i \in \mathrm{A}_i$ et $y_j \in \mathrm{A}_j$ de A, leur produit vérifie la propriété d'anticommunication :

$$y_j x_i = (-1)^{ij} x_i y_j.$$

Cette formule fait donc intervenir le *produit* des degrés, et seule la *parité* de ces degrés intervient. Les éléments de degré pair commutent avec tous les autres. Les éléments de ce degré impair anticommutent entre eux.

Exemple : l'algèbre extérieure sur E : $\wedge\mathrm{E}$ qui sera définie au chapitre V.

Si A et B sont deux algèbres sur le même corps, l'application de $\mathrm{A}\times\mathrm{B}\times\mathrm{A}\times\mathrm{B}$ dans $\mathrm{A}\otimes\mathrm{B}$: $(a, b, a', b') \to aa' \otimes bb'$ est quadrilinéaire. Elle définit une application linéaire : $\mathrm{A}\otimes\mathrm{B}\otimes\mathrm{A}\otimes\mathrm{B} = (\mathrm{A}\otimes\mathrm{B})\otimes(\mathrm{A}\otimes\mathrm{B}) \to \mathrm{A}\otimes\mathrm{B}$ qui est une multiplication sur $\mathrm{A}\otimes\mathrm{B}$. L'algèbre $\mathrm{A}\otimes\mathrm{B}$ ainsi définie est dite *produit tensoriel des algèbres* A *et* B. Le produit de deux éléments décomposables de $\mathrm{A}\otimes\mathrm{B}$ est : $(a\otimes b)\cdot(a'\otimes b') = aa' \otimes bb'$. Si A et B sont associatives, resp. commutatives, resp. avec des unités 1_{A} et 1_{B}, $\mathrm{A}\otimes\mathrm{B}$ est associative, resp. commutative, resp. avec une unité $1_{\mathrm{A}} \otimes 1_{\mathrm{B}}$.

Dans ce dernier cas $a \to a\otimes 1_{\mathrm{B}}$ et $b \to 1_{\mathrm{A}} \otimes \mathrm{B}$ sont des homomorphismes injectifs de A et de B dans $\mathrm{A}\otimes\mathrm{B}$ qui identifient A et B à des sous-algèbres de $\mathrm{A}\otimes\mathrm{B}$. Ces dernières commutent dans $\mathrm{A}\otimes\mathrm{B}$: $\forall a \in \mathrm{A}$ et $b \in \mathrm{B}$, $(a\otimes 1_{\mathrm{B}})\cdot(1_{\mathrm{A}}\otimes b) = a\otimes b = (1_{\mathrm{A}}\otimes b)\cdot(a\otimes 1_{\mathrm{B}})$

et elles sont linéairement disjointes : si $(a_i; i \in \mathrm{I})$ est une base de A et $(b_j; j \in \mathrm{J})$ une base de B, les $(a_i \otimes b_j; (i,j) \in \mathrm{I} \times \mathrm{J})$ sont linéairement indépendants et forment ainsi une base de $\mathrm{A} \otimes \mathrm{B}$.

Exemples : si E et F sont deux espaces vectoriels de dimension finie sur le même corps, $\mathcal{L}(\mathrm{E}) \otimes \mathcal{L}(\mathrm{F})$ s'identifie à $\mathcal{L}(\mathrm{E} \otimes \mathrm{F})$.

En prenant des bases on obtient l'isomorphisme entre algèbres de matrices : $\mathrm{K}(m) \otimes \mathrm{K}(n) \equiv \mathrm{K}(mn)$.

L'importance des algèbres anticommutatives conduit à définir une variante qui leur est adaptée du produit tensoriel d'algèbres.

Définition III.1.C :

Soient A et B deux algèbres sur le même corps K *graduées* par $\mathbb{N}$, $\mathbb{Z}$ ou $\mathbb{Z}_2$.

On appelle *produit tensoriel gauche* de ces algèbres et l'on note $\mathrm{A}\widehat{\otimes}\mathrm{B}$, l'espace gradué $\mathrm{A} \otimes \mathrm{B}$ muni de la multiplication de l'algèbre $\mathrm{A} \otimes \mathrm{B}$ modifiée comme suit : sur les éléments *homogène* a, $a' \in \mathrm{A}$ et $b, b' \in \mathrm{B}$, on définit $(a \otimes b) \cdot (a' \otimes b') = (-1)^{\deg b \cdot \deg a'} aa' \otimes bb'$.

Seuls interviennent pour le changement de signe les éléments b et a', qui échangent leur place, et seule intervient la parité de $\deg a'$ et $\deg b$. Des calculs simples permettent de vérifier que si A et B sont associatives, resp. *anticommutatives*, resp. avec des unités 1_{A} et 1_{B}, $\mathrm{A}\widehat{\otimes}\mathrm{B}$ est associative, resp. *anticommutative*, resp. avec une unité $1_{\mathrm{A}} \otimes 1_{\mathrm{B}}$.

III.12 – L'algèbre tensorielle sur un espace vectoriel

Soit E un espace vectoriel de dimension finie ou infinie sur un corps K. On peut former suivant le chapitre II toutes ses puissances tensorielles $\otimes^p \mathrm{E}$ et, en convenant que $\otimes^0 \mathrm{E} = \mathrm{K}$ et $\otimes^1 \mathrm{E} = \mathrm{E}$, considérer la somme directe :

$$\otimes \mathrm{E} = \bigoplus_{p=0}^{\infty} (\otimes^p \mathrm{E}) = \mathrm{K} \oplus \mathrm{E} \oplus (\otimes^2 \mathrm{E}) \oplus \dots$$

C'est un espace vectoriel gradué par les entiers naturels sur lequel la multiplication tensorielle définit une structure d'algèbre associative graduée avec unité. $\otimes \mathrm{E}$ est par définition *l'algèbre tensorielle sur* E.

L'importance de l'algèbre tensorielle réside dans sa *propriété universelle*, qui la définit à un isomorphisme canonique près :

Théorème III.12 : Toute application linéaire f de l'espace vectoriel E dans une algèbre A associative, avec unité, sur le même corps, se prolonge de façon unique en un homomorphisme d'algèbres avec unité $\widetilde{f}$ de $\otimes$E dans A.

Preuve : $\widetilde{f}$ est immédiatement déterminé sur $\otimes^0$E par $f(\lambda \cdot 1) = \lambda 1_{\mathrm{A}}$ et sur $\otimes^1\mathrm{E} = \mathrm{E}$ par $\widetilde{f}|_{\mathrm{E}} = f$. Son extension à $\otimes$E est unique : l'application p-linéaire : $(x_1, x_2, \ldots, x_p) \to f(x_1)f(x_2)\ldots f(x_p)$ de $\mathrm{E} \times \mathrm{E} \times \ldots \mathrm{E}$ dans A détermine l'application linéaire $\widetilde{f}$ de $\otimes^p$E dans A. Sur des tenseurs décomposables : $t \in \otimes^p\mathrm{E}$ et $u \in \otimes^q\mathrm{E}$, la définition que l'on vient de donner de $\widetilde{f}$ entraîne $\widetilde{f}(tu) = \widetilde{f}(t)\widetilde{f}(u)$. Par linéarité $\widetilde{f}$ est bien un isomorphisme de $\otimes$E dans A.

Corollaire III.12 : Une application linéaire f d'un espace vectoriel E dans un espace vectoriel F sur le même corps se prolonge en un homomorphisme d'algèbres unitaires graduées $f^{\otimes}$ de $\otimes$E dans $\otimes$F. C'est la propriété « fonctorielle » de l'algèbre tensorielle, car si h est maintenant une application linéaire de F dans H, on a : $h^{\otimes} \circ f^{\otimes} = (h \circ f)^{\otimes}$.

Remarque : Si $e = \{e_i; i \in \mathrm{I}\}$ est une base de E. Il y a une bijection entre les suites de p éléments de I : $(i_1 i_2 \ldots i_p)$ et les éléments de la base $\otimes^p e$ de $\otimes^p$E : $e_{i_1} \otimes e_{i_2} \ldots \otimes e_{i_p}$. Au produit des suites obtenu par juxtaposition, correspond le produit tensoriel des éléments de base qui leurs sont associés. On a donc un isomorphisme naturel entre $\otimes$E et l'algèbre libre sur K de l'ensemble I, puisque $\mathrm{E} \equiv \mathrm{K}^{(\mathrm{I})}$.

III.13 – Problèmes universels

Nous avons rencontré à plusieurs reprises des constructions algébriques présentées comme solutions de problèmes universels. Cette notion de problème universel est un concept éminemment simplificateur, et son usage est tel qu'il n'est pas inutile d'en exposer les éléments même très succintement. Cela concerne deux catégories différentes d'objets mathématiques. Pour nous une catégorie est formée d'*objets*, qui sont des ensembles munis d'une certaine structure : $\mathfrak{E}_K$, espaces vectoriels sur K, $\mathcal{T}$, espaces topologiques, $\mathfrak{M}$ espaces métriques, $\mathfrak{A}_K$ algèbres associatives avec unité sur K, $\mathfrak{g}_K$ algèbres de Lie sur K, etc..., et de *morphismes*, qui sont les applications conservant la structure : applications linéaires, continues, homomorphismes, etc... Soient deux catégories $\mathfrak{C}_0$ et $\mathfrak{C}_1$ et $\mathfrak{F}$ une famille d'applications f des objets A de $\mathfrak{C}_0$ dans les objets B de $\mathfrak{C}_1$ de telle sorte que pour tout morphisme a dans $\mathfrak{C}_0$ ou b dans $\mathfrak{C}_1$, $f \circ a$ et $b \circ f$ soient encore des applications de $\mathfrak{F}$. Dans ces conditions, on a le *Problème universel posé par* $(\mathfrak{C}_0, \mathfrak{C}_1, \mathfrak{F})$: faire correspondre à chaque objet A de $\mathfrak{C}_0$ un *objet* U(A) de $\mathfrak{C}_1$ et une *application* $u \in \mathfrak{F}$ de A dans U(A) tels que pour tout $f \in \mathfrak{F}$, de A dans B $\in \mathfrak{C}_1$ il existe un morphisme φ de U(A) dans B, *unique* tel que $f = \varphi \circ u$. Cela s'exprime par le « diagramme universel » :

$$\begin{tikzcd} A \arrow[rr, "f"] \arrow[dr, "u"'] & & B \\ & U(A) \arrow[ur, "\varphi"'] & \end{tikzcd} \qquad \text{avec } \begin{cases} f = \varphi \text{ ou} \\ \varphi \text{ unique} \end{cases}$$

U(A) représente A dans l'autre catégorie $\mathfrak{C}_1$ et u sert à représenter toutes les applications de A dans $\mathfrak{C}_1$ par des morphismes de $\mathfrak{C}_1$.

Si un problème universel possède une solution, elle est évidemment unique à un isomorphisme canonique près puisque du diagramme universel :

$$\begin{tikzcd} A \arrow[rr, "u_1"] \arrow[dr, "u_2"'] & & U_1(A) \arrow[dl, shift left, "\varphi_1"] \\ & U_2(A) \arrow[ur, shift left, "\varphi_2"] & \end{tikzcd} \qquad \text{avec } \begin{cases} u_2 = \varphi_2 \text{ ou }_1 \\ u_1 = \varphi_1 \text{ ou }_2 \end{cases}$$

il résulte : $u_2 = (\varphi_2 \circ \varphi_1) \circ u_2 = (\mathrm{Id}\ U_2) \circ u_2$ et $u_1 = (\varphi_1 \circ \varphi_2) \circ u_1 =$ (Id $U_1 \circ u_1$. Mais la condition d'unicité entraîne $\varphi_2 \circ \varphi_1 = \mathrm{Id}\ U_2$ et $\varphi_1 \circ \varphi_2 = \mathrm{Id}\ U_1$ ce qui montre que φ_1 et φ_2, uniques, sont des isomorphismes inverses l'un de l'autre.

Quelques exemples :

1) Si l'on prend pour $\mathfrak{C}_0$ la catégorie des ensembles et pour $\mathfrak{C}_1$ une sous-catégorie dont les objets sont des ensembles munis d'une structure algébrique donnée, la solution du problème universel ainsi

posé, lorsqu'elle existe, associe à chaque ensemble X un objet U(X) de $\mathfrak{C}_1$ appelé *objet libre* de $\mathfrak{C}_1$ engendré par X.
$\mathfrak{C}_1 = \mathrm{E_K}$, espaces K-vectoriels : $\mathrm{U(X)} = \mathrm{K^{(X)}}$
$\mathfrak{C}_1 = \{$K-algèbres associatives, commutatives, avec unité$\}$
U(X) = K[X], algèbre des polynômes en les éléments de X, sur K...

2) $\mathfrak{C}_0 = \{$couples d'espaces K– vectoriels (E, F) avec pour morphismes les couples d'applications linéaires$\}$,

$$\mathfrak{C}_1 = \mathfrak{E}_{\mathrm{K}}, \quad \mathfrak{F} = \{\text{applications bilinéaires}\} : \mathrm{U(E,F)} = \mathrm{E} \otimes \mathrm{F}.$$

3) Le produit tensoriel de deux algèbres associatives unitaires A et B sur le même corps K, est caractérisé par la propriété universelle suivante : quel que soit le couple (φ, ψ) d'homomorphismes de A et B dans une algèbre unitaire C qui *commutent*, il détermine un homomorphisme unique $\varphi \otimes \psi$ de $\mathrm{A} \otimes \mathrm{B}$ dans C, qui coïncide avec φ sur $\mathrm{A} \otimes 1_{\mathrm{B}}$ et avec ψ sur $1_{\mathrm{A}} \otimes \mathrm{B}$.

4) $\mathfrak{C}_0 = \{$algèbres de Lie sur le corps K$\}$, $\mathfrak{C}_1 = \{$algèbres associatives avec unité sur K$\}$, $\mathfrak{F} = \{$homomorphismes pour le crochet $f([x,y]) = [f(x), f(y)] = f(x)f(y) - f(y)f(x)\}$.
$\mathrm{U}(\mathfrak{g})$ = algèbre enveloppante de l'algèbre de Lie $\mathfrak{g}$.
$\mathrm{U}(\mathfrak{g})$ est l'algèbre unitaire quotient de l'algèbre tensorielle $\otimes\mathfrak{g}$ sur l'espace vectoriel $\mathfrak{g}$ par l'idéal bilatère engendré par les éléments (inhomogènes) : $\{x \otimes y - y \otimes x - [x,y] ; \forall x, y \in \mathfrak{g}\}$.

III.14 – Invariants. Extensions tensorielles des représentations linéaires d'algèbres de Lie

Etant donnée une représentation d'un ensemble muni d'une structure algébrique A par des endomorphismes d'un espace vectoriel E, on peut se demander si elle est susceptible de se prolonger aux espaces tensoriels bâtis sur E et de donner ainsi naissance à toute une famille de représentations linéaires associées. On a vu au §II.7, II.8 et III.4 ci-dessus que c'est bien le cas lorsque A est un groupe. Mais il n'en est pas de même, par exemple pour la structure d'algèbre associative, qui *ne permet pas* l'extension tensorielle de ses représentations linéaires. Les représentations linéaires d'algèbres associatives seront étudiées au §VI.4 à propos des algèbres de Clifford.

Parmi les algèbres, seule la structure d'algèbre de Lie autorise cette extension tensorielle, et cela n'est guère étonnant si l'on considère la structure d'algèbre de Lie comme une linéarisation de celle du groupe ! Rappelons la :

Définition III.14.A :

Une représentation linéaire d'une algèbre de Lie $\mathfrak{g}$ dans un espace vectoriel sur le même corps E est un homomorphisme ρ de $\mathfrak{g}$ dans l'algèbre de Lie $\mathcal{L}(\mathrm{E})_-$. On a donc :

$$\forall a, b \in \mathfrak{g} : \rho([a, b] = [\rho(a), \rho(b)] = \rho(a)\rho(b) - \rho(b)\rho(a).$$

En particulier, l'identité de Jacobi montre que les translations à gauche de $\mathfrak{g}$: $\forall a, x \in \mathfrak{g}$, $ada \cdot x = [a, x]$ définissent une représentation linéaire de $\mathfrak{g}$ dans elle-même, dite *représentation adjointe*.

La propriété universelle des algèbres enveloppantes se traduit dans le cas des représentations linéaires par le :

Théorème III.14.A : Soient $\mathfrak{g}$ une algèbre de Lie, $\mathrm{U}(\mathfrak{g})$ son algèbre enveloppante (§III.13), quotient de l'algèbre tensorielle $\otimes\mathfrak{g}$ par l'idéal bilatère engendré par les éléments $\{x\otimes y - y\otimes x - [x, y]\}$. Toute représentation linéaire ρ de $\mathfrak{g}$ dans un espace vectoriel E sur le même corps se prolonge de façon unique en un homomorphisme $\overline{\rho}$ d'algèbres (associatives) unitaires de $\mathrm{U}(\mathfrak{g})$ dans $\mathcal{L}(\mathrm{E})$. Réciproquement un tel homomorphisme détermine une représentation linéaire de $\mathfrak{g}$ dans E. Un $\mathfrak{g}$-module est donc automatiquement un $\mathrm{U}(\mathfrak{g})$-module.

Preuve : D'après le théorème III.12, ρ se prolonge de façon unique en un homomorphisme d'algèbre unitaires $\widetilde{\rho}$ de $\otimes\mathfrak{g}$ dans $\mathcal{L}(\mathrm{E})$ qui s'annule sur les générateurs de l'idéal bilatère : $\widetilde{\rho}(x \otimes y - y\otimes x - [x, y]) = \rho(x)\rho(y) - \rho(y)\rho(x) - \rho([x, y]) = 0$, et passe donc au quotient. Réciproquement, si φ est un homomorphisme d'algèbres unitaires de $\mathrm{U}(\mathfrak{g})$ dans $\mathcal{L}(\mathrm{E})$ la composée des applications :

$$\mathfrak{g} \longrightarrow \otimes\mathfrak{g} \longrightarrow \mathrm{U}(\mathfrak{g}) \xrightarrow{\varphi} \mathcal{L}(\mathrm{E})$$

est une représentation de $\mathfrak{g}$ dans E.

Nous allons maintenant examiner successivement les notions d'invariant et d'extensions des représentations dans le cas des algèbres de Lie.

Lorsqu'un groupe G opère comme groupe de transformations d'un ensemble E, un *invariant* est un élément x de E tel que $gx = x$ quel que soit $g \in$ G. Pour obtenir la définition correspondante pour les algèbres de Lie, on fait le raisonnement suivant : E étant un espace vectoriel réel de dimension finie et G un sous-groupe fermé de Gl(E), on considère un chemin différentiable $u(t)$ tout entier contenu dans G et tel que $u(0) = \mathrm{Id}$.

Si x est un invariant de G, on a $u(t)x = x$ quel que soit t, d'où $u'(0)x = 0$: x est annulé par tous les vecteurs tangents en l'élément neutre de G. D'où la :

Définition III.14.B :

Soient $\mathfrak{g}$ une algèbre de Lie sur un corps K, ρ une représentation linéaire de $\mathfrak{g}$ dans un espace vectoriel E sur le même corps. Un *vecteur x de* E *est dit Lie-invariant, ou simplement invariant si* $\rho(\mathrm{X}) \cdot x = 0$ *pour tout* $\mathrm{X} \in \mathfrak{g}$: $x \in \cap \mathrm{Ker}\, \rho(x)$. En particulier, tout vecteur x d'un espace vectoriel réel E, invariant par les opérateurs d'une représentation linéaire ρ d'un groupe de Lie G : $\rho(g)x = x$, $\forall g \in$ G *est aussi* un invariant de la représentation associée de l'algèbre de Lie $\mathfrak{g}$ de G.

Dans tout ce qui suit E désigne un espace vectoriel de dimension finie, E* son dual et $\langle x, \alpha \rangle$ le crochet de dualité entre $x \in$ E et $\alpha \in$ E*. Si $u \in \mathfrak{L}(\mathrm{E})$, son transposé ${}^t u \in \mathfrak{L}(\mathrm{E}^*)$ est défini par ${}^t u \cdot \alpha = \alpha \circ u$, ou $\langle ux, \alpha \rangle = \langle x, {}^t u\alpha \rangle$.

Si $u \in \mathrm{Gl(E)}$, le contragrédient $\check{u} = {}^{t}u^{-1}$ définit la représentation contragrédiente de Gl(E) dans E* de telle sorte que le crochet de dualité entre E et E* soit invariant : $\langle ux, \check{u}\alpha \rangle = \langle x, \alpha \rangle$. A toute représentation linéaire r d'un groupe G dans E est ainsi associée la représentation contragrédiente : $g \in \mathrm{G} \to \widecheck{r(g)}$ dans E*.

Le passage (par « dérivation ») d'une propriété des groupes à la propriété correspondante des algèbres de Lie entraîne la définition suivante :

Définition III.14.C :

Si ρ est une représentation de l'algèbre de Lie $\mathfrak{g}$ dans E, la *représentation contragrédiente* $\widetilde{\rho}$ de $\mathfrak{g}$ dans E* est définie par :

$$\widetilde{\rho}(\mathrm{X}) \cdot \alpha = -{}^t\rho(\mathrm{X}) \cdot \alpha = -\alpha \circ \rho(\mathrm{X}),$$

de telle sorte que le crochet de dualité soit Lie invariant (voir ci-dessous Exemple 2) :

$$\langle \rho(\mathrm{X})x, \alpha\rangle + \langle x, \widetilde{\rho}(\mathrm{X})\alpha\rangle = 0 \quad \forall \mathrm{X} \in \mathfrak{g},$$

(On vérifie bien que $\widetilde{\rho}([\mathrm{X}, \mathrm{Y}]) = ([\widetilde{\rho}(\mathrm{X}), \widetilde{\rho}(\mathrm{Y})])$.

Si r est une représentation linéaire du groupe G dans E, $g \in \mathrm{G} \to \otimes^{(v)} r(g)$ est la représentation « puissance tensorielle $(v)^{ème}$ de r'', de G dans $\otimes^{(v)}\mathrm{E}$.

Par « dérivation », on passe aux définitions et propriétés dites « de Lie » :

Définition III.14.D :

Extensions tensorielles et invariants de Lie d'un opérateur linéaire.

Soit $a \in \mathfrak{L}(\mathrm{E})$. *L'extension tensorielle « de Lie » de* a à $\otimes^{(v)}\mathrm{E}$, notée $\widetilde{\otimes}^{(v)} a$ (où le signe $\sim$ distingue cette extension de l'extension tensorielle usuelle « multiplicative ») : $\bar{\otimes}^{(v)} a)$ est l'opérateur linéaire de $\otimes^{(v)}\mathrm{E}$ somme de $|v|$ opérateurs : le j-ème est le produit tensoriel des identités dans chaque facteur à l'exception du j-ème où l'on prend a si ce facteur est E, $(-{}^t a)$ si ce facteur est E^*. On a : $\widetilde{\otimes}^{(v)}[a, b] = [\widetilde{\otimes}^{(v)} a, \widetilde{\otimes}^{(v)} b]$.

Les éléments du noyau de $\widetilde{\otimes}^{(v)} a$ sont appelés *les invariants de Lie de variance* (v) de l'opérateur a.

Remarque : On peut rassembler les extensions « de Lie » de $a \in \mathfrak{L}(\mathrm{E})$ dans toutes les puissances tensorielles $\otimes^p\mathrm{E}$. En complétant par la condition $a \cdot 1 = 0$, on obtient ainsi un opérateur linéaire $\widetilde{\otimes} a$ de $\otimes\mathrm{E}$: c'est la dérivation de l'algèbre tensorielle $\otimes\mathrm{E}$ associée à a ce que l'on retrouvera au §IV.13.

Définition III.14.E :

Extensions tensorielles d'une représentation linéaire et invariants d'une algèbre de Lie.

Si ρ est une représentation linéaire d'une algèbre de Lie $\mathfrak{g}$ dans E, elle détermine :

a) pour chaque variance (v) une représentation linéaire de $\mathfrak{g}$ dans $\otimes^{(v)}\mathrm{E}$, *extension tensorielle « de Lie » de* ρ : $\mathrm{X} \in \mathfrak{g} \to \widetilde{\otimes}^{(v)} \rho(\mathrm{X})$.

b) une représentation $\widetilde{\otimes}\rho$ de $\mathfrak{g}$ dans l'algèbre $\otimes$E où elle opère comme une algèbre de Lie de dérivations : $\forall X \in \mathfrak{g}$:

$$\widetilde{\otimes}\rho(X)(a \otimes b) = \{\widetilde{\otimes}\rho(X)a\} \otimes b + a \otimes \{\widetilde{\otimes}\rho(X)b\}.$$

c) d'une façon analogue, une représentation $\widetilde{\otimes}\check{\rho}$ dans $\otimes E^*$.

On a vu que le sous-espace $\mathrm{Inv}(\rho; (v)) = \{\cap \mathrm{Ker}\widetilde{\otimes}^{(v)}\rho(X)\,; \forall X \in \mathfrak{g}\}$ de $\widetilde{\otimes}^{(v)}E$ est appelé le sous-espace des invariants de variance (v) de ρ. L'ensemble de ces sous-espaces forme les *invariants tensoriels de la représentation* ρ. Réciproquement, si l'on se donne un ensemble d'éléments $f_i \in \otimes^{(v_i)}E$, les $a \in \mathfrak{L}(E)$ tels que $\widetilde{\otimes}^{(v_i)}a \cdot (f_i) = 0$ forment une sous-algèbre de Lie de $\mathfrak{L}(E)_-$ puisque ces conditions sont linéaires et stables par le crochet.

Si ρ est la représentation adjointe, les éléments de $\cap\mathrm{Ker}\widetilde{\otimes}^{(v)}\mathrm{ad}\, a$ dans $\otimes^{(v)}\mathfrak{g}$ sont les invariants de variance (v) de $\mathfrak{g}$. Leur ensemble forme les *invariants tensoriels de l'algèbre de Lie* $\mathfrak{g}$ dans $\mathfrak{T}(\mathfrak{g})$.

Comme pour les extensions tensorielles du groupe Gl(E) aux puissances tensorielles $\otimes^p E$ ou $\otimes^p E^*$ (théorèmes III.4.A et B), les extensions tensorielles d'une représentation linéaire ρ d'une algèbre de Lie $\mathfrak{g}$ dans E à $\otimes^p E$ ou $\otimes^p E^*$ *commutent avec les permutations de facteurs*, donc avec les opérations du groupe symétrique $\mathfrak{S}_p$ sur $\otimes^p E$ et $\otimes^p E^*$. Si u est un invariant, le symétrisé : $\mathrm{s}u$, et l'antisymétrisé : $\mathbf{a}u$, sont également des invariants.

Exemples d'invariants : ($\mathfrak{g}$ et E de dimension finies)

1) Si $(v) = (.*)$, $\otimes^{(v)}a = a \otimes \mathrm{Id} - \mathrm{Id} \otimes^t a$. Sur un élément $x \otimes \alpha$.

$$(\widetilde{\otimes}^{(v)}a) \cdot (x \otimes \alpha) = ax \otimes \alpha - x \otimes (\alpha \circ a).$$

Si l'on interprète un élément u de $E \otimes E^*$ comme un opérateur linéaire de E, l'action « de Lie » de $\otimes^{(v)}a$ sur u s'écrit :

$$((\widetilde{\otimes}^{(v)}a \cdot u)(x) = au(x) - u(ax)$$

et u est Lie-invariant par a si et seulement s'il commute avec a.

2) $(v) = (**)$. Un élément b de $E^* \otimes E^*$ peut être interprété (§II.3) comme une forme bilinéaire sur E. Soient φ et ψ les applications associées à droite et à gauche de E dans E^* :

$$b(x, y) = \langle x, \varphi y\rangle = \langle \psi x, y\rangle.$$

Considérons d'abord le cas où un groupe G opère dans E par un homomorphisme $G \to Gl(E)$, et dans E^* par la représentation contragrédiente : $g \to \check{g} = {}^t \overset{-1}{g}$. Dire que la forme bilinéaire b est invariante par G est équivalent à dire que φ ou ψ est un G-morphisme de E dans E^*.

En effet : $b(gx, gy) = \langle gx, \varphi gy)\rangle = \langle x, {}^t g\varphi gy\rangle$ et l'égalité : $b(gx, gy) = b(x, y)$ quels que soient $x, y \in E$ est équivalente à $\varphi = {}^t g\varphi g$ soit $\varphi g = {}^t \overset{-1}{g} \varphi$.

Réciproquement, tout G-morphisme φ de E dans E^* détermine une forme bilinéaire invariante sur E :

$$b(x, y) = \langle x, \varphi y\rangle \text{ puisque } b(gx, gy) =$$
$$\langle gx, \varphi gy\rangle = \langle x, {}^t g\varphi gy\rangle = \langle x, \varphi y\rangle = b(x, y).$$

Si b et b' sont deux formes bilinéaires sur E, invariantes par G, et si b est non-dégénérée (φ inversible), b' s'exprime à l'aide de b :

$$b'(x, y) = \langle x, \varphi' y\rangle = \langle x, \varphi \overset{-1}{\varphi} \varphi' y\rangle = b(x, \overset{-1}{\varphi} \varphi' y)$$

et de $\gamma = \overset{-1}{\varphi} \varphi'$, qui est un endomorphisme du G-module E.

On dit aussi que γ appartient au « *commutant* » de E. Réciproquement tout élément γ du commutant de E fait passer d'une forme bilinéaire quelconque invariante b à une forme $b'(x, y) = b(x, \gamma y)$ également invariante.

Voyons maintenant le cas où c'est une algèbre de Lie $\mathfrak{g}$ qui opère dans E par un homomorphisme d'algèbres de Lie $\rho : \mathfrak{g} \to \mathfrak{L}(E)_-$. Pour simplifier l'écriture, notons simplement, (comme dans le cas d'un groupe), par ax l'image de $x \in E$ par $\rho(a)$. La représentation contragrédiente dans E^* est donc $(-{}^t a) \cdot \alpha = -\alpha \circ a$, et l'opérateur $\widetilde{\otimes}^{(v)} a$ dans $E^* \otimes E^*$ s'écrit, sur $\alpha \otimes \beta$ décomposable :

$$\widetilde{\otimes}^{(v)} a \cdot (\alpha \otimes \beta) = -(\alpha \circ a) \otimes \beta - \alpha \otimes (\beta \circ a)$$

sur b :
$$\begin{aligned}(\otimes^{(v)} a \cdot b)(x, y) &= -b(ax, y) - b(x, ay) \\ &= -\langle x, ({}^t a\varphi + \varphi a) y\rangle \\ &= -\langle(\psi a + {}^t a\psi)x, y\rangle.\end{aligned}$$

Il en résulte que les trois propriétés suivantes sont équivalentes :

1) b est un invariant de Lie : $b(ax,y)+b(x,ay)=0$;

2) φ est un morphisme de $\mathfrak{g}$-modules de E dans E* : $\varphi a = -{}^t a\varphi$;

3) ψ est un morphisme de $\mathfrak{g}$-modules : $\psi a = -{}^t a\psi$.

Réciproquement, tout $\mathfrak{g}$-morphisme φ de E dans E* détermine une forme bilinéaire Lie-invariante sur E : $b(x,y)=\langle x,\varphi y\rangle$

Si b et b' sont deux formes bilinéaires Lie-invariantes sur E, et si b est non-dégénérée, on peut écrire $b'(x,y)=b(x,\overset{-1}{\varphi}\varphi' y)$ et $\overset{-1}{\varphi}\varphi'=\gamma$ est un endomorphisme du $\mathfrak{g}$-module de E, ou élément du commutant de E. Réciproquement, tout élément γ du commutant de E fait passer d'une forme bilinéaire Lie-invariante quelconque b à une autre $b'(x,y)=b(x,\gamma y)$ également Lie-invariante car :

$$b(ax,\gamma y)+b(x,\gamma ay)=b(ax,\gamma y)+b(x,a\gamma y)=0.$$

Les formes bilinéaires invariantes sur le $\mathfrak{g}$-module E forment un espace vectoriel $\mathfrak{B}_0$ identifiable au sous-espace des tenseurs invariants de $E^*\otimes E^*$, stable par la transposition et par les opérations contragrédientes du commutant de E.

Il en résulte de ce qui précède qu'aussi bien dans le cas d'un groupe G que d'une algèbre de Lie $\mathfrak{g}$, il n'existe de forme bilinéaire invariante *non-dégénérée* sur un module E de dimension finie que si et seulement si le module E est isomorphe au module E* muni de la représentation contragrédiente.

Si le G ou $\mathfrak{g}$-module est *simple*, c'est-à-dire ne possède aucun sous-module propre, tout G ou $\mathfrak{g}$-morphisme de E dans lui-même non nul ne peut être qu'un automorphisme : tous les éléments non nuls du commutant de E sont inversibles, et le commutant Γ de E est un corps, éventuellement non commutatif, extension de degré fini du corps K initial.

Lorsque Γ *se réduit au corps* K, si $\mathfrak{B}_0$ n'est pas nul, toute forme bilinéaire invariante b' se déduit de l'un d'elles b par multiplication par un scalaire : $b'(x,y)=b(x,\gamma y)=\gamma b(x,y)$ et $\dim\mathfrak{B}_0=1$. On a en particulier si $b\neq 0$: $b(y,x)=\lambda b(x,y)=\lambda^2 b(y,x)$ d'où $\lambda=\pm 1$: b est symétrique, ou antisymétrique. Nous avons démontré la :

Proposition III.14.A :

Soit E un module *simple* de dimension finie sur un groupe G ou une algèbre de Lie $\mathfrak{g}$ dont le commutant se réduit au corps K de E.

Si E est isomorphe à son module dual E^*, il existe sur E une forme bilinéaire unique à un scalaire près, qui est ou bien symétrique, ou bien antisymétrique.

3) si $(v) = (.**)$, l'extension à $E \otimes E^* \otimes E^*$ d'une représentation de l'algèbre de Lie $\mathfrak{g}$ dans E s'écrit sur un élément décomposable :

$$(\widetilde{\otimes}^{(v)} a) \cdot (x \otimes \alpha \otimes \beta) = ax \otimes \alpha \otimes \beta - x \otimes (\alpha \circ a) \otimes \beta \\ -x \otimes \alpha \otimes (\beta \circ a).$$

Un élément μ de l'espace $\otimes^{(v)}E = E \otimes E^* \otimes E^*$ est une *multiplication bilinéaire* sur E : si $\mu = \sum z_i \otimes \alpha_i \otimes \beta_i$, $\mu(x, y) = \sum \langle \alpha_i, x\rangle \langle \beta_i, y\rangle z_i$.

L'action « de Lie » de $\widetilde{\otimes}^{(v)} a$ sur μ s'écrit :

$$((\widetilde{\otimes}^{(v)} a) \cdot \mu)(x, y) = a\mu(x, y) - \mu(ax, y) - \mu(x, ay)$$

et *dire que μ est Lie-invariant par a revient à dire que a est une μ-dérivation.*

Si $\mathfrak{g}$ est une algèbre de Lie et $\mu \in \otimes^{(v)}\mathfrak{g}$ son crochet, puisque les $\operatorname{ad} a$, $\forall a \in \mathfrak{g}$ sont des dérivations, *μ est toujours un invariant de $\mathfrak{g}$.*

4) les algèbres de Lie classiques sont définies comme les sous-algèbres des algèbres de Lie $\mathfrak{L}(E)_-$ d'endomorphismes linéaires ayant un invariant donné. Pour $(v) = (**)$ $b \in E^* \otimes E^*$ représente une forme bilinéaire sur E. Soit $\mathfrak{g}(b)$ la sous-algèbre de Lie de $\mathfrak{L}(E)_-$ formée des $a \in \mathfrak{L}(E)$ laissant b invariant. Si b est non-dégénérée symétrique, $\mathfrak{g}(b)$ est l'algèbre de Lie orthogonale $\mathfrak{S}O(E; b)$, si b est non-dégénérée antisymétrique $\mathfrak{g}(b)$ est l'algèbre de Lie symplectique $\mathfrak{S}p(E; b)$. Si $(v) = (.**)$ et $\mu \in E \otimes E^* \otimes E^*$, la sous-algèbre de Lie $\mathfrak{g}(\mu)$ formée des $a \in \mathfrak{L}(E)$ laissant μ invariant est l'algèbre de Lie des dérivations de l'algèbre $(E; \mu)$.

Une sous-algèbre de Lie de $\mathfrak{L}(E)_-$ qui est identique à l'algèbre de Lie de ses invariants est dite *algébrique*. Elle est en effet définie par l'annulation d'expressions algébriques.

Remarque : On peut vérifier que $\widetilde{\otimes}^{(v)} \rho$ est une représentation linéaire de $\mathfrak{g}$ dans $\widetilde{\otimes}^{(v)}E$ en utilisant la proposition suivante, évidente, et qui a d'autres applications :

Proposition III.14.B :

Si $\rho_1, \rho_2, \ldots, \rho_p$ sont p représentations linéaires de $\mathfrak{g}$ dans un même espace vectoriel qui *commutent* entre elles : $[\rho_i(\mathrm{X}), \rho_j(\mathrm{X})] = 0\ \forall i \neq j$, $\mathrm{X} \in \mathfrak{g}$, alors : $\mathrm{X} \rightarrow \rho_1(\mathrm{X}) + \rho_2(\mathrm{X}) + \ldots + \rho_p(\mathrm{X})$ est une représentation linéaire de $\mathfrak{g}$ *somme des* ρ_j car $\left(\sum \rho_j)([\mathrm{X},\mathrm{Y}]\right) = \sum \rho_j(\mathrm{X}) \cdot \sum \rho_j(\mathrm{Y}) - \sum \rho_j(\mathrm{Y}) \cdot \sum \rho_j(\mathrm{X})$.

Exemples d'applications de la Proposition III.14.B :

1) Si ρ et ρ' sont des représentations linéaires de $\mathfrak{g}$ dans E et F, ρ et ρ' déterminent une représentation linéaire θ de $\mathfrak{g}$ dans $\mathfrak{L}(\mathrm{E};\mathrm{F})$:

si $f \in \mathfrak{L}(\mathrm{E};\mathrm{F})$, $\theta(\mathrm{X})f = \rho'(\mathrm{X}) \circ f - f \circ \rho(\mathrm{X})$

soit, si $x \in \mathrm{E}$: $(\theta(\mathrm{X})f)(x) = \rho'(\mathrm{X}) \cdot f(x) - f(\rho(\mathrm{X})x)$ une représentation β de $\mathfrak{g}$ dans l'espace vectoriel $\mathfrak{B}(\mathrm{E};\mathrm{F})$ des formes bilinéaires sur $\mathrm{E} \times \mathrm{F}$:

si $b \in \mathfrak{B}(\mathrm{E};\mathrm{F})$, $(\beta(\mathrm{X}) \cdot b)(x,y) = -b(\rho(\mathrm{X})x, y) - b(x, \rho'(\mathrm{X})y)$ une représentation $\rho \otimes \mathrm{I} + \mathrm{I} \otimes \rho'$ dans $\mathrm{E} \otimes \mathrm{F}$:

$$(\rho \otimes \mathrm{I} + \mathrm{I} \otimes \rho')(x \oplus y) = \rho(\mathrm{X})x \otimes y + x \otimes \rho'(\mathrm{X})y$$

qui redonne les deux précédentes lorsqu'on identifie $\mathfrak{L}(\mathrm{E};\mathrm{F})$ à $\mathrm{F} \otimes \mathrm{E}^*$ et $\mathfrak{B}(\mathrm{E};\mathrm{F})$ à $\mathrm{E}^* \otimes \mathrm{F}^*$.

2) dans l'espace vectoriel des applications p-linéaires de $\mathrm{E} \times \mathrm{E} \times \mathrm{E} \ldots \mathrm{E}$ dans F, ρ et ρ' déterminent une représentation linéaire θ par :

si $x_1, x_2, \ldots, x_p \in \mathrm{E}$:

$$(\theta(x)f)(x_1, \ldots, x_p) = \rho'(\mathrm{X}) \cdot f(x_1, \ldots, x_p) - \sum_{i=1}^{p} f(x_1, \ldots, \rho(\mathrm{X})x_i, \ldots, x_p).$$

Soient $g \in \mathrm{Gl}(\mathrm{E})$, $a \in \mathfrak{L}(\mathrm{E})$, $t \in \otimes^{(v)}\mathrm{E}$, $u \in \otimes^{(w)}\mathrm{E}$. Le produit de t et de u est un élément $t \otimes u$ de $\otimes^{(vw)}\mathrm{E}$. De la même façon que l'on a : $g \cdot (t \otimes u) = gt \otimes gu$ par distribution de l'opérateur g à tous les facteurs, on a pour l'extension de Lie de a :

$$(\widetilde{\otimes}^{(vw)}a) \cdot (t \otimes u) = (\widetilde{\otimes}^{(v)}a \cdot t) \otimes u + t \otimes (\widetilde{\otimes}^{(w)}a \cdot u).$$

Il en résulte que :

- le produit de deux invariants de Lie a est un invariant de Lie;
- si l'on se restreint aux tenseurs contravariants, l'extension de Lie de a est une dérivation de l'algèbre tensorielle $\otimes\mathrm{E}$;
- si l'on se restreint aux tenseurs covariants, l'extension de Lie de a est une dérivation de l'algèbre $\otimes\mathrm{E}^*$;
- d'une façon générale, l'extension de Lie de a est une dérivation de l'algèbre tensorielle complète sur E : $\mathfrak{T}(\mathrm{E}) = \sum\limits_{(v)} \otimes^{(v)}\mathrm{E}$;
- si ρ est une représentation linéaire de l'algèbre de Lie $\mathfrak{g}$ dans l'espace vectoriel E, les extensions tensorielles de Lie de ρ réalisent $\mathfrak{g}$ comme une algèbre de Lie de dérivations des diverses algèbres tensorielles sur E;
- puique le noyau d'une dérivation d'une algèbre unitaire est une sous-algèbre unitaire (vérification immédiate), les invariants de l'algèbre de Lie $\mathfrak{g}$ dans $\otimes\mathfrak{g}$, $\otimes\mathfrak{g}^*$, $\mathfrak{T}(\mathfrak{g}) = \sum\limits_{(v)} \otimes^{(v)}\mathfrak{g}$, $\mathrm{U}(\mathfrak{g})$ forment des sous-algèbres unitaires de ces diverses algèbres.

De même, les invariants d'une représentation linéaire ρ de $\mathfrak{g}$ dans E forment des sous-algèbres unitaires de $\otimes\mathrm{E}$, $\otimes\mathrm{E}^*$ et $\mathfrak{T}(\mathrm{E})$.

Proposition III.14.C

L'extension de Lie de $a \in \mathfrak{L}(\mathrm{E})$ à l'algèbre tensorielle complète $\mathfrak{T}(\mathrm{E}) = \sum\limits_{(v)} \otimes^{(v)}\mathrm{E}$ commute avec les contractions. Le contracté d'un invariant est donc invariant.

Preuve : C'est une conséquence de l'invariance du crochet de dualité et il suffit de le vérifier sur un exemple. Soit γ la contraction des second et troisième facteurs dans $\otimes^{(v)}\mathrm{E}$ avec $(v) = (..**)$.

C'est l'application de $\otimes^{(v)}\mathrm{E}$ dans $\otimes^{(.*)}\mathrm{E}$ définie par :

$$\gamma(x \otimes y \otimes \alpha \otimes \beta) = \langle y, \alpha\rangle x \otimes \beta$$

On a :

$$\begin{aligned}\widetilde{a}(x \otimes y \otimes \alpha \otimes \beta) &= ax \otimes y \otimes \alpha \otimes \beta + x \otimes ay \otimes \alpha \otimes \beta \\ &\quad - x \otimes y \otimes^t a\alpha \otimes \beta - x \otimes y \otimes \alpha \otimes^t a\beta\end{aligned}$$

et $\gamma\widetilde{a}(x \otimes y \otimes \alpha \otimes \beta) = \langle y, \alpha\rangle(ax \otimes \beta - x \otimes^t a\beta) = \widetilde{a}\gamma(x \otimes y \otimes \alpha \otimes \beta)$ puisque

$$\langle ay, \alpha\rangle = \langle y, {}^t a\alpha\rangle. \quad \text{c.q.f.d.}$$

Applications : soit $\mu = \sum u_i \otimes \alpha_i \otimes \beta_i$ l'élément de l'espace tensoriel $\otimes^{(.**)}\mathfrak{g}$ qui représente le crochet de l'algèbre de Lie $\mathfrak{g}$: $\mu(y,z) = [y,z] = \sum u_i\langle\alpha_i, y\rangle\langle\beta_i, z\rangle$.

On a vu ci-dessus que μ est un invariant de $\mathfrak{g}$ et ad $y = \sum\langle\alpha_i, y\rangle u_i \otimes \beta_i$.

Si $e = (e_1, e_2, \ldots, e_n)$ est une base de $\mathfrak{g}$, $\mu = \sum \mathrm{C}^k_{ij} e_k \otimes e^{*i} \otimes e^{*j}$ où les C^k_{ij} sont les constantes de structure associées à la base e, composantes du tenseur μ.

a) le contracté $\gamma_{13}\mu = \sum\langle u_i, \beta_i\rangle\alpha_i$ est la *forme linéaire invariante* sur $\mathfrak{g}$: $x \longrightarrow \mathrm{Tr}(\mathrm{ad}\ x)$ puisque $\gamma_{13}\mu(x) =$ Trace de l'opérateur $\sum\langle\alpha_i, x\rangle u_i \otimes \beta_i$. L'invariance traduit la propriété : $\mathrm{Tr}(\mathrm{ad}[a,x])$ $= 0$, $\forall a$.

b) Le double contracté $\gamma_{16}\gamma_{34}(\mu \otimes \mu)$ est la *forme bilinéaire symétrique de Killing sur* $\mathfrak{g}$. $(x,y) \longrightarrow \mathrm{Tr}(\mathrm{ad}\ x \circ \mathrm{ad}\ y)$. En effet : $\mu \otimes \mu = \sum u_i \otimes \alpha_i \otimes \beta_i \otimes u_j \otimes \beta_j \otimes \alpha_j$ et l'opérateur ad $x \circ$ ad y est l'opérateur

$$\Big(\sum\langle\alpha_i, x\rangle u_i \otimes \beta_i\Big) \circ \Big(\sum\langle\alpha_j, y\rangle u_j \otimes \beta_j\Big)$$
$$= \sum\langle\alpha_i, x\rangle\langle\alpha_j, y\rangle\langle\beta_i, u_j\rangle u_i \otimes \beta_j$$

dont la trace est $\sum\langle\alpha_i, x\rangle\langle\alpha_j, x\rangle\langle\beta_i, u_j\rangle\langle\beta_j, u_i\rangle$.

C'est bien la valeur pour (x,y) de la forme bilinéaire $\gamma_{16}\gamma_{34}$ $(\mu \otimes \mu) = \sum \alpha_i \otimes \alpha_j\langle\beta_j, u_i\rangle\langle\beta_i, u_j\rangle$. L'invariance s'exprime par : $\mathrm{Tr}([\mathrm{ad}\, a, \mathrm{ad}\, x]\mathrm{ad}\, y) + \mathrm{Tr}(\mathrm{ad}\, x[\mathrm{ad}\, a, \mathrm{ad}\, y]) = 0$, $\forall a$.

En utilisant une base e de $\mathfrak{g}$ et les constantes de structures associées, on obtient pour la forme de Killing : $\sum g_{ij} e^{*i} \otimes e^{*j}$ avec $g_{ij} = \sum \mathrm{C}^l_{ik}\mathrm{C}^k_{jl}$.

c) Plus généralement, la *forme p-linéaire sur* $\mathfrak{g}$: $\mathfrak{T}_p(x_1, x_2, \ldots, x_p)$ $= \mathrm{Tr}(\mathrm{ad}\ x_1 \mathrm{ad}\ x_2 \ldots \mathrm{ad}\ x_p)$ est l'invariant, élément de $\otimes^p \mathfrak{g}^*$, obtenu par la contraction multiple : $\gamma_{34}\gamma_{67} \cdots \gamma_{3k,3k+1}\gamma_{3p,1}(\otimes^p \mu)$, les $(p-1)$premières $\gamma_{3k,3k+1}$ représentant la composition des opérateurs, et la dernière $\gamma_{3p,1}$ représentant la trace. Cette forme s'écrit dans une base e de $\mathfrak{g}$:

$$\mathfrak{T}_p = \sum \mathrm{C}^{k_p}_{i_1 k_1}\mathrm{C}^{k_1}_{i_2 k_2} \ldots \mathrm{C}^{k_{p-1}}_{i_p k_p} e^{*i_1} \otimes e^{*i_1} \ldots \otimes e^{*i_p}$$

L'invariance de cette forme p-linéaire sur $\mathfrak{g}$ s'exprime par :

$$\sum \mathrm{Tr}(\mathrm{ad}\, x_1 \circ \ldots \mathrm{ad}\, x_{i-1} \circ [\mathrm{ad}\, a, \mathrm{ad}\, x_i]\mathrm{ad}\, x_{i+1} \circ \ldots \mathrm{ad}\, x_p)$$

$$\sum_i \mathrm{Tr}(\mathrm{ad}\, x_1 \circ \ldots \circ (ad\, a \;\mathrm{ad}\, x_i - \mathrm{ad}\, x_i \;\mathrm{ad}\, a) \circ \ldots \mathrm{ad}\, x_p) = 0 \;\forall a \in \mathfrak{g}.$$

Cette forme p-linéaire invariante n'a aucune raison d'être symétrique. On en déduit une forme p-linéaire symétrique en la symétrisant :

$$\mathrm{T}_p(x_1, \ldots, x_p) = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \mathrm{Tr}(\mathrm{ad}\, x_{\sigma 1} \circ \mathrm{ad}\, x_{\sigma 2} \circ \ldots \mathrm{ad}\, x_{\sigma p})$$

où $\mathfrak{S}_p$ est le groupe symétrique de permutations de $(1, 2, \ldots, p)$, et une fonction polynomiale invariante sur $\mathfrak{g}$ par :

$$\mathrm{P}_p(x) = \mathrm{Tr}(\mathrm{ad}\; x)^p$$

d) L'expression des invariants précédents à l'aide de la représentation adjointe conduit naturellement à une construction analogue à partir, cette fois, d'une représentation linéaire de dimension finie quelconque ρ de $\mathfrak{g}$ dans un espace vectoriel E.
La forme p-linéaire $\mathfrak{R}_p$ sur $\mathfrak{g}$ définie par :

$$\mathfrak{R}_p(x_1, x_2, \ldots x_p) = \mathrm{Tr}(\rho(x_1)\rho(x_2) \ldots \rho(x_p))$$

est invariante puisque, quel que soit $a \in \mathfrak{g}$:

$$\sum_i \mathfrak{R}_p(x_1, \ldots [a, x_i], \ldots x_p)$$
$$= \sum_i \mathrm{Tr}(\rho(x_1) \ldots \rho(x_{i-1})[\rho(a), \rho(x_i)] \ldots \rho(x_p)) = 0.$$

On en déduit une forme p-linéaire symétrique et une fonction polynomiale invariantes :

$$\mathrm{R}_p(x_1, x_2, \ldots, x_p) = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \mathrm{Tr}(\rho(x_{\sigma_2}) \ldots \rho(x_{\sigma_p})),$$
$$\mathrm{P}_p^{(\rho)}(x) = \mathrm{Tr}\, \rho(x)^p.$$

Remarquons que $\mathfrak{R}_p$ s'obtient par composition à travers l'algèbre enveloppante $\mathrm{U}(g)$ de $\mathfrak{g}$.

$$\otimes^p \mathfrak{g} \to \mathrm{U}(\mathfrak{g}) \xrightarrow{\overline{\rho}} \mathfrak{L}(\mathrm{E}) \xrightarrow{\mathrm{Tr}} \mathrm{K},$$

puisque $\overline{\rho}(x_1, x_2, \ldots, x_p) = \rho(x_1)\rho(x_2) \ldots \rho(x_p)$.

Théorème III.14.B : Soit $\mathfrak{g}$ une algèbre de Lie. La représentation $\widetilde{\mathrm{ad}}$ de $\mathfrak{g}$ comme algèbre de dérivations de l'algèbre $\otimes\mathfrak{g}$, extension tensorielle de sa représentation adjointe, (Définition III.14.) passe au quotient et détermine une représentation $\widetilde{\mathrm{ad}}$ de $\mathfrak{g}$ comme algèbre de dérivations de l'algèbre enveloppante $\mathrm{U}(\mathfrak{g}) = \otimes\mathfrak{g}/\mathrm{J}$. La projection Π de $\otimes\mathfrak{g}$ sur $\mathrm{U}(\mathfrak{g})$ est un épimorphisme de $\mathfrak{g}$-modules : $\Pi(\widetilde{\mathrm{ad}}\ a \cdot t) = \overline{\mathrm{ad}}\ a \cdot (\Pi t)$, $\forall a \in \mathfrak{g}$, $t \in \otimes\mathfrak{g}$. L'image par Π d'un invariant de $\otimes\mathfrak{g}$ est donc un invariant de $\mathrm{U}(\mathfrak{g})$. La sous-algèbre unitaire des invariants de $\mathrm{U}(\mathfrak{g})$ est le centre $\mathrm{Z}(\mathfrak{g})$ de $\mathrm{U}(\mathfrak{g})$. Si ρ est une représentation linéaire de $\mathfrak{g}$ dans E, les opérateurs de E appartenant à l'image du centre $\mathrm{Z}(\mathfrak{g})$ par $\overline{\rho}$ *commutent* avec tous les opérateurs de la représentation : ils appartiennent au commutant du $\mathfrak{g}$-module E.

Preuve : L'idéal bilatère J est engendré par les éléments $j(x,y) = x \otimes y - y \otimes x - [x,y]$. Quel que soit $a \in \mathfrak{g}$, $\widetilde{\mathrm{ad}}\, a.\ \mathrm{J} \subset \mathrm{J}$ puisque d'après l'identité de Jacobi :

$$\begin{aligned}\widetilde{\mathrm{ad}}\, a j(x,y) &= [a,x] \otimes y - y \otimes [a,x] - [[a,x],y] \\ &\quad + x \otimes [a,y] - [a,y] \otimes x - [x,[a,y]] \in \mathrm{J}.\end{aligned}$$

Si $t \in \mathfrak{g}$ et $u = \Pi t = t \bmod \mathrm{J}$, on définit l'action de a sur u par :

$$\overline{\mathrm{ad}}\, a \cdot u = \overline{\mathrm{ad}}\, a \cdot \Pi t = \overline{\mathrm{ad}}\, a(t \bmod \mathrm{J}) = (\widetilde{\mathrm{ad}}\, a \cdot t) \bmod \mathrm{J} = \Pi(\widetilde{\mathrm{ad}}\, a \cdot t).$$

Sur un élément décomposable $t = x_1 \otimes x_2 \otimes \ldots x_p$ de $\otimes\mathfrak{g}$:

$$\widetilde{\mathrm{ad}}\, a \cdot x_1 \otimes x_2 \otimes \ldots x_p = [a,x_1] \otimes x_2 \otimes \ldots x_p + x_1 \otimes [a,x_2] \otimes \ldots x_p + \ldots$$

Par souci de clarté, utilisons les mêmes lettres pour désigner les éléments de $\mathfrak{g} \subset \otimes\mathfrak{g}$ et leurs images par Π dans $\mathrm{U}(\mathfrak{g})$. L'image de $[x,y] \in \mathfrak{g}$ dans $\mathrm{U}(\mathfrak{g})$ est ainsi, puisque $j(x,y) \in \mathrm{J}$, égale à $xy - yx$.

On a donc dans $\mathrm{U}(\mathfrak{g})$, d'après de qui précède :

$$\begin{aligned}\overline{\mathrm{ad}}\, a \cdot x_1 x_2 \ldots x_p &= (a x_1 - x_1 a) x_2 \ldots x_p + x_1 (a x_2 - x_2 a) x_3 \ldots x_p \\ &\quad + x_1 x_2 \ldots x_{p-1}(a x_p - x_p a)\end{aligned}$$

qui se simplifie en :

$$\overline{\mathrm{ad}}\, a \cdot x_1 x_2 \ldots x_p = a x_1 x_2 \ldots x_p - x_1 x_2 \ldots x_p a = [a, x_1 x_2 \ldots x_p]$$

Quel que soit $u \in \mathrm{U}(\mathfrak{g}))$, on a donc $\overline{\mathrm{ad}}\, a \cdot u = au - ua = [a, u]$, et $\overline{\mathrm{ad}}\, a$ est une dérivation de $\mathrm{U}(\mathfrak{g})$. Puisque $\mathfrak{g}$ engendre $\otimes\mathfrak{g}$ et $\mathrm{U}(\mathfrak{g})$, $\overline{\mathrm{ad}}\, a \cdot u = 0$ quel que soit $a \in \mathfrak{g}$ est donc équivalent à la propriété : $uv = vu$, $\forall v \in \mathrm{U}(\mathfrak{g})$, c.q.f.d.

Un raisonnement analogue conduit à l'extension suivante qui illustre l'interdépendance des structures d'algèbres associatives et de Lie et dont la preuve est immédiate.

Théorème III.14.C : Soient $\mathfrak{g}$ une algèbre de Lie et A une algèbre associative unitaire sur le même corps. Un homomorphisme ρ d'algèbres de Lie de $\mathfrak{g}$ dans A_- détermine une structure de $\mathfrak{g}$-module sur A, un élément x de $\mathfrak{g}$ opérant sur A par : $\mathrm{ad}\,\rho(x)a = [\rho(x), a] = \rho(x)a - a\rho(x)$. Dès lors, ρ, ainsi que les homomorphismes d'algèbres unitaires $\widetilde{\rho} : \otimes\mathfrak{g} \to \mathrm{A}$ et $\overline{\rho} : \mathrm{U}(\mathfrak{g}) \to \mathrm{A}$ associés, sont des morphismes de $\mathfrak{g}$-modules (ils commutent avec, les opérations de $\mathfrak{g}$). Les invariants de $\mathfrak{g}$ dans $\mathfrak{g}$, $\otimes\mathfrak{g}$, $\mathrm{U}(\mathfrak{g})$ ont donc pour images des invariants de $\mathfrak{g}$ dans A.

Soit maintenant $\mathfrak{g}$ une algèbre de Lie dont la forme de Killing $\mathrm{B}(x, y) = \mathrm{Tr}(\mathrm{ad}\, x \circ \mathrm{ad}\, y)$ est non-dégénérée : $\mathfrak{g}$ est alors dite *semi-simple*, et l'application linéaire associée à B de $\mathfrak{g}$ dans $\mathfrak{g}^*$, définie par $\mathrm{B}(x, y) = \langle x, \varphi(y)\rangle = \langle \varphi(x), y\rangle$ est bijective. Puisque B est invariante, c'est un isomorphisme de $\mathfrak{g}$-modules de $\mathfrak{g}$ sur $\mathfrak{g}^*$:

$$\langle x, \varphi([a, y])\rangle = \mathrm{B}(x, \mathrm{ad}\, a \cdot y) = \mathrm{B}(-\mathrm{ad}\, a \cdot x, y) = \langle x, -{}^t(\mathrm{ad}\, a)\varphi(y)\rangle$$

L'inverse $\overset{-1}{\varphi}$ est un isomorphisme de $\mathfrak{g}$-modules de $\mathfrak{g}^*$ sur $\mathfrak{g}$. C'est aussi un élément Γ de $\mathfrak{g} \otimes \mathfrak{g}$ qui est donc invariant « inverse » de B.

L'invariant Γ, ou son image C dans l'algèbre enveloppante $\mathrm{U}(\mathfrak{g})$, également invariant, sont indifféremment appelés *élément de Casimir de* $\mathfrak{g}$.

L'isomorphisme entre $\mathfrak{g}^*$ et $\mathfrak{g}$ que définit B permet de transformer des invariants covariants, appartenant aux puissances tensorielles de $\mathfrak{g}^*$, en invariants contravariants, éléments des $\otimes^p\mathfrak{g}$. (ainsi : $\Gamma = (\overset{-1}{\varphi} \otimes \overset{-1}{\varphi}) \cdot \mathrm{B})$ et de prendre leurs images dans $\mathrm{U}(\mathfrak{g})$.

On définit les *éléments de Casimir d'ordre supérieur* comme les images dans $\otimes\mathfrak{g}$ des formes p-linéaires *symétriques* invariantes T_p définies en c) :

$\Gamma_p = (\otimes^p \overset{-1}{\varphi}) \cdot \mathrm{T}_p$ ou dans $\mathrm{U}(g)$: $\mathrm{C}_p =$ image de Γ_p par l'application $\otimes\mathfrak{g} \to \mathrm{U}(\mathfrak{g})$.

Si e est une base de $\mathfrak{g}$, e^* sa base duale, soit $|g_{ij}|$ la matrice de B, qui est aussi celle de φ, et $|g^{ij}|$ son inverse, qui est celle de $\overset{-1}{\varphi}$. On a :
$\mathrm{B} = \sum g_{ij} e^{*i} \otimes e^{*j}$ et $\Gamma = \sum g^{ij} e_i \otimes e_j = \sum\limits_i (e_i \otimes \sum\limits_j g^{ij} e_j)$ $= \sum e_i \otimes e'^i$ et $\mathrm{C} = \sum e_i e'^i = \sum g_{ij} e'^i e'^j = \sum \mathrm{C}_{ik}^m \mathrm{C}_{jm}^k e'^i e'^j$, où $e' = \overset{-1}{\varphi} e^*$ est la base B-supplémentaire de e : $\mathrm{B}(e_i, e'^j) = \delta_i^j$. L'élément de Casimir C_p s'écrit dans $\mathrm{U}(\mathfrak{g})$:

$$\mathrm{C}_p = \sum \mathrm{C}_{i_1 k_p}^{k_p} \mathrm{C}_{i_2 k_2}^{k_1} \ldots \mathrm{C}_{i_p k_p}^{k_{p-1}} e'^{i_1} e'^{i_2} \ldots e'^{i_p}.$$

Nous terminerons ce paragraphe par une notion particulière d'invariance qui généralise l'invariance des formes bilinéaires sur les algèbres de Lie, et à ses applications :

Définition III.14.F :

Soit A une algèbre *quelconque* de dimension finie. Une forme bilinéaire symétrique b sur A est dite *invariante*, ou encore « *associative* » si quels que soient $x, y, z \in \mathrm{A}$, on a : $b(xy, z) = b(x, yz)$.

Lorsque A est une algèbre de Lie, la condition d'invariance : $b([y, x], z) + b(x, [y, z]) = 0$ exprime que b est un invariant de A.

Exemples : les formes trace : Si u, v sont des opérateurs linéaires d'un espace vectoriel de dimension finie E, la forme bilinéaire symétrique $b(u, v) = \mathrm{Tr}(u \circ v)$ est une forme invariante sur chacune des trois algèbres : associative $\mathfrak{L}(\mathrm{E})$, de Lie $\mathfrak{L}(\mathrm{E})_-$, de Jordan $\mathfrak{L}(\mathrm{E})_+$. En effet :

$$b(ux, v) = \mathrm{Tr}((ux)v) = \mathrm{Tr}(u(xv)) = b(u, xv) \quad \text{(associative)}$$

$$b([u, x], v) = \mathrm{Tr}(uxv) - \mathrm{Tr}(xuv) = \mathrm{Tr}(uxv) - \mathrm{Tr}(uvx) = b(u, [x, v]) \quad \text{(de Lie)}$$

$$b(u \circ x, v) = \frac{1}{2}\,\mathrm{Tr}(uxv) + \frac{1}{2}\,\mathrm{Tr}(xuv) = b(u, x \circ v) \quad \text{(de Jordan)}$$

Si ρ est un homomorphisme de l'algèbre A associative, de Lie, de Jordan, dans $\mathfrak{L}(\mathrm{E})$, resp. $\mathfrak{L}(\mathrm{E})_-$, resp. $\mathfrak{L}(\mathrm{E})_+$, la forme $b_\rho(x, y) = \mathrm{Tr}\, \rho(x)\rho(y)$ est invariante sur A : c'est la *forme trace* de la représentation.

Remarque :

1) Lorsque b est *non-dégénérée*, l'invariance exprime que les translations à gauche et à droite de A sont *adjointes* : $L_x^* = R_x$. Les $(L_x - R_x)$ sont donc antisymétriques.

Si b est une forme invariante, la relation d'orthogonalité qui lui est associée est « compatible » avec la structure d'algèbre. On a en effet les propriétés suivantes :

Propriétés générales :

1) le sous-espace $\mathfrak{I}^\perp$ orthogonal d'un idéal $\mathfrak{I}$ de A est également un idéal.

$$\text{si} \quad x \in \mathfrak{I}^\perp,\ y \in \mathfrak{I} \text{ et } a \in A \text{ on a puisque } ay \text{ et } ya \in \mathfrak{I} :$$
$$b(x \cdot a, y) = b(x, a \cdot y) = 0 \text{ et } b(y, a \cdot x) = b(y \cdot a, x) = 0$$
$$\text{d'où} \quad x \cdot a \quad \text{et} \quad a \cdot x \in \mathfrak{I}^\perp$$

$\mathfrak{I} \cap \mathfrak{I}^\perp$ est totalement isotrope : b y est identiquement nulle

2) en particulier le radical $A^\perp$ de $(a; b)$ est un idéal de A.

Propriétés particulières lorsque b est non-dégénérée :
$J = \mathfrak{I} \cap \mathfrak{I}^\perp$ est un idéal de carré nul. En effet, si x et $y \in J$, quel que soit $a \in A$ on a : $b(x \cdot y, a) = b(x, y \cdot a) = 0$ puisque x et y sont orthogonaux, d'où $xy = 0$.

Théorème III.14.D. de la semi-simplicité. Soit A une algèbre de dimension finie possédant les *deux* propriétés :

a) il existe sur A une forme bilinéaire symétrique invariante non-dégénérée.

b) A ne possède aucun idéal non nul de carré nul.

A est alors somme directe d'idéaux bilatères simples : A est dite *semi-simple*.

Preuve : Soit $\mathfrak{I}$ un idéal (bilatère) non nul *minimal*. Puisque (Propriété ci-dessus) $\mathfrak{I} \cap \mathfrak{I}^\perp$ est un idéal de carré nul, $\mathfrak{I} \cap \mathfrak{I}^\perp = 0$ et $A = \mathfrak{I} \oplus \mathfrak{I}^\perp$.

Tout idéal $\mathfrak{I}$ étant un idéal de A, et $\mathfrak{I}^2$ étant non nul, donc égal à $\mathfrak{I}$, $\mathfrak{I}$ est une algèbre simple.

La restriction de b à $\mathfrak{I}^\perp$ est une forme bilinéaire symétrique invariante non-dégénérée, et tout idéal de $\mathfrak{I}^\perp$ étant un idéal de A est de carré non nul. $\mathfrak{I}^{\perp}$ vérifie les hypothèses initiales faites sur A et on achève la démonstration par une récurrence sur la dimension.

De par leur structure, les algèbres usuelles, associatives, de Lie, ou de Jordan, sont munies de façon naturelle de formes bilinéaires invariantes que l'on appelle fondamentales et pour lesquelles on utilise la notion du produit scalaire : $(|)$.

Définition III.4.G. Formes fondamentales :

On les définit pour chacune des catégories d'algèbres suivantes (de dimensions finies) comme des formes traces liées aux translations à gauche de l'algèbre :

- algèbres *associatives* : la forme fondamentale est la forme trace de sa représentation par les translations à gauche :
  $(x \mid y) = \text{Trace } (\mathrm{L}_x\mathrm{L}_y) = \text{Trace } (\mathrm{L}_{xy})$. Elle est symétrique puisque $\mathrm{Tr}([\mathrm{L}_x, \mathrm{L}_y]) = 0$ et invariante au sens de la Définition III.14 ci-dessus.

Pour une *algèbre de Jordan*, la forme fondamentale :

$$\boxed{(x \mid y) = \text{Trace } \mathrm{L}_{xy}},$$

est la trace de l'opérateur linéaire de translation par (xy). L'identité classique des algèbres de Jordan, dont la démonstration est laissée à titre d'exercice :

$$\mathrm{L}_{x(yz)} - \mathrm{L}_{(xy)z} = [[\mathrm{L}_x, \mathrm{L}_z], \mathrm{L}_y],$$

montre qu'elle est « invariante » (N.B. la forme bilinéaire $\mathrm{Tr}(\mathrm{L}_x\mathrm{L}_y)$ est symétrique mais n'est pas invariante car $x \longrightarrow \mathrm{L}_x$ *n'est pas* une représentation).

Pour une *algèbre de Lie* la forme fondamentale, dite *forme de Killing*, est la forme trace de la représentation adjointe :

$$\boxed{(x \mid y) = \mathrm{Tr}(\text{ad } x \cdot \text{ad } y)}.$$

Proposition III.14.D :

Les formes fondamentales ont un caractère héréditaire : la forme relative à un idéal $\mathfrak{I}$ de l'algèbre A coïncide avec la restriction à $\mathfrak{I}$ de la forme fondamentale de A.

Preuve : En effet, il suffit de prendre une base de A : $e_1, e_2, \ldots,$ e_n dont les p premiers vecteurs forment une base de $\mathfrak{I}$. Relativement à une telle base, les matrices des translations à gauche L_x pour $x \in \mathfrak{I}$, ont leurs termes $(\mathrm{L}_x)_{p+i}^{p+j} = (\mathrm{L}_x e_{p+i})^{p+j} = 0$, pour $i, j > 0$ et les traces de L_{xy} ou de $\mathrm{L}_x\mathrm{L}_y$ pour $x, y \in \mathfrak{I}$, dans A, ou de leurs restrictions à $\mathfrak{I}$, sont les mêmes c.q.f.d.

Proposition III.14.E :

Si l'algèbre A est somme directe d'idéaux : $\mathrm{A} = \mathfrak{I}_1 \oplus \mathfrak{I}_2 \oplus \ldots \mathfrak{I}_p$, les idéaux sont orthogonaux pour la forme fondamentale, qui est donc somme des formes fondamentale des $\mathfrak{I}_k$.

Preuve : En effet, si $x \in \mathfrak{I}_j$ et $y \in \mathfrak{I}_k$ avec $j \neq k$, $x \cdot y = 0$ et les opérateurs L_{xy} ou ad $x \cdot$ ad y dans le cas de Lie, sont des opérateurs nuls d'où $(x \mid y) = 0$.

Théorème de la forme fondamentale III.14.E : Soit A une algèbre de dimension finie associative, de Lie, ou de Jordan dont la forme fondamentale est régulière. Alors :

1) A ne contient aucun idéal de carré nul et est *donc semi-simple* d'après le théorème III.14. D.

2) L'application linéaire : $x \in \mathrm{A} \to \mathrm{L}_x \in \mathfrak{L}(\mathrm{A})$ (où $\mathrm{L}_x u = x \cdot u$) est injective ($x \to \mathrm{L}_x$ est une représentation linéaire de A si A est associative ou de Lie, ce *n'est pas* une représentation linéaire si A est de Jordan).

3) Si A est associative ou de Jordan, elle possède une unité (algorithme de Casimir).

Preuve :

1) Soient $\mathfrak{I}$ un idéal de carré nul de A, $a \in \mathfrak{I}$ et x, y deux éléments quelconques de A. Nous allons prouver que $\mathfrak{I}$ est nécessairement nul.

a) si A est une algèbre de Lie, on a $[\mathfrak{I}, \mathfrak{I}] = 0$ d'où :

$$y \in \mathrm{A} \xrightarrow{\text{ad } a} [a, y] \in \mathfrak{I} \xrightarrow{\text{ad } x} [x, [a, y]] \in \mathfrak{I} \xrightarrow{\text{ad } a} 0.$$

L'opérateur ad$x \circ$ ad a est de carré nul, donc de trace nulle, d'où il résulte que $(x \mid a) = \mathrm{Tr}(\text{ad } x \circ \text{ad } a) = 0$ quel que soit x, et puisque $(|)$ est régulière, $a = 0$: $\mathfrak{I} \equiv 0$.

b) si A est associative ou de Jordan, $\mathfrak{I}\mathfrak{I} = 0$ et $ax \in \mathfrak{I}$, d'où :

$$y \in \mathrm{A} \xrightarrow{\mathrm{L}_{ax}} (ax)y \in \mathfrak{I} \xrightarrow{\mathrm{L}_{ax}} 0.$$

L'opérateur L_{ax} est de carré nul, donc de trace nulle, et $(a \mid x) = \mathrm{Tr}\, \mathrm{L}_{ax} = 0$ quel que soit x, d'où $a = 0$ et $\mathfrak{I} \equiv 0$.

2)

a) Si A est de Lie, le noyau de la représentation adjointe est le centre $\mathfrak{Z}$ de A : $\mathfrak{Z} = \{x \in \mathrm{A}; [x, y] = 0\ \forall \mathrm{A}\}$.
Le centre $\mathfrak{Z}$ est un idéal de carré nul, donc nul.

b) Si A est associative ou de Jordan $\mathrm{L}_x = 0$ signifie que $xu = 0$ quel que soit $u \in \mathrm{A}$ et dans ce cas : $(x \mid u) = \mathrm{Tr}\, \mathrm{L}_{xu} = 0$ d'où $x = 0$, c.q.f.d.

3) Choisissons dans A deux bases $e = (e_1, e_2, \ldots, e_n)$ et $e' = (e'_1, e'_2, \ldots, e'_n)$ complémentaires relativement à la forme fondamentale c'est-à-dire telles que $(e_i \mid e'_j) = \delta_{ij}$. (la base e' est l'image réciproque de la base e^*, duale de e dans l'espace A^*, par l'isomorphisme de A sur A^* associé à $(|)$).

Si a est un opérateur linéaire de l'espace vectoriel A, on a :

$$\sum_i (ae_i \mid e'_i) = \sum_i \Big(\sum_j e_j a_i^j \mid e'_i \Big) = \sum_i a_i^i = \mathrm{Tr}\, a.$$

Si A est associative ou de Jordan, soit $e = \sum e_i e'_i \in \mathrm{A}$. Quels que soient $x, y \in \mathrm{A}$, on a :

$$(x \mid ye) = \sum (xy \mid e_i e'_i) = \sum ((xy)e_i \mid e'_i) = \mathrm{Tr}(\mathrm{L}_{xy}) = (x \mid y),$$

ce qui prouve que $ye = y$ quel que soit $y \in \mathrm{A}$. Mais $(x \mid ey) = (xe \mid y) = (x \mid y)$ et on a aussi $ey = y$ quel que soit $y \in \mathrm{A}$, c.q.f.d.

EXERCICES SUR LE CHAPITRE III

EX. III.1 :

Soit E un espace vectoriel de dimension n. On considère pour chaque entier $p : 1 \leqslant p \leqslant n$ l'ensemble de composantes :

$$\delta_{i_1 \ldots i_p, j_1 \ldots j_p} = \text{dét}\, |\delta_{i_k, j_h}| \quad \text{les } i_i, j_r \text{ variant de 1 à } n.$$

(exemple : $\delta_{i_1,j_1 i_2,j_2} = \delta_{i_1,j_1}\delta_{i_2,j_2} - \delta_{i_1,j_2}\delta_{i_2,j_1}$) :

1) Montrer que ce sont les composantes d'un tenseur et calculer les contractés sur les p indices de $\delta_{i_1\dots i_p,j_1\dots j_p}$ avec les tenseurs $t^{j_1\dots j_p}$ et $t_{i_1\dots i_p}$.
2) Si (a^i_j) est la matrice des composantes d'un opérateur linéaire de E, calculer $\sum a^{i_1}_{j_1}\dots a^{i_n}_{j_n}\delta_{j_1\dots j_n,i_1\dots i_n}$ et en déduire les formules :
si $n=2$: dét $a = \dfrac{1}{2}\,[(\mathrm{Tr}\,a)^2 - \mathrm{Tr}\,a^2]$,
si $n=3$: dét $a = \dfrac{1}{3!}\,[(\mathrm{Tr}\,a)^3 - 3\mathrm{Tr}\,a\cdot\mathrm{Tr}\,a^2 + 2\mathrm{Tr}\,a^3]$.
3) Montrer que $\delta_{i_1\dots i_n,j_1\dots j_n}$ est le tenseur de Kronecker (II.6).
4) Calculer et interpréter les contractés :

$$\sum \delta^{i_1\dots i_{n-1}k}_{j_1\dots j_{n-1}k},\quad \sum \delta^{i_1\dots i_p k_1\dots k_{n-p}}_{j_1\dots j_p k_1\dots k_{n-p}}.$$

EX. III.2 :

Soient $a_{i,j}$ les composantes d'un tenseur une fois contravariant et une fois covariant sur espace vectoriel E de dimension finie $n > 1$.

Les cofacteurs du déterminant dét$|a^i_j|$ sont-ils les composantes d'un tenseur ? Même question pour un tenseur a_{ij} ou a^{ij} deux fois covariant ou deux fois contravariant.

EX. III.3 :

Soit K(n) l'espace vectoriel des matrices $n \times n$ sur le corps commutatif K. Montrer que toute forme linéaire f sur K(n) peut s'écrire d'une manière et d'une seule : si $X \in K(n)$, $f(X) = \mathrm{Trace}\,FX$ où $F \in K(n)$. Montrer que si f possède la propriété : $f(XY) = f(YX)$ quels que soient X et $Y \in K(n)$, alors $F = \lambda I$ et $f(X) = \lambda\ \mathrm{Trace}\,X$.

EX. III.4 :

Soit A une algèbre de dimensions *finie* sur le corps K. Montrer que s'il existe un élément a de A tel que $ax = 0$ ou $xa = 0$ impliquent $x = 0$, A possède une unité et a est inversible.

EX. III.5 :

Soient A et B deux algèbres unitaires sur le même corps K, C le centre de A, D celui de B. Montrer que le commutant de la sous-algèbre $A \otimes 1_B$

dans $A \otimes B$ est une sous-algèbre qui s'identifie à $C \otimes B$, et que le centre de $A \otimes B$ s'identifie à $C \otimes D$.

EX. III.6 :

Soient E un espace vectoriel de dimension finie n, T son algèbre tensorielle (§III.12).

a) Montrer que si a et b sont des éléments de T tels que $ab = 0$, l'un au moins de ces éléments est nul.

b) Montrer que si a, b, c, d sont des éléments de T tels que $ac = bd$, on a nécessairement $a = bh$ ou $b = ak$.

c) En déduire que si le degré de a est > 0, les seuls éléments de T qui commutent avec a sont les puissances de a.

d) Montrer si $n > 1$ le centre de T se réduit aux scalaires.

EX. III.7 :

Soit $\mathfrak{g}$ une algèbre de Lie de dimension finie sur le corps K. Montrer que :

a) si a est une forme bilinéaire invariante sur $\mathfrak{g}$, la forme trilinéaire $t(x, y, z) = b([x, y], z)$ est antisymétrique invariante sur $\mathfrak{g}$.

b) si la forme de Killing de $\mathfrak{g}$ est non-dégénérée toute dérivation D de $\mathfrak{g}$ est intérieure : il existe $a \in \mathfrak{g}$ tel que $Dx = \operatorname{ad} a \cdot x = [a, x]$ quel que soit x.

CHAPITRE IV

Algèbre extérieure et Algèbre symétrique d'un espace vectoriel de dimension finie ou infinie

Dans le chapitre III précédent l'espace vectoriel de base E était supposé de dimension finie ce qui permettait à son dual E^* d'intervenir systématiquement sur un pied d'égalité avec E. Dans ce chapitre IV, les constructions sont effectuées sur le seul espace E, sans intervention de E^*, ce qui permet de n'imposer aucune restriction à la dimension de E qui peut être infinie. $\otimes^p E$ désigne le produit tensoriel de p copies de E suivant la définition du §II.7.

En analyse, les dérivées secondes et d'ordre supérieur des fonctions différentiables de plusieurs variables sont symétriques.

L'antisymétrisation les élimine et permet de ne conserver que les dérivées du premier ordre, pour lesquelles les changements de variables sont linéaires : c'est un instrument de linéarisation.

IV.1 – Puissances extérieures et symétriques d'un espace vectoriel

Remarque préliminaire : Nous verrons que la puissance extérieure p-ème, $\wedge^p$E, de l'espace vectoriel E est naturellement isomorphe au sous-espace des tenseurs antisymétriques de $\otimes^p$E, tandis que la puissance symétrique p-ème, $\vee^p$E est naturellement isomorphe au sous-espace des tenseurs symétriques de $\otimes^p$E. On pourrait donc penser qu'il suffit de s'en tenir à ces sous-espaces, de définition bien claire. Or, partout où ils interviennent, $\wedge^p$E et $\vee^p$E apparaissent bien plutôt comme des quotients de $\otimes^p$E. La structure d'espace quotient est moins simple que celle de sous-espace, mais cette complication au départ est largement compensée par le caractère naturel de toutes les constructions ultérieures.

Soient F un sous-espace vectoriel de $\otimes^p$E, H l'espace quotient, $\bar{t}$ l'image dans H de $t \in \otimes^p$E. Si l'on veut que le groupe symétrique $\mathfrak{S}_p$ opère de façon triviale sur H : $\forall \bar{t} \in$ F, $\sigma\bar{t} = \bar{t}$, il faut et il suffit que F contienne tous les éléments de la forme $t - \sigma t$. De même, si l'on veut que $\mathfrak{S}_p$ opère sur H par la signature : $\sigma\bar{t} = \varepsilon(\sigma)\bar{t}$, il faut et il suffit que F contienne tous les éléments de la forme $t - \varepsilon(\sigma)\sigma t$.

Définition IV.I.A.

1) Le sous-espace vectoriel L_p de $\otimes^p$E engendré par les éléments décomposables ayant deux facteurs égaux, est aussi engendré par les éléments de la forme $\{t - \varepsilon(\sigma)\sigma t;\ \sigma \in \mathfrak{S}_p\}$ et est appelé le sous-espace d'alternation ou d'antisymétrisation de $\otimes^p$E. L'espace quotient $\otimes^p\mathrm{E}/\mathrm{L}_p = \wedge^p$E est la *puissance extérieure p-ième de* E. L'image de $x_1 \otimes x_2 \otimes \ldots x_p$ est notée $x_1 \wedge x_2 \wedge \ldots \wedge x_p$ et est appelée produit extérieur de $x_1, x_2, \ldots x_p$.

2) Le sous-espace vectoriel M_p de $\otimes\mathrm{E}$ engendré par les éléments de la forme $\{t - \sigma t;\ \sigma \in \mathfrak{S}_p\}$ est appelé le sous-espace de symétrisation de $\otimes^p\mathrm{E}$. L'espace quotient $\otimes^p\mathrm{E}/\mathrm{M}_p = \vee^p\mathrm{E}$ est la *puissance symétrique p-ème de* E. L'image de $x_1 \otimes x_2 \otimes \ldots x_p$ est notée $x_1 \vee x_2 \ldots \vee x_p$ et appelée produit symétrique de $x_1, x_2 \ldots x_p$. Les éléments de $\vee^p\mathrm{E}$ sont aussi appelés *polynômes de degré p* sur E, tandis que les éléments de $\vee^p\mathrm{E}^*$ sont appelés *fonctions polynomiales de degré p* sur E.

Commentaire sur ces définitions :

1) Si t est décomposable avec deux facteurs égaux, la transposition τ de ces deux facteurs laisse t invariant, et $t = \dfrac{1}{2}(t + \tau t)$. Réciproquement, si $\sigma = \tau_1\tau_2 \ldots \tau_m$ est le produit de m transpositions, $\varepsilon(\sigma) = (-1)^m$, et $t - (-1)^m\sigma t = (t + t\tau_1) - (t\tau_1 + t\tau_1\tau_2) \ldots - (-1)^m(t\tau_1 \ldots \tau_{m-1} + t\tau_1 \ldots \tau_m)t$ étant somme d'éléments décomposables, si τ permute le i-ème et le j-ème facteur $x_1 \otimes \ldots \otimes x_p + \tau(x_1 \otimes \ldots \otimes x_p) = x_{1\otimes} \ldots (x_i + x_j) \otimes \ldots \otimes (x_i + x_j) \otimes \ldots x_p - x_1 \otimes \cdots x_i \otimes \ldots x_i \otimes \ldots x_p - x_{1\otimes} \ldots x_j \otimes \ldots x_j \otimes \ldots x_p$.

Il en résulte que tout élément de la forme $t - \varepsilon(\sigma)\sigma t$ est somme d'éléments décomposables ayant deux facteurs égaux.

2) Si σ est paire $t - \varepsilon(\sigma)\sigma t = t - \sigma t$ appartient à la fois à L_p et M_p.

3) L_p et M_p sont stables par les opérations du groupe linéaire $\mathrm{G}l(\mathrm{E})$. Ces opérations passent donc au quotient ce qui représente linéairement $\mathrm{G}l(\mathrm{E})$ dans $\vee^p\mathrm{E}$ et $\wedge^p\mathrm{E}$. Ainsi $g \cdot (x_1 \wedge x_2 \wedge \ldots x_p) = gx_1 \wedge gx_2 \wedge \ldots gx_p$ et $g \cdot (x_1 \vee x_2 \vee \ldots x_p) = gx_1 \vee gx_2 \vee \ldots gx_p$.

Proposition IV.1.A.

Dans $\otimes^p\mathrm{E}$:

a) les sous-espaces L_p, d'antisymétrisation, et $\mathrm{A}^p\mathrm{E} = a(\otimes^p\mathrm{E})$, des tenseurs antisymétriques, sont supplémentaires. L_p est le noyau de a.

b) les sous-espaces M_p, de symétrisation, et $\mathrm{S}^p\mathrm{E} = s(\otimes^p\mathrm{E})$, des tenseurs symétriques, sont supplémentaires. M_p est le noyau de s.

$$\otimes^p\mathrm{E} = \mathrm{Ker}\, a \oplus \mathrm{Im}\, a = \mathrm{L}_p \oplus \mathrm{A}^p\mathrm{E} \text{ et } \otimes^p\mathrm{E} = \mathrm{Ker}\, s \oplus \mathrm{Im}\, s = \mathrm{M}_p \oplus \mathrm{S}^p\mathrm{E}.$$

Preuve :

1) Si $u \in \mathrm{M}_p$, soit $u = \Sigma\lambda_j(t_j - \sigma_j t_j)$ on a $su = \Sigma\lambda_j(st_j - st_j) = 0$, d'où $\mathrm{M}_p \subset \operatorname{Ker} s$. Si $u \in \mathrm{M}_p \cap \mathrm{S}^p\mathrm{E}$, $u = su = 0$.

D'autre part, quel que soit $t \in \otimes^p\mathrm{E}$, on a $t = st + (t - st)$ et $t - st = \dfrac{1}{p!}(p!t - \Sigma\sigma t) = \dfrac{1}{P!}\Sigma(t - \sigma t)$, ce qui montre que $t - st \in \mathrm{M}_p$ et prouve b).

2) de la même façon, si $u \in \mathrm{L}_p \cap \mathrm{A}^p\mathrm{E}$, $au = u$ et $u = \Sigma\lambda_j(t_j - \varepsilon(\sigma_j)\sigma_j)t_j$, ce qui donne $au = 0$.

Puisque $t = at + (t - at)$, on a bien *a)*.

Proposition IV.1.B.

L'application linéaire φ de $\otimes^p\mathrm{E}$ sur son quotient $\wedge^p\mathrm{E} = \otimes^p\mathrm{E}/\mathrm{L}_p$ détermine un isomorphisme naturel Φ de $\mathrm{A}^p\mathrm{E}$, supplémentaire de L_p, sur $\wedge^p\mathrm{E}$, qui envoie l'antisymétrisé $a(x_1 \otimes x_2 \otimes \ldots x_p)$
$= \dfrac{1}{p!}\Sigma\varepsilon(\sigma)x_{\sigma 1} \otimes x_{\sigma 2} \ldots \otimes x_{\sigma p}$ sur le produit extérieur $x_1 \wedge x_2 \wedge \ldots x_p$.

De la même façon, l'application linéaire ψ de $\otimes^p\mathrm{E}$ sur son quotient $\vee^p\mathrm{E} = \otimes^p\mathrm{E}/\mathrm{M}_p$ détermine un isomorphisme naturel ψ de $\mathrm{S}^p\mathrm{E}$, supplémentaire de M_p, sur $\vee^p\mathrm{E}$, qui envoie le symétrisé $s(x_1 \otimes x_2 \otimes \ldots x_p) = \dfrac{1}{p!}\Sigma x_{\sigma 1} \otimes x_{\sigma 2} \otimes \ldots x_{\sigma p}$ sur le produit symétrique $x_1 \vee x_2 \vee \ldots x_p$.

Ces isomorphismes linéaires sont aussi des isomorphismes de $\mathrm{G}l(\mathrm{E})$ – modules.

Preuve :

$$\varphi\left(\frac{1}{p!}\Sigma\varepsilon(\sigma)x_{\sigma 1} \otimes \ldots x_{\sigma p}\right)$$
$$= \frac{1}{p!}\sum_{\sigma \in \mathfrak{S}_p}^{p} \varepsilon(\sigma) \ldots x_{\sigma 1} \wedge x_{\sigma 2} \wedge \ldots x_{\sigma p} = x_1 \wedge x_2 \wedge \ldots x_p$$

et de même $\Psi\left(\dfrac{1}{p!}\Sigma x_{\sigma 1} \otimes \ldots x_{\sigma p}\right) = \dfrac{1}{p!}\displaystyle\sum_{\sigma \in \mathfrak{S}_p}^{p} x_{\sigma 1} \vee x_{\sigma 2} \vee \ldots x_{\sigma p} x$
$= x_1 \vee x_2 \vee \ldots x_p$.

Nous allons maintenant obtenir une caractérisation de $\wedge^p\mathrm{E}$ et $\vee^p\mathrm{E}$ comme solutions de problèmes universels.

Soient E un espace vectoriel donné, V un espace vectoriel quelconque, tous deux sur le même corps K. Considérons les applications p-linéaires φ.

$$\underset{p}{\mathrm{E} \times \mathrm{E} \times \ldots \times \mathrm{E}} \xrightarrow{\varphi} \mathrm{V}$$

satisfaisant à l'une ou à l'autre des propriétés suivantes :

a) φ est *symétrique* : $\forall \sigma \in \mathfrak{S}_p$, $\sigma\varphi = \varphi$.

b) φ est *antisymétrique* : $\forall \sigma \in \mathfrak{S}_p$, $\sigma\varphi = \varepsilon(\sigma)\varphi$.

c) φ est *alternée* : φ est nulle sur tout p-uple d'éléments de E présentant deux termes égaux.

Les conditions b) et c) sont équivalentes. En effet, si φ est alternée, $\varphi(x_1, \ldots x_j + x_k, x_{j+1}, \ldots, x_k + x_j, x_{k+1}, \ldots x_p) = 0$ entraîne $\tau\varphi = -\varphi$ pour toute transposition de deux éléments soit, puisque toute permutation est un produit de transpositions : $\sigma\varphi = \varepsilon(\sigma)\varphi$ et φ est antisymétrique. Réciproquement supposons φ antisymétrique. Un p-uple $(x_1, x_2, \ldots x_p)$ qui présente deux termes égaux reste inchangé par la transposition τ qui les échange et $\varphi(x_1, x_2, \ldots x_p) = \varphi(x_{\tau 1}, x_{\tau 2}, \ldots x_{\tau p}) = -\varphi(x_1, x_2, \ldots x_p) = 0$, et φ est alternée.

Ces applications p-linéaires satisfaisant à *a)* ou à *b)c)* posent un problème universel : trouver une application p-linéaire u dans un espace vectoriel fixe H tels que pour toute application φ dans V, il existe une application linéaire unique Φ de H dans V telle que $\varphi = \Phi \circ u$:

$$\begin{tikzcd}
\mathrm{E} \times \mathrm{E} \times \ldots \times \mathrm{E} \arrow[rr, "\varphi"] \arrow[dr, "u"'] & & \mathrm{V} \\
& \mathrm{H} \arrow[ur, "\phi"'] &
\end{tikzcd}$$

ϕ linéaire unique avec $\varphi = \phi \circ u$

Or, puisque φ est multilinéaire, elle détermine une application linéaire $\widetilde{\varphi} : \otimes^p\mathrm{E} \longrightarrow \mathrm{V}$ telle que sur les éléments décomposables de $\otimes^p\mathrm{E}$: $\widetilde{\varphi}(x_1 \otimes x_2 \otimes \ldots x_p) = \varphi(x_1, x_2, \ldots x_p)$. Dès lors, φ est symétrique, resp. antisymétrique, si et seulement s'il en est de même de $\widetilde{\varphi}$, c'est-à-dire si et seulement si $\widetilde{\varphi}$ est nulle sur M_p, resp.

L_p et $\widetilde{\varphi}$ se factorise alors par l'application de $\otimes^p E$ sur son quotient par M_p ou L_p d'où le

Théorème IV.1.A. L'application naturelle u obtenue par la composition : $E \times E \times \ldots E \to \otimes^p E \to \otimes^p E / L_p = \wedge^p E$ est p-linéaire antisymétrique universelle : elle factorise toute application p-linéaire antisymétrique (ou alternée) de E dans V :

$$\begin{tikzcd} E \times E \times \ldots E \arrow[rr, "\varphi"] \arrow[dr, "u"'] & & V \\ & \Lambda PE \arrow[ur, "\phi"'] & \end{tikzcd} \qquad \varphi = \phi \circ u$$

par une application linéaire unique Φ de $\wedge^p E$ dans V telle que $\Phi(x_1 \wedge x_2 \wedge \ldots x_p) = \varphi(x_1, x_2, \ldots x_p)$. On a ainsi une bijection canonique : $\mathrm{Alt}_p(E; V) \equiv \mathfrak{L}(\wedge^p E;\ V)$.

De la même façon, l'application p-linéaire symétrique v obtenue par composition : $E \times E \times \ldots E \to \otimes^p E \to \vee^p E$ est universelle. Elle factorise toute application p-linéaire symétrique ψ de E dans V par une application linéaire unique Ψ de $\vee^p E$ dans V telle que

$$\Psi(x_1 \vee x_2 \vee \ldots x_p) = \psi(x_1, x_2, \ldots x_p).$$

On a ainsi une bijection canonique :

$$\mathrm{Sym}_p(E; V) \equiv \mathfrak{L}(\vee^p E; V).$$

Proposition IV.1.C.

On a deux isomorphismes naturels de $Gl(E)$–modules, qui sont utilisés l'UN ET L'AUTRE :

a) $\overline{A}$ *et* $\overline{a}$ *de* $\wedge^p E$ *sur* $A^p E$, par passage au quotient des applications alternées A et a et $\otimes^p E$ sur $A^p E$:

$$\begin{aligned} \overline{A}(x_1 \wedge x_2 \wedge \ldots x_p) &= A(x_1 \otimes x_2 \otimes \ldots x_p) \\ &= \Sigma \varepsilon(\sigma) x_{\sigma 1} \otimes x_{\sigma 2} \otimes \ldots x_{\sigma p} \\ \overline{a}(x_1 \wedge x_2 \wedge \ldots x_p) &= a(x_1 \otimes x_2 \otimes \ldots x_p) \\ &= \frac{1}{p!} \Sigma \varepsilon(\sigma) x_{\sigma 1} \otimes x_{\sigma 2} \otimes \ldots x_{\sigma p}. \end{aligned}$$

b) $\overline{\mathrm{S}}$ *et* $\overline{s}$ *de* $\vee^p\mathrm{E}$ *sur* $\mathrm{S}^p\mathrm{E}$, par passage au quotient des applications symétriques S et s de $\otimes^p\mathrm{E}$ sur $\mathrm{S}^p\mathrm{E}$:

$$\overline{\mathrm{S}}(x_1 \vee x_2 \vee \dots x_p) = \Sigma x_{\sigma 1} \otimes x_{\sigma 2} \otimes \dots x_{\sigma p},$$
$$\overline{s}(x_1 \vee x_2 \vee \dots x_p) = \frac{1}{p!}\Sigma x_{\sigma 1} \otimes x_{\sigma 2} \otimes \dots x_{\sigma p}.$$

Il est clair que $\overline{a}$ et $\overline{s}$ sont les inverses des applications Φ et Ψ de la proposition IV.1.A.

IV.2 – Bases de $\vee^p\mathrm{E}$ et $\wedge^p\mathrm{E}$ associées à une base de E

Soit $e = \{e_j;\ j \in \mathrm{S}\}$ une base de l'espace vectoriel E indexée par un ensemble totalement ordonné S. Notons, pour toute suite $\mathrm{J} = \{j_1, j_2, \dots j_p\}, \otimes_{\mathrm{J}} e = e_{j_1} \otimes e_{j_2} \otimes \dots e_{j_p}$ et $\check{e}_{\mathrm{J}} = e_{j_1} \vee e_{j_2} \vee \dots e_{j_p}$ son image dans la puissance symétrique p-ème $\vee^p\mathrm{E}$. Les $\otimes_{\mathrm{J}} e$ forment *la base* $\otimes^p e$ *de* $\otimes^p\mathrm{E}$.

Considérons sur cette base la relation d'équivalence déterminée par l'action du groupe symétrique $\mathfrak{S}_p$ sur $\otimes^p\mathrm{E}$. Deux éléments sont équivalents si et seulement s'ils ne diffèrent que par l'ordre des facteurs. Dans chaque classe il y a un éléments et un seul dont les indices forment une suite croissante. Soient $\mathfrak{C}_p$ l'ensemble des suites croissantes $\mathrm{J} = \{j_1 \leqslant j_2 \leqslant \dots \leqslant j_p\}$ $\mathfrak{D}_p$ le sous-ensemble des suites strictement croissantes $l = \{i_1 < i_2 < \dots < i_p\}$ ($\mathfrak{D}_p$ est vide si $p > \dim \mathrm{E}$).

a) base de $\vee^p\mathrm{E}$. Les éléments d'une même classe d'équivalence de $\otimes^p\mathrm{E}$ ont même image dans $\vee^p\mathrm{E}$. Il en résulte que les $\check{e}_{\mathrm{J}}$ pour $\mathrm{J} \in \mathfrak{C}_p$ engendrent $\vee^p\mathrm{E}$. Une application linéaire f de $\otimes^p\mathrm{E}$ dans F est entièrement déterminée par ses valeurs sur la base $\otimes^p e$, et si elle est symétrique, elle prend la même valeur sur les éléments d'une même classe d'équivalence. Mais cette condition, nécessaire, est aussi suffisante puisqu'alors, quel que soit :

$$t = \Sigma \lambda_{k_1 k_2 \dots k_p} e_{k_1} \otimes e_{k_2} \otimes \dots e_{k_p}$$
$$f(\sigma \cdot t) = \Sigma \lambda_{k_1 \dots k_p} f(\sigma \cdot e_{k_1} \otimes \dots e_{k_p})$$
$$= \Sigma \lambda_{k_1 \dots k_p} f(e_{k_1} \otimes \dots e_{k_p}) = f(t).$$

Pour définir une application linéaire symétrique f de $\otimes^p\mathrm{E}$ dans F, on peut donc se donner arbitrairement la valeur de f sur les

$\otimes_{\mathrm{J}} e$, $J \in \mathfrak{C}_p$. La propriété universelle de $\vee^p\mathrm{E}$ montre alors que les $\{\check{e}_{\mathrm{J}}; \forall \mathrm{J} \in \mathfrak{C}_p\}$ forment une base de $\vee^p\mathrm{E}$. *Cette base est notée* $\vee^p e$.

Un élément $e_{j_1} \vee e_{j_2} \vee \ldots e_j$ de la base $\vee^p\mathrm{E}$; avec $j_i \leqslant j_z \leqslant j_p$ s'écrit plus simplement, en remplaçant la notation assez lourde du produit symétrique : $\vee$ par une simple juxtaposition, et en rassemblant les éléments égaux, sous la forme : $(e_{i_1})^{k_1} \ldots (e_{i_m})^{k_m}$ avec $k_1 + k_2 + \ldots k_m = p$. Les éléments de $\vee^p\mathrm{E}$, combinaisons linéaires finies des éléments de $\vee^p\mathrm{E}$, s'identifient donc aux *polynômes de degré p par rapport à l'ensemble « d'indéterminées »* $\{e_j;\ j \in \mathrm{S}\}$.

Lorsque E est de dimension finie n, les fonctions polynômiales sur E de degré p s'écrivent comme des polynômes :

$$\sum_{k_1+\ldots k_n=p} \lambda_{k_1\ldots k_n}(e^{*1})^{k_1} \ldots (e^{*n})^{k_n},$$

de degré p en les formes linéaires de base $e^{*1}, \ \ldots e^{*n}$.

b) base de $\wedge^p\mathrm{E}$. Il est d'usage de noter $e_{\mathrm{I}} = e_{i_1} \wedge e_{i_2} \wedge \ldots e_{i_p}$ pour toute suite strictement croissante $\mathrm{I} = \{i_1 < i_2 < \ldots < i_p\} \in \mathfrak{D}_p$.

Les e_{I} engendrent $\wedge^p\mathrm{E}$. Pour démontrer qu'ils sont linéairement indépendants, on peut soit faire usage de la propriété universelle comme on l'a fait pour $\vee^p\mathrm{E}$, soit, plus simplement, observer que les images des e_{I}, $\mathrm{I} \in \mathfrak{C}_p$ par l'application $\bar{\mathrm{A}}$ définie à la proposition IV.1.B ci-dessus sont linéairement indépendants dans $\otimes^p e$ puisque sommes d'éléments de $\otimes^p e$ tous distincts. Résumons :

Théorème IV.2. Une base e de l'espace vectoriel E détermine une base $\wedge^p e = \{e_{\mathrm{I}};\ \forall \mathrm{I} \in \mathfrak{D}_p\}$, ensemble des produits extérieurs $e_{\mathrm{I}} = e_{i_1} \wedge e_{i_2} \wedge \ldots e_{i_p}$ dont les indices forment une suite strictement croissante

– une base $\vee^p e = \{\vee_{\mathrm{J}} e;\ \forall \mathrm{J} \in \mathfrak{C}_p\}$ ensemble des produits symétriques $\vee_{\mathrm{J}} e = e_{j_1} \vee e_{j_2} \vee \ldots e_{j_p}$ dont les indices forment une suite croissante.

Corollaire IV.2. Si $\dim \mathrm{E} = n$ est finie, la dimension de $\wedge^p\mathrm{E}$ est $\begin{pmatrix} n \\ p \end{pmatrix}$ et $\dim \wedge^p \mathrm{E} = \dim \wedge^{n-p} \mathrm{E}$; $\dim \wedge^n \mathrm{E} = 1$. Si e est une base de E, $e_1 \wedge e_2 \wedge \ldots e_n$ est la base associée de la droite $\wedge^n\mathrm{E}$.

La différence de comportement entre les puissances extérieures et symétriques est mise en évidence dans la :

Proposition IV.2.

Le produit extérieur de p vecteurs : $x_1 \wedge x_2 \wedge \dots x_p$ est non nul si et seulement si les vecteurs sont linéairement indépendants.

Le produit symétrique de p vecteurs : $x_1 \vee x_2 \vee \dots x_p$ est non nul si et seulement si les vecteurs sont non nuls.

Preuve : Si l'un des x_j est combinaison linéaire des autres, le produit extérieur est alors combinaison linéaire de produits extérieurs comportant deux facteurs égaux donc nuls.

Réciproquement si $x_1, x_2 \dots x_n$ sont linéairement indépendants il existe une base de E dont ils sont les p premiers éléments et leur produit extérieur est un élément de la base associée de $\wedge^p \mathrm{E}$ donc non nul.

Quant au produit symétrique de vecteurs non nuls, l'ordre des facteurs étant indifférent, on peut placer au début q des p vecteurs de telle sorte qu'ils soient linéairement indépendants et que les $r = p - q$ autres en soient des combinaisons linéaires. On ordonne ces q vecteurs : $e_1, e_2, \dots e_q$ et en écrivant les autres sous la forme $x_j = \Sigma \lambda_{k,j} e_k$ l'un des $\lambda_{k,j} \neq 0$, on voit immédiatement que le produit symétrique est non nul en vertu du théorème de la base IV.1.

Définition IV.2.

Les éléments de $\wedge^p \mathrm{E}$ sont appelés les *p-vecteurs* de E. Un *p-vecteur* qui est le produit extérieur de p vecteurs de E : $x_1 \wedge x_2 \wedge \dots x_p$ est appelé un p-vecteur décomposable. Les éléments du dual $(\wedge^p \mathrm{E})^*$ sont appelés les *p-formes* sur E, les éléments de $\wedge^p \mathrm{E}^*$ sont les *p-covecteurs*.

VI.3 – Puissances extérieures et symétriques d'une application linéaire

Soit a une application linéaire de E dans F. Les applications p-linéaires :

$$(x_1, x_2, \dots, x_p) \longrightarrow ax_1 \wedge ax_2 \wedge \dots ax_p$$
$$(x_1, x_2, \dots, x_p) \longrightarrow ax_1 \vee ax_2 \vee \dots ax_p$$

sont, la première, une application alternée, à valeurs dans $\wedge^p$E, la seconde une application symétrique, à valeurs dans $\vee^p$E. Elles se factorisent donc par des applications linéaires $\overline{\wedge}^p a$ et $\overline{\vee}^p a$, appelés respectivement *puissance extérieure p-ème* et *puissance symétrique p-ème de a* suivant les diagrammes :

$$\begin{tikzcd}
E \times E \times \dots E \arrow[rr] \arrow[dr, "(x_1, x_2, \dots x_p)"'] & & \Lambda^p E \\
& \substack{\otimes^p E \\ (x_1 \otimes x_2 \otimes \dots x_p)} \arrow[r] & \substack{\Lambda^p E \\ (x_1 \Lambda x_2 \Lambda \dots x_p)} \arrow[u, "\substack{(ax_1 \Lambda ax_2 \Lambda \dots ax_p) \\ (\bar{\Lambda}^p a)}"']
\end{tikzcd}$$

$$\begin{tikzcd}
E \times E \times \dots E \arrow[rr] \arrow[dr, "(x_1, x_2, \dots x_p)"'] & & V^p E \\
& \substack{\otimes^p E \\ (x_1 \otimes x_2 \otimes \dots x_p)} \arrow[r] & \substack{V^p E \\ (x_1 V x_2 V \dots x_p)} \arrow[u, "\substack{(ax_1 V ax_2 V \dots ax_p) \\ (\bar{V}^p a)}"']
\end{tikzcd}$$

[Les barres des notations $\overline{\vee}^p a$ et $\overline{\wedge}^p a$ sont essentielles : il faut éviter de confondre, par exemple $\overline{\wedge}^2 a$ avec $\wedge^2 a = a \wedge a = 0$ carré extérieur de l'élément a de l'espace vectoriel $\mathcal{L}$(E; F).]

La construction précédente peut aussi s'interpréter par le fait que la puissance tensorielle p-ème de a : $\overline{\otimes}^p a$ applique le sous-espace d'antisymétrisation de $\otimes^p$E dans celui de $\otimes^p$F et le sous-espace de symétrisation de $\otimes^p$E dans celui de $\otimes^p$F. $\overline{\otimes}^p a$ passe donc au quotient par chacun d'eux pour donner $\overline{\wedge}^p a$ et $\overline{\vee}^p a$.

Si b est une application linéaire de F dans H, on a : $\overline{\wedge}^p(b \circ a) = \overline{\wedge}^p b \circ \overline{\wedge}^p a$ et $\overline{\vee}^p(b \circ a) = \overline{\vee}^p b \circ \overline{\vee}^p a$. Il en résulte, en prenant H $=$ E et $b = \overline{a}^1$ que si a est un isomorphisme, il en est de même de $\overline{\wedge}^p a$ et de $\overline{\vee}^p a$. Si a est un automorphisme de l'espace vectoriel E, $\overline{\wedge}^p a$ est un automorphisme de $\wedge^p$E. $\overline{\vee}^p a$ un automorphisme de $\vee^p$E, et

$$\overset{-1}{(\overline{\wedge}^p a)} = \overline{\wedge}^p(\overset{-1}{a}) \quad \overset{-1}{(\overline{\vee}^p a)} = \overline{\vee}^p(\overset{-1}{a}).$$

Il résulte de ce qui précède que le groupe Gl(E) ainsi que tous ses sous-groupes sont représentés linéairement dans chaque $\wedge^p$E et chaque $\vee^p$E. Plus généralement, si θ est une représentation linéaire d'un groupe G dans E, les $\{\overline{\wedge}^p \theta(g);\ g \in \mathrm{G}\}$ forment la puissance extérieure p-ème de θ, représentation de G dans $\wedge^p$E, les $\{\overline{\vee}^p \theta(g); g \in \mathrm{G}\}$ forment la puissance symétrique p-ème de θ, représentation de G dans $\vee^p$E.

Si maintenant ρ est une représentation de l'algèbre de Lie $\mathfrak{g}$ dans E (cf Définition III.14.), son extension $\widetilde{\rho}$ à $\otimes^p E$ applique en eux-mêmes les sous-espaces d'alternation L_p et de symétrisation M_p (cf Définition IV.1.A.) comme on le vérifie immédiatement. Elle passe donc au quotient dans $\otimes^p E/L_p$ et $\otimes^p E/M_p$ et définit des extensions $\widehat{\rho}$ et $\check{\rho}$ de ρ dans $\wedge^p E$ et $\vee^p E$. Sur des éléments décomposables les opérations de $a \in \mathfrak{g}$ sont (on écrit $\rho(a)x = ax$) :

$$\begin{aligned}
\widehat{\rho}(a) \cdot x_1 \wedge x_2 \wedge \dots x_p &= ax_1 \wedge x_2 \wedge \dots x_p + x_1 \wedge ax_2 \wedge \dots x_p \\
&\quad + \dots x_1 \wedge x_2 \wedge \dots ax_p, \\
\check{\rho}(a) \cdot x_1 \vee x_2 \vee \dots x_p &= ax_1 \vee x_2 \vee \dots x_p + \dots x_1 \vee x_2 \vee \dots ax_p.
\end{aligned}$$

Les épimorphismes de $\otimes^p E$ sur ses quotients $\wedge^p E$ et $\vee^p E$ sont naturellement des $\mathfrak{g}$-morphismes (Définition III.11.A). En rassemblant ces extensions, on obtient les représentations associées de $\mathfrak{g}$ comme algèbre de dérivations des algèbres $\wedge E$ et $\vee E$ (voir les § IV.9 et § IV.13). En particulier par extension de la représentation adjointe, $\mathfrak{g}$ opère dans $\wedge\mathfrak{g}$ et $\vee\mathfrak{g}$ ainsi que dans $\wedge\mathfrak{g}^*$ et $\vee\mathfrak{g}^*$. Les invariants de ces représentations sont appelés les *invariants antisymétriques et symétriques de* $\mathfrak{g}$.

On sait que l'on appelle rang d'une application linéaire a de E dans F la dimension du sous-espace image Ima. Lorsque ce rang est fini, dire que le rang de a est r signifie donc que l'on peut trouver r vecteurs $x_1, x_2, \dots x_r$ de E tels que $ax_1, ax_2, \dots ax_r$ soient linéairement indépendants, tandis que quels que soient $(r+1)$ vecteurs de E, leurs images par a sont liées. Le rang de a est donc le plus grand entier p tel que $\overline{\wedge}^p a \neq 0$.

Lorsque la dimension de l'espace vectoriel E est *finie*, soit $\dim E = n$, $\dim \wedge^n E = 1$: $\wedge^n E$ est une droite. On appelle alors *déterminant* d'un opérateur a de E, et on note *dét* a le scalaire, rapport de l'homothétie $\overline{\wedge} a$ de $\wedge^n E$. Quels que soient les n vecteurs $x_1, x_2, \dots x_n$ de E :

$$(\overline{\wedge}^n a)(x_1 \wedge x_2 \wedge \dots x_n) = ax_1 \wedge ax_2 \wedge \dots ax_n = (\text{dét}\, a) x_1 \wedge x_2 \wedge \dots x_n.$$

Si $x_1, x_2, \dots x_n$ sont linéairement indépendants, c'est-à-dire forment une base de E, leur produit extérieur est non nul et il résulte de la Proposition IV.2 que $ax_1, ax_2, \dots ax_n$ sont linéairement indépendants, autrement dit a inversible, si et seulement si $\text{dét}\, a \neq 0$. On a comme cas particulier de ce qui précède, $\text{dét}(b \circ a) = \text{dét}\, b \cdot \text{dét}\, a$ et $\text{dét}\, \overset{-1}{a} = \overset{-1}{(\text{dét}\, a)}$.

Si $A = |a^i_j|$ est la matrice de a, $ae_k = \Sigma a^{jk}_k e_{j_k}$

et $(\overset{-n}{\wedge} a)(e_1 \wedge e_2 \wedge \ldots e_n) = \Sigma a^{j1}_1 a^{j2}_2 \ldots a^{jn}_n e_{j_1} \wedge e_{j_2} \wedge \ldots e_{j_n}$.

Les produits extérieurs de vecteurs de base apparaissant dans la somme du second membre ne sont différents de zéro que si leurs indices sont tous distincts, formant une permutation de $(1, 2, \ldots n)$:

$$1 \longrightarrow j_1 = \sigma 1, 2 \longrightarrow j_2 = \sigma 2, \ldots n \longrightarrow j_n = \sigma n.$$

Chaque permutation σ de $\mathfrak{S}_n$ apparait ainsi une fois et une seule et l'expression se réduit à : $\sum\limits_{\sigma} a^{\sigma 1}_1 a^{\sigma 2}_2 \ldots a^{\sigma n}_n e_{\sigma 1} \wedge e_{\sigma 2} \wedge \ldots e_{\sigma n}$ d'où

$$\boxed{\text{dét}\, a = \Sigma_{,\sigma}\, \varepsilon(\sigma) a^{\sigma 1}_1 a^{\sigma 2}_2 \ldots a^{\sigma n}_n}\,.$$

$\wedge^n$E s'identifie évidemment à la droite des pseudoscalaires contravariants et $\wedge^n E^*$ à la droite des pseudoscalaires covariants (§ III.6.).

Cas particulier : les puissances extérieures et symétriques de l'injection φ d'un sous-espace vectoriel F de E dans E sont des injections : $\overline{\wedge}^p \varphi$ et $\overline{\vee}^p \varphi$ qui permettent d'identifier les puissances extérieures ou symétriques des sous-espaces de E, à des sous-espaces des puissances, extérieures ou symétriques de E.

En effet le théorème de la base incomplète affirme que l'on peut compléter une base $\{e_i;\ i \in I\}$ de F en une base $e = \{e_j; j \in J$, avec $I \subset J\}$ de E. Dès lors, chacune des puissances $\overline{\wedge}^p \varphi$ ou $\overline{\vee}^p \varphi$ est l'identité sur la base de $\wedge^p$F ou $\vee^p$F associée.

IV.4 – Multiplication extérieure et symétrique

Il existe deux formes (duales) de multiplication extérieure ou symétrique : la multiplication directe d'éléments (des $\wedge^p$E ou des $\vee^p$E), à laquelle est consacré ce paragraphe, et la multiplication des applications multilinéaires, alternées ou symétriques, qui sera vue aux § IV.8 et 9. En dimension finie, ces deux formes de multiplication peuvent être identifiées et se réduisent donc à une seule.

Théorème IV.4.A. Soit E un espace vectoriel de dimension finie ou infinie. Il existe des applications bilinéaires naturelles (multiplications) : $\wedge^p\mathrm{E} \times \wedge^p\mathrm{E} \longrightarrow \wedge^{p+q}\mathrm{E}$: ***produit extérieur***, noté $\wedge$ et $\vee^p\mathrm{E} \times \vee^q\mathrm{E} \longrightarrow \vee^{p+q}\mathrm{E}$: ***produit symétrique***, noté $\vee$.

Si $x_1, x_2, \ldots x_p, y_1, y_2, \ldots y_q$ sont $p + q$ vecteurs de E, ces produits s'écrivent sur les éléments décomposables correspondants :

$$\begin{aligned}(x_1 \wedge x_2 \wedge \ldots x_p) \wedge (y_1 \wedge y_2 \wedge \ldots y_q)& \\ &= x_1 \wedge x_2 \wedge \ldots x_p \wedge y_1 \wedge \ldots y_q\\ (x_1 \vee x_2 \vee \ldots x_p) \vee (y_1 \vee y_2 \vee \ldots y_q)& \\ &= x_1 \vee x_2 \vee \ldots x_p \vee y_1 \vee \ldots y_q.\end{aligned}$$

Ces produits sont associatifs. Le produit symétrique est commutatif. Le produit extérieur est anticommutatif : si $u \in \wedge^p\mathrm{E}$ et $v \in \wedge^q\mathrm{E}$, $v \wedge u = (-1)^{pq} u \wedge v$.

En particulier, si $u \in \wedge^p\mathrm{E}$ et p impair, $u \wedge u = 0$.

Preuve : Considérons les applications composées de la multiplication tensorielle : $\otimes^p\mathrm{E} \times \otimes^q\mathrm{E} \longrightarrow \otimes^{p+q}\mathrm{E}$ et des projections de $\otimes^{p+q}\mathrm{E}$ sur ses quotients $\wedge^{p+q}\mathrm{E}$ et $\vee^{p+q}\mathrm{E}$.

Nous allons voir que ces applications bilinéaires passent au quotient par les sous-espaces d'antisymétrisation et de symétrisation respectivement. Cela va résulter du lemme suivant :

Lemme IV.4. Soient $\mathrm{A}, \mathrm{B}, \mathrm{C}$ trois espaces vectoriels sur le même corps, M un sous-espace de A, N un sous-espace de B, f une application bilinéaire de $\mathrm{A} \times \mathrm{B}$ dans C telle que $f(m, b) = 0$ et $f(a, n) = 0$ $\forall a \in \mathrm{A}$, $b \in \mathrm{B}$, $m \in \mathrm{M}$, $n \in \mathrm{N}$. f « passe alors au quotient » et détermine une application bilinéaire $\overline{f}$ de $(\mathrm{A}/\mathrm{M}) \times (\mathrm{B}/\mathrm{N})$ dans C telle que $\overline{f}\,(a \bmod \mathrm{M}, b \bmod \mathrm{N}) = f(a, b)$.

En effet, on a d'après la bilinéarité de f :

$$f(a + m, b + n) = f(a, b) + f(m, b) + f(a, n) + f(m, n) = f(a, b).$$

On peut donc définir : $\overline{f}(a \bmod \mathrm{M}, b \bmod \mathrm{N}) = f(a, b)$.

Si maintenant L_p est le sous-espace d'antisymétrisation de $\otimes^p\mathrm{E}$, il est engendré par les éléments $\{t - \varepsilon(\sigma)\sigma t;\ \forall t \in \otimes^p\mathrm{E};\ \forall \sigma \in \mathfrak{S}_p\}$. En appelant $\widetilde{\sigma}$ la permutation de $\otimes^{p+q}\mathrm{E}$ égale à σ sur les p premiers facteurs et à l'identité sur les q derniers, on a quels que soient $t \in \otimes^p\mathrm{E}$ et $u \in \otimes^q\mathrm{E}$:

$$(t - \varepsilon(\sigma) \cdot \sigma t) \otimes u = t \otimes u - \varepsilon(\widetilde{\sigma})\widetilde{\sigma} \cdot t \otimes u.$$

Ce raisonnement et les raisonnements analogues montrent que la multiplication tensorielle : $\otimes^p\mathrm{E} \times \otimes^q\mathrm{E} \to \otimes^{p+q}\mathrm{E}$ applique : $\mathrm{L}_p \times \otimes^q\mathrm{E}$ et $\otimes^p\mathrm{E} \times \mathrm{L}_q$ dans L_{p+q} et pour les sous-espaces de symétrisation : $\mathrm{M}_p \times \otimes^q\mathrm{E}$ et $\otimes^p\mathrm{E} \times \mathrm{M}_q$ dans M_{p+q}, c.q.f.d. L'associativité de ces produits résulte de l'associativité de la multiplication tensorielle. Il reste à vérifier l'anticommutativité du produit extérieur. Or, si $u \in \wedge^p\mathrm{E}$ et $\mathrm{V} \in \wedge^p\mathrm{E}$ sont décomposables, on passe de $u \wedge v$ à $v \wedge u$ en faisant passer chacun des q facteurs de v au-dessus des p facteurs de u, occasionnant à chaque fois p changements de signe.

Le raisonnement qui a servi de preuve au théorème précédent permet d'obtenir la généralisation suivante :

Théorème IV.4.B. Soient $\mathrm{E}, \mathrm{E}', \mathrm{V}$ trois espaces vectoriels sur le même corps, f une application $(p+q)$-linéaire : $\mathrm{E} \times \ldots \mathrm{E} \times \mathrm{E}' \times \ldots \mathrm{E}' \to \mathrm{V}$, soit :

$$(x_1, \ldots x_p,\ y_1, \ldots y_q) \to f(x_1, \ldots x_p,\ y_1, \ldots y_q),$$

telle que pour $(y_1, \ldots y_q)$ fixés, l'application partielle p-linéaire : $(x_1, \ldots x_p) \to f(x_1, \ldots x_p, y_1, \ldots y_q)$ soit alternée (resp. symétrique) et que pour $(x_1, \ldots x_p)$ fixés, l'application partielle : $(y_1, \ldots y_q) \to f(x_1, \ldots x_p, y_1, \ldots y_q)$ soit alternée (resp. symétrique).

f détermine alors une application bilinéaire $\widetilde{f}$ de $\wedge^p\mathrm{E} \times \wedge^p\mathrm{E}'$ (resp. de $\vee^p\mathrm{E} \times \vee^q\mathrm{E}'$) dans V telle que $\widetilde{f}(x_1 \wedge \ldots x_p;\ y_1 \wedge \ldots y_q) = f(x_1, \ldots x_p,\ y_1, \ldots y_q)$ (resp. $\widetilde{f}(x_1 \vee \ldots x_p;\ y_1 \vee \ldots y_q) = f(x_1, \ldots x_p,\ y_1, \ldots y_q)$.

Preuve : On a un diagramme commutatif d'applications :

$$\begin{tikzcd}
\mathrm{E} \times \mathrm{E} \ldots \mathrm{E} \times \mathrm{E}' \times \ldots \mathrm{E}' \arrow[r, "f"] \arrow[d] & \mathrm{V} \\
(\otimes^p\mathrm{E}) \times (\otimes^q\mathrm{E}') \arrow[r, "\mathrm{F}", "\mu"'] & (\otimes^p\mathrm{E}) \otimes (\otimes^q\mathrm{E}') \arrow[u]
\end{tikzcd}$$

Avec les notations précédentes, l'hypothèse sur f entraîne que F s'annule sur les sous-espaces $\mathrm{L}_p \times (\otimes^q\mathrm{E}')$ et $(\otimes^p\mathrm{E}) \times \mathrm{L}'_q$ (resp. $\mathrm{M}_p \times (\otimes^q\mathrm{E}')$ et $(\otimes^p\mathrm{E}) \times \mathrm{M}'_q$) et passe donc au quotient suivant le lemme ci-dessus.

Proposition IV.4.

Soient $x_1, x_2, \dots x_p$ p vecteurs linéairement indépendants de E, F le sous-espace dont ils forment une base, $u = x_1 \wedge x_2 \wedge \dots x_p \in \wedge^p\mathrm{E}$. Alors :

a) pour un vecteur x de E : $x \in \mathrm{F} \leftrightarrow x \wedge u = 0$.

b) si $y_1, y_2, \dots, y_p$ sont p vecteurs linéairement indépendants de E, on a $y_1 \wedge y_2 \wedge \dots y_p = \lambda x_1 \wedge x_2 \wedge \dots x_p$ si et seulement si les sous-espaces ayant pour bases $(x_1, x_2, \dots x_p)$ et $(y_1, y_2, \dots y_p)$ sont les mêmes.

Preuve :

a) si $x \in \mathrm{F}$, $x \wedge u = 0$; réciproquement, si $x \wedge u = 0$, $x, x_1, x_2, \dots x_p$ sont liés d'où il résulte que $x \in \mathrm{F}$.

b) Si $y_1, y_2, \dots y_p$ est une autre base de F, les y_j sont des combinaisons linéaires des x_k d'où une égalité : $y_1 \wedge y_2 \wedge \dots y_p = \lambda x_1 \wedge x_2 \wedge \dots x_p$ avec λ non nul. Réciproquement une telle égalité entraîne que les y_j appartiennent à F en vertu de *a)* et ils en forment donc une base.

Remarque : on a une bijection naturelle entre les sous-espaces de dimension p de E et les droites de $\wedge^p\mathrm{E}$ portant des p-vecteurs *décomposables*.

IV.5 – Dualité

Soient $\alpha_1, \alpha_2, \dots \alpha_p$, p formes linéaires sur E. Leur produit tensoriel en tant qu'applications de E dans le corps K : $\alpha_1 \overline{\otimes} \alpha_2 \overline{\otimes} \dots \alpha_p$ est une forme linéaire sur $\otimes^p\mathrm{E}$, donc un élément de $(\otimes^p\mathrm{E})^*$, tandis que leur produit tensoriel en tant qu'éléments de E^* est l'élément $\alpha_1 \otimes \alpha_2 \otimes \dots \alpha_p$ de $\otimes^p\mathrm{E}^*$. L'application : $(\alpha_1, \alpha_2, \dots \alpha_p) \longrightarrow \alpha_1 \overline{\otimes} \alpha_2 \overline{\otimes} \dots \alpha_p$ étant p-linéaire détermine d'ailleurs une application canonique de $\otimes^p\mathrm{E}^*$ dans $(\otimes^p\mathrm{E})^*$ dont on voit aisément, en prenant une base de E, que c'est une injection. C'est une bijection lorsque la dimension de E est finie. Le crochet de dualité entre $\otimes^p\mathrm{E}$ et $\otimes^p\mathrm{E}^* \subset (\otimes^p\mathrm{E})^*$ s'écrit sur *les éléments décomposables :*

$$\langle x_1 \otimes x_2 \otimes \dots x_p;\ \alpha_1 \otimes \alpha_2 \otimes \dots \alpha_p \rangle = \langle x_1; \alpha_1 \rangle \langle x_2; \alpha_2 \rangle \dots \langle x_p; \alpha_p \rangle.$$

On peut aussi interpréter cette égalité comme l'expression de la puissance tensorielle p-ème de la forme bilinéaire crochet de $\mathrm{E} \times \mathrm{E}^*$ sur les tenseurs décomposables. On peut enfin l'obtenir directement en composant l'application linéaire de $\mathrm{E} \otimes \ldots \mathrm{E} \otimes \mathrm{E}^* \otimes \ldots \mathrm{E}^*$ dans K, associée à l'application $2p$-linéaire : $(x_1, \ldots x_p, \alpha_1, \ldots \alpha_p) \to \langle x_1; \alpha_1\rangle \ldots \langle x_p; \alpha_p\rangle$, avec la multiplication tensorielle de $(\otimes^p \mathrm{E}) \times (\otimes^p \mathrm{E})^*$ dans $(\otimes^p \mathrm{E}) \otimes (\otimes^p \mathrm{E})^*$.

Soit $\sigma \in \mathfrak{S}_p$. Dans le crochet précédent, il est équivalent d'effectuer la permutation σ sur les x_j ou la permutation $\overset{-1}{\sigma}$ sur les α_k :

$$\langle \sigma \cdot (x_1 \otimes \ldots x_p); \alpha_1 \otimes \ldots \alpha_p\rangle = \Pi \langle x_{\overset{-1}{\sigma} j}; \alpha_j\rangle = \Pi \langle x_k; \alpha_{\sigma k}\rangle$$

$$= \langle x_1 \otimes \ldots x_p;\ \overset{-1}{\sigma} \cdot (\alpha_1 \otimes \ldots \alpha_p)\rangle.$$

σ s'écrit en effet : $\begin{pmatrix} 1 & 2 & \ldots p \\ \sigma_1 & \sigma_2 & \ldots \sigma_p \end{pmatrix} \equiv \begin{pmatrix} \overset{-1}{\sigma}1 & \overset{-1}{\sigma}2 \ldots & \overset{-1}{\sigma}p \\ 1 & 2 & \ldots p \end{pmatrix}$.

En sommant l'égalité précédente sur l'ensemble $\mathfrak{S}_p$, on voit qu'il est équivalent d'appliquer l'opérateur de symétrisation $\mathrm{S} = \Sigma\sigma = \Sigma\sigma^{-1}$, au premier argument ou au second. On obtient ainsi le *crochet symétrisé :*

$$\langle \mathrm{S} \cdot (x_1 \otimes \ldots x_p); \alpha_1 \otimes \ldots \alpha_p\rangle = \langle x_1 \otimes \ldots x_p; \mathrm{S}(\alpha_1 \otimes \ldots \alpha_p)\rangle.$$

Puisqu'une permutation et son inverse ont la même signature, en multipliant par $\varepsilon(\sigma) = \varepsilon(\overset{-1}{\sigma})$ avant de sommer sur $\mathfrak{S}_p$, on voit de la même façon qu'il est équivalent d'appliquer l'opérateur d'antisymétrisation $\mathrm{A} = \Sigma\varepsilon(\sigma)\sigma = \Sigma\varepsilon(\sigma)\overset{-1}{\sigma}$ au premier argument ou au second, ce qui donne le *crochet antisymétrisé :*

$$\langle \mathrm{A} \cdot (x_1 \otimes \ldots x_p); \alpha_1 \otimes \ldots \alpha_p\rangle = \langle x_1 \otimes \ldots x_p; \mathrm{A} \cdot (\alpha_1 \otimes \ldots \alpha_p)\rangle.$$

Définition IV.5.A

On appelle *déterminant* de p vecteurs $x_1, x_2, \ldots x_p$ et de p

covecteurs $\alpha_1, \alpha_2, \dots \alpha_p$ d'un espace vectoriel E le scalaire :

$$\begin{aligned}
\mathrm{dét}(x_1, \dots x_p; \alpha_1, \dots \alpha_p) &= \langle \mathrm{A}(x_1 \otimes \dots x_p); \alpha_1 \otimes \dots \alpha_p \rangle \\
&= \langle x_1 \otimes \dots x_p; \mathrm{A}(\alpha_1 \otimes \dots \alpha_p) \rangle \\
&= \sum_\sigma \varepsilon(\sigma) \Pi \langle x_{\sigma j}; \alpha_j \rangle \\
&= \sum \varepsilon(\sigma) \Pi \langle x_j, \alpha_{\sigma j} \rangle = \mathrm{dét}|\langle \alpha_i; x_j \rangle|
\end{aligned}$$

déterminant de la matrice des $\langle \alpha_i; x_j \rangle$.

On appelle *permanent* de p vecteurs et p covecteurs de E le scalaire :

$$\begin{aligned}
\mathrm{per}(x_1, \dots x_p; \alpha_1, \dots \alpha_p) &= \langle \mathrm{S}(x_1 \otimes \dots x_p); \alpha_1 \otimes \dots \alpha_p \rangle \\
&= \langle x_1 \otimes \dots x_p; \mathrm{S}(\alpha_1 \otimes \dots \alpha_p) \rangle \\
&= \sum_\sigma \Pi \langle x_{\sigma j}; \alpha_j \rangle = \sum_\sigma \Pi \langle x_j; \alpha_{\sigma j} \rangle \\
&= \mathrm{per}|\langle \alpha_i, x_j \rangle|,
\end{aligned}$$

en appelant permanent d'une matrice carrée $|a_{ij}|$ (notion définie simultanément par Cauchy et Binet en 1812) le scalaire

$$\mathrm{per}|a_{ij}| = \sum_\sigma a_{1,\sigma 1}, a_{2,\sigma 2} \cdots a_{p,\sigma p} = \sum_\sigma a_{\sigma 1,1} a_{\sigma 2,2} \cdots a_{\sigma p,p}.$$

Remarque : Ce sont les opérateurs d'antisymétrisation A et de symétrisation S qui sont utilisés et *NON* a et s qui amèneraient avec eux d'insupportables coefficients $\dfrac{1}{p!}$ à traîner dans toutes les formules.

Par application immédiate du théorème IV.4.B ci-dessus, les crochets symétrisés et antisymétrisés passent au quotient définissant ainsi des crochets de dualité entre $\wedge^p \mathrm{E}$ et $\wedge^p \mathrm{E}^*$ et entre $\vee^p \mathrm{E}$ et $\vee^p \mathrm{E}^*$ qui s'écrivent *sur les éléments décomposables* de ces espaces :

$$\begin{aligned}
\langle x_1 \wedge \dots x_p; \alpha_1 \wedge \dots \alpha_p \rangle &= \mathrm{dét}|\langle \alpha_i, x_j \rangle| \\
\langle x_1 \vee \dots x_p; \alpha_1 \vee \dots \alpha_p \rangle &= \mathrm{per}|\langle \alpha_i, x_j \rangle|.
\end{aligned}$$

Supposons maintenant E de dimension finie n

Si $e = \{e_s; s \in \mathrm{S}\}$ est une base de E, $e^* = \{e^{*s}; s \in \mathrm{S}\}$ la base duale de E* où l'ensemble d'indices S est supposé totalement ordonné, on a vu au § IV.2. que les $e_\mathrm{I} = e_{i_1} \wedge e_{i_2} \wedge \dots e_{i_p}$ où $\mathrm{I} = (i_1 < i_2 < \dots i_p)$ parcourt l'ensemble D_p des suites strictement croissantes d'éléments de S forment la base associée $\wedge^p e$ de $\wedge^p \mathrm{E}$ tandis que les $\{e^{*\mathrm{I}}; \mathrm{I} \in \mathrm{D}_p\}$ forment la base associée $\wedge^p e^*$ de $\wedge^p \mathrm{E}^*$. Ces bases sont évidemment duales : $\langle e_{\mathrm{I}}; e^{*\mathrm{I}'}\rangle = \delta_{\mathrm{I}}^{\mathrm{I}'} = \{1 \text{ si } \mathrm{I} = \mathrm{I}'; 0 \text{ si } \mathrm{I} \neq \mathrm{I}'\}$.

De la même façon les bases $\vee^p e$ et $\vee^p e^*$ de $\vee^p \mathrm{E}$ et $\vee^p \mathrm{E}^*$ sont « presque duales ». On a en effet : $\langle e_{i_1} \vee e_{i_2} \vee \dots e_{i_p};\ e^{*j_1} \vee e^{*j_2} \vee \dots e^{*j_p}\rangle = 0$ si les suites croissantes $\mathrm{I} = \{i_1 \leqslant i_2 \leqslant \dots i_p\}$ et $\mathrm{J} = \{j_2 \leqslant j_2 \leqslant \dots j_p\}$ sont distinctes, et $= m_{\mathrm{I}}$ si $\mathrm{I} = \mathrm{J}$, m_{I} étant le nombre de permutations qui laissent I invariante ou « multiplicité de I ». Il en résulte que le crochet met « en dualité » $\wedge^p \mathrm{E}$ et $\wedge^p \mathrm{E}^*$, $\vee^p \mathrm{E}$ et $\vee^p \mathrm{E}^*$ suivant la définition suivante :

Définition IV.5.B.

On dit qu'une forme bilinéaire b sur le couple d'espaces vectoriels $\mathrm{E} \times \mathrm{F}$ les met en dualité si elle possède les propriétés suivantes :

1) quel que soit le vecteur non nul x de E, il existe $y \in \mathrm{F}$ tel que $b(x, y) \neq 0$

2) quel que soit le vecteur non nul y de F, il existe $x \in \mathrm{E}$ tel que $b(x, y) \neq 0$.

Si ρ est l'application linéaire de F dans E^* associée à droite de b par : $b(x, y) = \langle x, \rho(y)\rangle$ et γ l'application linéaire de E dans F^* associée à gauche de b par : $b(x, y) = \langle \gamma(x), y\rangle$, les conditions 1) et 2) expriment que γ et ρ sont injectives. En dimension finie, puisque $\gamma =^t \rho$, il en résulte qu'elles sont alors bijectives.

Dans les applications, la dualité intervient souvent non pas, comme on vient de le voir, entre les puissances extérieures ou symétriques de même degré sur E et E^*, mais directement entre E et les puissances extérieures et symétriques de E^*. Les éléments de $\wedge^p \mathrm{E}^*$ et $\vee^p \mathrm{E}^*$ sont alors considérés comme des fonctions sur $\mathrm{E} \times \mathrm{E} \times \dots \mathrm{E}$ (p fois) p-linéaires, alternées ou symétriques, point de vue développé au § IV.8. Dans le cas de $\vee^p \mathrm{E}^*$, on préfère même considérer ses éléments f non pas comme des fonctions de p vecteurs distincts de E : $f(x_1, x_2, \dots x_p)$ mais plus simplement comme des fonctions d'un *seul* vecteur x : $F(x) = f(x, x, \dots x)$ ou *fonctions polynomiales de degré p sur* E. A partir de F, on retrouve f, par l'algorithme de polarisation. Dans le cas $p = 2$, c'est la relation usuelle entre formes bilinéaires symétriques et formes quadratiques.

IV.6 – Développement de déterminants. Formules de Lagrange et de Laplace

Dans un espace vectoriel E de dimension finie n, le déterminant de n vecteurs $x_1, x_2, \dots x_n$ relativement à une base ordonnée $e = (e_1, e_2, \dots e_n)$ est le déterminant de ces vecteurs et des n formes coordonnées :

$$\begin{aligned} \text{dét}_e(x_1, x_2, \dots x_n) &= \text{dét}(x_1, x_2, \dots x_n; e^{*1}, e^{*2}, \dots e^{*n}) \\ &= \langle x_1 \wedge x_2 \wedge \dots x_n;\ e^{*1} \wedge e^{*2} \wedge \dots e^{*p} \rangle \\ &= \text{dét}|x^i_j| = \text{dét}\,\mathrm{X} \end{aligned}$$

où X est la matrice dont la j-ème colonne X_j est la matrice des composantes x^i_j de x_j dans e. On a donc :

$$\begin{aligned} x_i \wedge x_2 \wedge \dots x_n &= (\text{dét}\,\mathrm{X}) e_1 \wedge e_2 \wedge \dots e_n \\ &= \text{dét}_e(x_1, \dots x_n) e_1 \wedge e_2 \wedge \dots e_n. \end{aligned}$$

Si a est un opérateur linéaire de E :

$$\begin{aligned} a(x_1) \wedge a(x_2) \wedge \dots a(x_n) &= (\text{dét}\, a) x_1 \wedge x_2 \wedge \dots x_n \\ &= \text{dét}\, a \text{dét}\,\mathrm{X} \cdot e_1 \wedge e_2 \wedge \dots e_n \end{aligned}$$

d'où : $\text{dét}_e(a(x_1), a(x_2) \dots a(x_n)) = \text{dét}\, a \cdot \text{dét}_e(x_1, x_2, \dots x_n)$ et $\text{dét}\, a = \text{dét}_e(a(e_1), a(e_2), \dots a(e_n))$. On retrouve sur $\wedge^n \mathrm{E}$ le fait que l'opérateur linéaire a multiplie les « volumes algébriques » (cf. III.6) par le scalaire $\text{dét}\, a$.

Si $\mathrm{A} = |a^i_j|$ est une matrice numérique à m lignes et n colonnes, soient $p \leqslant \text{Inf}\,(m, n)$, $L = \{1 \leqslant i_1 < i_2 < \dots i_p \leqslant m\}$ une suite strictement croissante de p indices définissant un ensemble de lignes et $\mathrm{C} = \{1 \leqslant j_1 < j_2 < \dots j_p \leqslant n\}$ une suite strictement croissante de p indices définissant un ensemble de colonnes. On appelle *mineur de* A pour les *p-lignes de* L *et les p colonnes de* C et on note $\mathrm{A}^{\mathrm{L}}_{\mathrm{C}}$ le déterminant de la matrice carrée partielle dont les lignes et colonnes sont celles de L et C.

Si a est une application linéaire de l'espace vectoriel E, de dimension n, dans l'espace vectoriel F, de dimension m, et A sa matrice relativement à des bases e et f de E et F, les mineurs d'ordre p de A sont les éléments de la matrice de $\wedge^p a$ relativement aux bases $\wedge^p e$ et $\wedge^p f$ de $\wedge^p \mathrm{E}$ et $\wedge^p \mathrm{F}$. En effet, les $e_{\mathrm{C}} = e_{j_1} \wedge e_{j_2} \wedge \dots e_{j_p}$, où $\mathrm{C} = \{1 \leqslant j_1 < j_2 < \dots j_p \leqslant n\}$ parcourt l'ensemble des suites croissantes de p indices entre 1 et n, forment la base $\wedge^p e$ et les $f^*_{\mathrm{L}} = f^*_{i_1} \wedge f^*_{i_2} \wedge \dots f^*_{i_p}$ la base $\wedge^p f^*$ de $\wedge^p \mathrm{F}^*$. D'où :

$$\begin{aligned}&\langle \wedge^p a(e_{\mathrm{C}}), f^{*\mathrm{L}} \rangle \\ &\qquad = \langle a(e_{j_1}) \wedge a(e_{j_2}) \wedge \dots a(e_{j_p});\; f^{*i_1} \wedge f^{*i_2} \wedge \dots f^{*i_p} \rangle = \mathrm{A}^{\mathrm{L}}_{\mathrm{C}}.\end{aligned}$$

Proposition IV.6

Le p-ème coefficient c_p du polynôme caractéristique d'un opérateur linéaire a d'un espace vectoriel E de dimension finie :

$$\mathrm{P}_a(t) = \mathrm{dét}(t\mathbb{1} - a) = t^n + c_1 t^{n-1} + c_2 t^{n-2} + \dots c_n$$

est $c_p = (-1)^p$ Trace $\overline{\wedge}^p a$.

Preuve : Dans une base quelconque e de E, on a $(te_1 - ae_1) \wedge (te_2 - ae_2) \wedge \dots \wedge (te_n - ae_n) = \mathrm{dét}(t\mathbb{1} - a) \cdot e_1 \wedge e_2 \wedge \dots e_n$

Chaque terme du développement du premier membre est obtenu en retenant les facteurs $(-ae_j)$ pour les indices d'une suite $\mathrm{J} = \{1 \leqslant j_1 < j_2 \dots j_p \leqslant n\}$. En réordonnant les facteurs de façon à placer devant les ae_j le terme s'écrit $\varepsilon(\mathrm{J}) t^{n-p} (-1)^p \overline{\wedge}^p a(e_{\mathrm{J}}) \wedge e_{\mathrm{J}}$, où $\varepsilon(\mathrm{J})$ est la signature de la permutation faisant passer de $(1, 2, \dots n)$ à la suite obtenue en écrivant J' à la suite de J. Mais si A est la matrice de a dans e, le produit $\overline{\wedge}^p a(e_{\mathrm{J}}) \wedge e_{\mathrm{J}'}$, est égal à $\mathrm{A}^{\mathrm{J}}_{\mathrm{J}} e_{\mathrm{J}} \wedge e_{\mathrm{J}'} = \varepsilon(\mathrm{J}) \mathrm{A}^{\mathrm{J}}_{\mathrm{J}}\, e_1 \wedge e_2 \wedge \dots e_n$, où $\mathrm{A}^{\mathrm{J}}_{\mathrm{J}}$ est le mineur de A associé aux lignes et aux colonnes de J.

Pour un entier p fixé, la somme $\sum\limits_{\mathrm{J}} \mathrm{A}^{\mathrm{J}}_{\mathrm{J}} = \mathrm{Trace} \overline{\wedge}^p a$, on a donc :

$$\boxed{\begin{aligned}&\mathrm{dét}(t \cdot \mathbb{1} - a) = t^n - (\mathrm{Tr}\, a) t^{n-1} + \dots \\ &\qquad\qquad + (-1)^p (\mathrm{Tr}\, \overline{\wedge}^p a) t^{n-p} + \dots (-1)^n \mathrm{dét}\, a\end{aligned}}\,.$$

Plus généralement, si $x_1, x_2, \dots x_p$ sont p vecteurs de l'espace vectoriel E de dimension $n, p \leqslant n$ et si X est la matrice à n lignes et p colonnes de leurs composantes dans une base e, on a :

$$x_1 \wedge x_2 \wedge \dots x_p = \sum_{\mathrm{L}} \mathrm{X}^{\mathrm{L}} e_{\mathrm{L}}$$

avec $\mathrm{X}^{\mathrm{L}} = \langle x_1 \wedge x_2 \wedge \ldots x_p;\ e^{*i_1} \wedge e^{*i_2} \wedge \ldots e^{*i_p}\rangle$ = mineur d'ordre p de la matrice X correspondant aux p lignes de $\mathrm{L} = \{1 \leqslant i_1 < i_2 < \ldots i_p \leqslant n\}$.

Les X^{L} sont appelés les *coordonnées grassmaniennes* du p-vecteur décomposable $x_1 \wedge x_2 \wedge \ldots x_p$ dans la base e.

Si l'on désigne par x_j^{L} la composante du vecteur x_j dans le sous-espace défini par les vecteurs de base $e_{i_1}, e_{i_2}, \ldots e_{i_p}$ de L, on a :

$$\mathrm{X}^{\mathrm{L}} e_{\mathrm{L}} = x_1^{\mathrm{L}} \wedge x_2^{\mathrm{L}} \wedge \ldots \wedge x_p^{\mathrm{L}} \text{ et } x_1 \wedge x_2 \wedge \ldots x_p = \sum_{\mathrm{L}} x_1^{\mathrm{L}} \wedge x_2^{\mathrm{L}} \wedge \ldots \wedge x_p^{\mathrm{L}}.$$

On peut dualement considérer p covecteurs $\alpha^1, \alpha^2, \ldots \alpha^p$ et la matrice à p lignes et n colonnes A de leurs composantes dans e^*, la *ligne* A^i étant celle des composantes de α^i : $\alpha^i = \sum \alpha^i_j e^{*j}$. On a alors :

$$\alpha^1 \wedge \alpha^2 \wedge \ldots \alpha^p = \sum_{\mathrm{C}} \mathrm{A}_{\mathrm{C}} e^{*\mathrm{C}}$$

où $\mathrm{A}_{\mathrm{C}} = \langle e_{j_1} \wedge e_{j_2} \wedge \ldots e_{j_p};\, \alpha^1 \wedge \alpha^2 \wedge \ldots \alpha^p\rangle$ = mineur d'ordre p de A correspondant aux p colonnes de $\mathrm{C} = \{1 \leqslant j_1 < j_2 < \ldots j_p \leqslant n\}$.

En développant le crochet de dualité entre $\wedge^p\mathrm{E}$ et $\wedge^p\mathrm{E}^*$:

$$\begin{aligned}&\langle x_1 \wedge x_2 \wedge \ldots x_p;\ \alpha^1 \wedge \alpha^2 \wedge \ldots \alpha^p\rangle = \text{dét}|\langle x_j, \alpha^i\rangle| \\ &= \langle \sum_{\mathrm{L}} \mathrm{X}^{\mathrm{L}} e_{\mathrm{L}};\ \sum_{\mathrm{C}} \mathrm{A}_{\mathrm{C}} e^{*\mathrm{C}}\rangle = \sum_{\mathrm{L}} \mathrm{X}^{\mathrm{L}} \mathrm{A}_{\mathrm{L}}.\end{aligned}$$

Or dans la base e : $\langle x_j, \alpha^i\rangle = \langle \sum x_j^k e_k,\ \sum \alpha^i_l e^{*l}\rangle = \sum x_j^k \alpha^i_k$ est obtenu en multipliant terme à terme la i-ème ligne de la matrice $(p \times n)$A par la j-ème colonne de la matrice $(n \times p)$X.

D'où : $|\langle x_j; \alpha^i\rangle| = \mathrm{AX}$, ce qui donne la *formule de Lagrange* exprimant le déterminant du produit d'une matrice $(p \times n)$A par une matrice $(n \times p)$X avec $p \leqslant n$.

$$\boxed{\text{dét}(\mathrm{AX}) = \sum_{\mathrm{L}} \mathrm{A}_{\mathrm{L}} \mathrm{X}^{\mathrm{L}}}.$$

où L parcourt les suites $\{1 \leqslant j_1 < j_2 < \ldots j_p \leqslant n\}$.

Si l'on prend pour A la transposée de X, ${}^t\mathrm{X}_\mathrm{L} = \mathrm{X}^\mathrm{L}$ et

$$\boxed{\text{dét}({}^t\mathrm{XX}) = \sum_\mathrm{L} (\mathrm{X}^\mathrm{L})^2}.$$

Une interprétation euclidienne donne une généralisation du théorème de Pythagore.

Pour deux vecteurs x, y d'un espace de dimension trois, c'est la formule élémentaire :

$$\text{dét} \begin{vmatrix} x^1 x^2 x^3 \\ \\ y^1 y^2 y^3 \end{vmatrix} \begin{vmatrix} x^1 y^3 \\ x^2 y^2 \\ x^3 y^3 \end{vmatrix} = (x^1y^2 - x^2y^1)^2 + (x^2y^3 - x^3y^2)^2$$
$$+ (x^3y_1 - x^1y^3)^2$$
$$= ((x^1)^2 + (x^2)^2 + (x^3)^2)((y^1)^2 + (y^2)^2 + (y^3)^2)$$
$$- (x^1y^1 + x^2y^2 + x^3y^3)^2.$$

De la même façon, le *développement de Laplace* du déterminant d'une matrice $(n \times n)$X s'obtient en considérant les colonnes de X comme les composantes dans une base e de n vecteurs : $x_1, x_2, \dots x_n$, soit : $\text{dét}\,\mathrm{X} = \text{dét}_e(x_1, x_2, \dots x_n)$.

Si $\mathrm{J} = \{1 \leqslant j_1 < j_2 < j_p \leqslant n\}$ est une suite strictement croissante de p indices et J′ la suite strictement croissante complémentaire, notons $\varepsilon(\mathrm{J}) = (-1)^{j_1+j_2+\dots j_p-(1+2+\dots p)}$ la signature de la permutation qui permet de passer de $(1, 2, \dots n)$ à la suite JJ′ formée de J suivie de J′, X(J) et X(J′) les sous-matrices de X formées respectivement des colonnes de J et de J′. On a :

$$x_1 \wedge x_2 \wedge \dots x_n = \varepsilon(\mathrm{J}) x_{j_1} \wedge \dots x_{j_p} \wedge x_{j'_1} \wedge \dots x_{j'_{n-p}}$$
$$x_{j_1} \wedge \dots x_{j_p} = \textstyle\sum_{,\mathrm{I}} \mathrm{X}(\mathrm{J})^\mathrm{I} e_\mathrm{I} \text{ où } \mathrm{I} = \{1 \leqslant i_1 < i_2 < \dots i_p \leqslant n\}$$
$$x_{j'_1} \wedge \dots x_{j'_p} = \textstyle\sum_{,\mathrm{L}} \mathrm{X}(\mathrm{J}')^\mathrm{L} e_\mathrm{L} \text{ où } \mathrm{L} = \{1 \leqslant l_1 < l_2 < \dots l_{n-p} \leqslant n\}.$$

Mais le mineur $\mathrm{X}(\mathrm{J})^\mathrm{I}$ de X(J) est le mineur $\mathrm{X}^\mathrm{I}_\mathrm{J}$ de X, le mineur $\mathrm{X}(\mathrm{J}')^\mathrm{L}$ de X(J′) est le mineur $\mathrm{X}^\mathrm{L}_{\mathrm{J}'}$ de X, d'où :

$$x \wedge x_2 \wedge \dots x_n = \varepsilon(\mathrm{J}) \sum_\mathrm{L} \mathrm{X}^\mathrm{I}_\mathrm{J} \mathrm{X}^\mathrm{L}_{\mathrm{J}'} e_\mathrm{I} \wedge e_\mathrm{L}$$

$e_{\mathrm{I}} \wedge e_{\mathrm{L}}$ est nul sauf dans le seul cas où $\mathrm{L} = \mathrm{I}'$ et $e_{\mathrm{I}} \wedge e_{\mathrm{I}'}$ est alors égal à $\varepsilon(\mathrm{I}) e_1 \wedge e_2 \wedge \dots e_n$. On obtient finalement le *développement de Laplace de* $\operatorname{dét} \mathrm{X}$ *suivant les colonnes de* J :

$$\operatorname{dét} \mathrm{X} = \varepsilon(\mathrm{J}) \sum_{\mathrm{I}} \varepsilon(\mathrm{I}) \mathrm{X}^{\mathrm{I}}_{\mathrm{J}} \mathrm{X}^{\mathrm{I}'}_{\mathrm{J}'},$$

(sommation sur toutes les suites $\mathrm{I} = \{1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n\}$) somme multipliée par $\varepsilon(\mathrm{J})$, des produits de tous les mineurs de $\mathrm{X}(\mathrm{J})$ par les mineurs complémentaires de $\mathrm{X}(\mathrm{J}')$ affectés des signes ad hoc.

Par contre, si H est une suite strictement croissante de $(n-p)$ indices différente de J' l'expression $\sum \varepsilon(\mathrm{I}) \mathrm{X}^{\mathrm{I}}_{\mathrm{J}} \mathrm{X}^{\mathrm{I}}_{\mathrm{H}}$ est nulle puisque le produit extérieur des vecteurs x_j, $j \in \mathrm{J}$ et x_h, $h \in \mathrm{H}$ contient alors deux termes égaux, et est donc nul.

Un développement identique vaut naturellement pour les lignes.

Exemples :

1) pour $p = 1$, $\mathrm{J} = \{j\}$:

$$\operatorname{dét} \mathrm{X} = (-1)^{j-1} \sum_{i=1}^{n} (-1)^{i-1} x^i_j \mathrm{X}^i_j = \sum_{i=1}^{n} (-1)^{i+j} \mathrm{X}^i_j x^i_j,$$

en notant pour simplifier X^i_j le déterminant de la matrice obtenue en supprimant de X la i-ème ligne et la j-ème colonne $(-1)^{i+j} \mathrm{X}^i_j$ est souvent appelé le mineur de l'élément de x^i_j avec son signe. La matrice transposée $\zeta = |\zeta^j_i|$ avec $\zeta^j_i = (-1)^{i+j} \mathrm{X}^i_j$ satisfait donc à l'égalité : $\zeta \mathrm{X} = \operatorname{dét} \mathrm{X} \mathbb{1}$. Si X est inversible, la matrice inverse est donc $\overset{-1}{\mathrm{X}} = \dfrac{1}{\operatorname{dét} \mathrm{X}} \zeta$ avec $(\overset{-1}{\mathrm{X}})^i_j = (-1)^{i+j} \dfrac{\mathrm{X}^j_i}{\operatorname{dét} \mathrm{X}}$.

2) une application immédiate du développement de Laplace donne les égalités où A et C sont des sous-matrices carrées :

$$\operatorname{dét} \begin{vmatrix} \mathrm{AB} \\ \mathrm{OC} \end{vmatrix} = \operatorname{dét} \mathrm{A} \cdot \operatorname{dét} \mathrm{C}; \operatorname{dét} \begin{vmatrix} \mathrm{AO} \\ \mathrm{DC} \end{vmatrix} = \operatorname{dét} \mathrm{A} \cdot \operatorname{dét} \mathrm{C}.$$

Nous allons maintenant procéder à une analyse plus détaillée du développement de Laplace et démontrer qu'il revient à une antisymétrisation effectuée en deux étapes.

Distinguons dans le groupe symétrique $\mathfrak{S}_n$ de permutations de n objets, numérotés $(1, 2, \dots n)$, les sous-groupes $\mathfrak{S}_p$, permutant

$(1, 2, \ldots p)$ et laissant fixes les $(n-p)$ derniers, et $\mathfrak{S}'_{n-p}$ permutant $(p+1, p+2, \ldots n)$ et laissant fixes les p premiers. Nous allons décrire l'ensemble des classes à gauche de $\mathfrak{S}_n$ modulo le sous-groupe produit $\mathfrak{S}_p \times \mathfrak{S}'_{n-p}$ et montrer qu'il s'identifie à l'ensemble $\mathrm{T}_{p,q}$ des permutations τ de $(1, 2, \ldots, n)$ qui respectent l'ordre partiel des p premiers éléments $\tau 1 < \tau 2 \ldots \tau p$ ainsi que l'ordre partiel des $(n-p)$ derniers : $\tau(p+1) < \ldots < \tau n$.

La donnée d'une suite strictement croissante $\mathrm{J} = \{1 \leqslant j_1 < j_2 < \ldots < j_p \leqslant n\}$, c'est-à-dire de p éléments distincts de $(1, 2, \ldots n)$ détermine la suite complémentaire J' et une permutation τ de $T_{p,q}$. L'ensemble $T_{p,q}$ contient donc $\binom{n}{p} = \dfrac{n!}{p!(n-p)!}$ éléments. La signature $\varepsilon(t)$ de τ est égale à $\varepsilon(\mathrm{J})$ suivant la notation ci-dessus puisque τ applique $(1, 2, \ldots n)$ sur JJ'.

Soit maintenant σ une permutation quelconque de $(1, 2, \ldots n)$. En plaçant dans l'ordre naturel les éléments $(\sigma 1, \sigma 2, \ldots \sigma p)$ on définit de façon unique une permutation μ de $(1, 2, \ldots p)$ telle que $\sigma\mu 1 < \sigma\mu 2 < \ldots \sigma\mu p$. De la même façon il existe une permutation unique ν de $(p+1, p+2, \ldots n)$ telle que $\sigma\nu(p+1) < \sigma\nu(p+2) < \ldots \sigma\nu(n)$. Il existe donc un élément unique $\mu \times \nu \in \mathfrak{S}_p \times \mathfrak{S}'_{n-p}$ tel que $\sigma \circ (\mu \times \nu) = \tau$ appartienne à $\mathrm{T}_{p,q}$. Cela revient à dire que σ s'écrit d'une façon et d'une seule : $\sigma = \tau \circ (\overset{-1}{\mu} \times \overset{-1}{\nu})$ avec $\overset{-1}{\mu} \times \overset{-1}{\nu} \in \mathfrak{S}_p \times \mathfrak{S}'_{n-p}$ et démontre que $\mathfrak{S}_n / \mathfrak{S}_p \times \mathfrak{S}'_{n-p}$ s'identifie au sous-ensemble $\mathrm{T}_{p,q}$ de $\mathfrak{S}_n$. On a ainsi une application naturelle bijective :

$$\mathfrak{S}_n \longrightarrow \mathrm{T}_{p,q} \times (\mathfrak{S}_p \times \mathfrak{S}'_{n-p}),$$

qui permet de décomposer l'opérateur d'antisymétrisation A de $\mathfrak{S}_n$ en antisymétrisation partielle par rapport aux p premiers et $(n-p)$ derniers éléments, suivie d'une antisymétrisation « transverse » par les éléments de T :

$$\boxed{\mathrm{A} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\sigma = \sum_{\tau \in \mathrm{T}p,q} \varepsilon(\tau)\tau \sum_{\mu,\nu} \varepsilon(\mu)\varepsilon(\nu)(\mu \times \nu)}\,.$$

Appliquons cette décomposition à l'expression du déterminant de n vecteurs et de n formes :

$$
\begin{aligned}
&\text{dét}(\alpha_1, \alpha_2, \ldots x_n; x_1, x_2, \ldots x_n) \\
&\qquad = \langle \mathrm{A}(\alpha_1 \otimes \alpha_2 \otimes \ldots \alpha_n); x_1 \otimes x_2 \otimes \ldots x_n \rangle \\
&\qquad = \sum_{\tau \in \mathrm{T}} \varepsilon(\tau) \langle \sum_{\mu} \varepsilon(\mu) \alpha_{\tau\mu 1} \otimes \ldots \alpha_{\tau\mu p}; x_1 \otimes \ldots x_p \rangle \\
&\qquad \langle \sum_{\nu} \varepsilon(\nu) \alpha_{\tau\nu(p+1)} \otimes \ldots \alpha_{\tau\nu n}; x_{p+1} \otimes \ldots x_n \rangle .
\end{aligned}
$$

Dans le premier crochet, τ étant fixe, on effectue l'antisymétrisation sur $\alpha_{\tau 1} \otimes \alpha_{\tau 2} \otimes \ldots \alpha_{\tau p}$. Dans le second crochet, τ étant fixe, on effectue l'antisymétrisation sur $\alpha_{\tau(p+1)} \otimes \alpha_{\tau(p+2)} \otimes \ldots \alpha_{\tau n}$.

Compte-tenu des raisonnements faits au § IV.4 sur le passage au quotient dans les crochets antisymétrisés, on obtient *le développement de Laplace du déterminant de n vecteurs et de n formes :*

$$
\boxed{
\begin{aligned}
&(\langle \alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_n; x_1 \wedge x_2 \wedge \ldots x_n \rangle \\
&= \sum_{\tau \in \mathrm{T}p,q} \varepsilon(\tau) \langle \alpha_{\tau 1} \wedge \ldots \alpha_{\tau p}; x_1 \wedge \ldots x_p \rangle \\
&\qquad \langle \alpha_{\tau(p+1)} \wedge \ldots \alpha_{\tau n}; x_{p+1} \wedge \ldots x_n \rangle \\
&= \sum_{\tau \in \mathrm{T}p,q} \varepsilon(\tau) \langle \alpha_1 \wedge \ldots \alpha_p; x_{\tau 1} \wedge \ldots x_{\tau p} \rangle \\
&\qquad \langle \alpha_{p+1} \wedge \ldots \alpha_n; x_{\tau(p+1)} \wedge \ldots x_{\tau n} \rangle
\end{aligned}
}.
$$

que l'on aurait pu obtenir directement en prenant les développements de dét$|\langle \alpha_i, x_j \rangle|$ par rapport aux p premières colonnes, correspondant à $x_1, x_2, \ldots x_p$ pour la première égalité, ou par rapport aux p premières lignes correspondant à $\alpha_1, \alpha_2, \ldots \alpha_p$ pour la seconde.

IV.7 – Analogues des formules de Lagrange et de Laplace pour les permanents (Binet, Cauchy, 1812)

Dans un espace vectoriel E de dimension finie n, le permanent de n vecteurs $x_1, x_2, \ldots x_n$ relativement à une base $e = (e_1, e_2, \ldots e_n)$ est le permanent de ces vecteurs et des n formes coordonnées, c'est-à-dire la composante du produit symétrique $x_1 \vee x_2 \vee \ldots x_n$ relative au vecteur $e_1 \vee e_2 \vee \ldots e_n$ de la base

$\vee e$ de $\vee^n E$ associée à e :

$$\begin{aligned}\operatorname{per}_e(x_1, x_2, \dots x_n) &= \operatorname{per}(x_1, x_2, \dots x_n; e^{*1}, e^{*2}, \dots e^{*n}) \\ &= \langle x_1 \vee x_2 \vee \dots x_n; e^{*1} \vee e^{*2} \vee \dots e^{*n} \rangle \\ &= \operatorname{per}|x_j^i| = \operatorname{per} X = \sum_{\sigma \in \mathfrak{S}_n} x_1^{\sigma 1} x_2^{\sigma 2} \dots x_n^{\sigma n}\end{aligned}$$

où X est comme précédemment la matrice dont la j-ème colonne X_j est formée des composantes x_j^i de x_j.

Plus généralement, si $x_1, x_2, \dots x_p$ sont p vecteurs de E, la propriété universelle de $\vee^p E$ du théorème IV.1.A montre que toute fonction numérique p-linéaire symétrique de $x_1, x_2, \dots x_p$ s'obtient par une forme linéaire sur $\vee^p E$, donc par un élément de $\vee^p E^*$. On a vu au § IV.2. qu'une base e de E détermine une base $\{\check{e}_I ; I = \{1 \leqslant i_1 \leqslant i_2 \leqslant \dots i_p \leqslant n\}$ de l'espace vectoriel $\vee^p E^*$ de ces formes où I parcourt l'ensemble $C_{p,n}$ des suites croissantes de p éléments de $(1, 2, \dots n)$ et :

$$x_1 \vee x_2 \vee \dots x_p = \sum_{I \in C_{p,n}} \check{X}^I \check{e}_I.$$

Les fonctions coordonnées pour les suites I de $C_{p,n}$:

$$(x_1, x_2, \dots x_p) \to \check{X}^I = \frac{1}{m_I} \sum_{\sigma \in \mathfrak{S}_p} x_{\sigma 1}^{i_1} x_{\sigma 2}^{i_2} \dots x_{\sigma p}^{i_p} x_{\sigma 2}^{i_2} \dots x_{\sigma p}^{i_p}$$

(où m_I est la multiplicité de I : si r_j est le nombre de fois où l'entier j apparaît dans I, $m_I = (r_1!)(r_2!) \dots (r_n!)$ en convenant que $0! = 1$) sont les analogues des coordonnées grassmaniennes d'un p-vecteur.

Si X est une matrice numérique $|x_j^i|$ à n lignes et p colonnes, en interprétant les colonnes comme les composantes de p vecteurs d'un espace vectoriel de dimension n, on peut lui associer les fonctions : $X \to \check{X}^I$ suivant la définition ci-dessus qui jouent, dans le cas symétrique le rôle des mineurs. Ces derniers sont des déterminants. On peut donner à $\check{X}^I$ l'habillage d'un permanent de la façon suivante : en appelant X(I) la matrice $p \times p$ ayant pour k-ème ligne la ligne d'indice i_k de X (ce n'est donc une sous-matrice de X *que* si I est strictement croissante), on a $\check{X}^I = \dfrac{1}{m_I} \operatorname{per} X(I)$.

En interprétant de la même façon une matrice numérique A à p lignes et n colonnes comme la matrice des composantes de p

covecteurs $\alpha_1, \alpha_2, \dots \alpha_p$ d'un espace vectoriel E de dimension n, on a :

$$\alpha^1 \vee \alpha^2 \vee \dots \alpha^p = \sum_{\mathrm{J}\in \mathrm{C}p,n} \check{\mathrm{A}}_\mathrm{J}\check{e}^{*\mathrm{J}} \text{ avec } \check{\mathrm{A}}_\mathrm{J} = \frac{1}{m_\mathrm{J}}\mathrm{per}\ \mathrm{A}(\mathrm{J})$$

où A(J) est cette fois la matrice $p \times p$ ayant pour k-ème colonne la colonne d'indice j_k de A.

Comme précédemment la matrice $p \times p$: $|\langle \alpha^i, x_j\rangle|$ est le produit des m matrices A et X. D'où :

$$\begin{aligned}\mathrm{per}|\langle \alpha^i, x_j\rangle| &= \mathrm{per}\ (\mathrm{AX}) = \mathrm{per}(\alpha^1, \alpha^2, \dots \alpha^p; x_1, x_2, \dots x_p)\\ &= \langle \alpha^1 \vee \alpha^2 \vee \dots \alpha^p; x_1 \vee x_2 \vee \dots \vee x_p\rangle\\ &= \langle \textstyle\sum \check{\mathrm{A}}_\mathrm{J}\check{e}^{*\mathrm{J}}; \sum \mathrm{X}^\mathrm{I}\check{e}_\mathrm{I}\rangle\end{aligned}$$

ce qui donne l'analogue de la formule de Lagrange :

$$\mathrm{per}(\mathrm{AX}) = \sum_{\mathrm{I}\in\mathrm{C}p,n} m_\mathrm{I}\check{\mathrm{A}}_\mathrm{I}\check{\mathrm{X}}^\mathrm{I} = \sum_{\mathrm{I}\in\mathrm{C}p,n} \frac{1}{m_\mathrm{I}}\mathrm{perA}(\mathrm{I})\mathrm{perX}(\mathrm{I})$$

et $\mathrm{per}({}^t\mathrm{XX}) = \sum\limits_{\mathrm{I}\in\mathrm{C}np} \dfrac{1}{m_\mathrm{I}}(\mathrm{perX}(\mathrm{I}))^2.$

Le développement de Laplace du permanent d'une matrice $(n \times n)$X peut s'obtenir, comme pour le déterminant, en considérant per X comme le permanent de ses vecteurs colonnes $x_1, x_2, \dots x_n$ relativement à la base e. Si $\mathrm{J} = \{1 \leqslant j_1 < j_2 < \dots j_p \leqslant n\}$ est une suite d'indices de colonnes de X, J′ la suite complémentaire, X(J) et X(J′) les sous-matrices de X formées des colonnes de J et de celles de J′, on a :

$$\begin{aligned}x_1 \vee x_2 \vee \dots x_n &= x_{j_1} \vee x_{j_2} \vee \dots \vee x_{jp} \vee x_{j'_1}, \vee x_{j'_2} \vee \dots x_{j'_{n-p}}\\ &= (\sum_\mathrm{I} \check{\mathrm{X}}(\mathrm{J})^\mathrm{I}\check{e}_\mathrm{I})(\sum_\mathrm{L} \check{\mathrm{X}}(\mathrm{J}')^\mathrm{L}\check{e}_\mathrm{L} = (\sum_{\mathrm{I},\mathrm{L}} \check{\mathrm{X}}(\mathrm{J})^\mathrm{I}\check{\mathrm{X}}(\mathrm{J}')^\mathrm{L}\check{e}_\mathrm{I} \vee \check{e}_\mathrm{L}\end{aligned}$$

où I et L parcourent les suites croissantes de p ou de $(n-p)$ indices de 1 à n. Pour qu'un terme appartienne à la composante suivant $e_1 \vee e_2 \vee \dots e_n$ il est nécessaire et suffisant que la réunion de I et de L soit une permutation de $(1, 2, \dots n)$ ce qui impose que I et L soient strictement croissantes et que L soit la complémentaire I′ de

I. Dans ce cas $m_{\mathrm{I}} = 1 = m_{\mathrm{I}'} \cdot \check{\mathrm{X}}(\mathrm{J})^{\mathrm{I}}$ est le permanent de la sous-matrice de X : X(I, J) formée des éléments appartenant aux lignes de I et colonnes de J, et de même pour $\check{\mathrm{X}}(\mathrm{J}')^{\mathrm{I}'}$.

On obtient ainsi le développement analogue au développement de Laplace correspondant au choix des p colonnes de J :

$$\operatorname{per} \mathrm{X} = \sum_{\mathrm{I}} \operatorname{per} \mathrm{X}(\mathrm{I}, \mathrm{J}) \operatorname{per} \mathrm{X}(\mathrm{I}', \mathrm{J}')$$

où la somme est étendue aux suites strictement croissantes de p éléments de $(1, 2, \dots n)$.

IV.8 – Produit extérieur d'applications multilinéaires alternées sur un espace vectoriel de dimension quelconque E

$\wedge^p \mathrm{E}$ est défini comme le *quotient* $\otimes^p \mathrm{E}/_{\mathrm{L}_p}$ de $\otimes^p \mathrm{E}$ par le sous-espace d'antisymétrisation (ou alternation) L_p. Son dual $(\wedge^p \mathrm{E})^*$ se définit par conséquent comme un *sous-espace* de $(\otimes^p \mathrm{E})^*$ à savoir le sous-espace des formes linéaires sur $\otimes^p \mathrm{E}$ nulles sur L_p ou p-formes. Plus généralement, en vertu de la propriété universelle de $\wedge^p \mathrm{E}$, les applications p-linéaires alternées de E dans un espace vectoriel F, c'est-à-dire telles que, quel que soit $\sigma \in \mathfrak{S}_p$, on ait $f \circ \sigma = \varepsilon(\sigma) f$, s'identifient aux applications linéaires de $\otimes^p \mathrm{E}$ dans F qui s'annulent sur L_p, puisque ce dernier est engendré par les éléments $\sigma t - \varepsilon(\sigma) t$ de $\otimes^p \mathrm{E}$.

Si h est une application de $\otimes^p \mathrm{E}$ dans F, on appelle antisymétrisée de h l'application composée de h et de l'opérateur d'antisymétrisation sur $\otimes^p \mathrm{E}$: $\mathrm{A} = \Sigma \varepsilon(\sigma)\sigma$ ou $a = \dfrac{1}{p!}\mathrm{A}$ suivant les cas. Par exemple, sur un élément décomposable, $h \cdot a(x_1 \otimes x_2 \otimes \dots x_p) = \dfrac{1}{p!}\Sigma \varepsilon(\sigma) h(x_{\sigma_1} \otimes x_{\sigma_2} \otimes \dots x_{\sigma_p})$ est la « moyenne antisymétrisée » des valeurs de h. On a bien sûr, $(h \circ a) \circ a = h \circ a$. D'autre part, toute forme linéaire alternée sur $\otimes^p \mathrm{E}$ est une antisymétrisée puisque $f \circ a = \dfrac{1}{p!}\Sigma \varepsilon(\sigma) f \circ \sigma = f$.

Soient maintenant E, F, H trois espaces vectoriels sur le même corps de scalaires K, f une application p-linéaire alternée de E dans F, h une application q-linéaire alternée de E dans H : le cas le plus important est celui où les espaces de valeurs F et H sont le corps K, f et h étant des formes multilinéaires alternée.

Le *produit extérieur* $f \wedge h$ est l'application $(p+q)$-linéaire alternée de E dans le produit tensoriel $F \otimes H$, donc dans K si $F = H = K$ puisque $K \otimes K \equiv K$, obtenue en complétant l'antisymétrisation partielle existant déjà relativement aux p premiers et q derniers arguments, par l'antisymétrisation transverse décrite à la fin du § IV.6 du produit tensoriel de f et h. C'est la procédure minimale, et aussi la plus rationnelle, d'obtention de l'antisymétrisation totale. On définit donc :

$$(f \wedge h)(x_1, x_2, \ldots x_{p+q})$$
$$= \sum_{\tau \in T} \varepsilon(\tau) f(x_{\tau 1}, \ldots x_{\tau p}) \otimes h(x_{\tau p+1}), \ldots x\tau(p+q)$$

où $T = T_{p,q}$ est, comme au § IV.6, le sous-ensemble de $\mathfrak{S}_{p+q}$ formé des permutations τ qui respectent les ordres partiels des p premiers et des derniers éléments :

$$\tau_1 < \tau_2 < \ldots < \tau_p \text{ et } \tau(p+1) < \tau(p+2) < \ldots \tau(p+q)$$

et où le produit tensoriel est à remplacer par le produit ordinaire si $F = H = K$.

Exemples :

Si $p = q = 1$, $(f \wedge h)(x_1, x_2) = f(x_1) \otimes h(x_2) - f(x_2) \otimes h(x_1)$.
Si $p = 1$ et q quelconque :

$$(f \wedge h)(x_0, x_1, x_2, \ldots x_q) = \sum_{i=0}^{q} (-1)^i f(x_i) \otimes h(x_0, x_1, \ldots x_i, \ldots x_q).$$

Remarque : la décomposition de l'opérateur d'antisymétrisation du § IV.6 montre que l'application $f \wedge h$ définie ci-dessus est antisymétrique. Mais cela peut aussi se démontrer directement : il suffit de vérifier que si deux arguments consécutifs, par exemple x_j et x_{j+1} sont égaux, $f \wedge h$ s'annule. Or, pour une permutation τ telle que $\overset{-1}{\tau} j$ et $\overset{-1}{\tau}(j+1)$ appartiennent tous deux à $(1, 2, \ldots p)$, c'est $f(x_{\tau 1}, \ldots x_{\tau p})$ qui est nulle; pour τ telle que $\overset{-1}{\tau} j$ et $\overset{-1}{\tau}(j+1)$ appartiennent à $(p+1, \ldots p+q)$, c'est h qui s'annule; enfin les autres permutations τ s'associent par paires de signatures opposées donnant une somme nulle.

La décomposition « de Laplace » du groupe symétrique $\mathfrak{S}_{p+q} = \mathrm{T}_{p,q} \times (\mathfrak{S}_p \times \mathfrak{S}'_q)$ permet également une définition du produit symétrique de deux applications p et q-linéaires symétriques f et g de E dans les espaces F et H, par une symétrisation transverse « minimale » de leur produit tensoriel, soit :

$$(f \vee g)(x_1, x_2, \ldots x_{p+q}) = \sum_{\tau \in \mathrm{T}p,q} f(x_{\tau 1}, \ldots x_{\tau p}) \otimes (x_{\tau(p+1)}, \ldots x_{\tau(p+q)}.$$

Cette application $f \vee g$ est alors une application $(p+q)$-linéaire symétrique de E dans le produit tensoriel $\mathrm{F} \otimes \mathrm{H}$.

Plus généralement, considérons une partition de l'ensemble $(1, 2, \ldots n)$ en parties consécutives : $\mathrm{P}_1 = (1, 2, \ldots p_1)$, $\mathrm{P}_2 = (p_1 + 1, \ldots p_1 + p_2) \ldots \mathrm{P}_k = (p_1 + p_2 + \ldots p_{k-1} + 1, \ldots, n)$ de cardinaux respectifs $p_1, p_2, \ldots p_k$. Soient $\mathfrak{S}_{(j)}$ le sous-groupe des permutations de $(1, 2, \ldots n)$ qui permute P_j en laissant invariants les éléments des P_i, $i \neq j$, et T la partie de $\mathfrak{S}_n$ formée des permutations de $(1, 2, \ldots n)$ dont les restrictions à *chacune* des P_j conserve l'ordre.

Le même raisonnement que dans le cas où $k = 2$ montre que $\mathfrak{S}_n$ se décompose de façon naturelle en le produit cartésien : $\mathfrak{S}_n = \mathrm{T} \times (\mathfrak{S}_{(1)} \times \mathfrak{S}_{(2)} \ldots \mathfrak{S}_{(k)})$. On peut ainsi définir le produit extérieur $f_1 \wedge f_2 \wedge \ldots f_k$ de k applications multilinéaires alternées $f_1 f_2, \ldots f_k$ de E dans des espaces vectoriels $\mathrm{F}_1, \mathrm{F}_2, \ldots \mathrm{F}_k$ comme l'application multilinéaire alternée obtenue par l'antisymétrisation « transverse » par les éléments de T, du produit tensoriel de $f_1, f_2, \ldots f_k$ soit :

$$(f_1 \wedge f_2 \wedge \ldots f_k)(x_1, x_2, \ldots x_n) = \sum_{\tau \in \mathrm{T}} \varepsilon(\tau) f_1(x_{\tau 1}, \ldots x_{\tau p_1})$$
$$\otimes (f_2(x_{\tau(p_1+1)}, \ldots x_{\tau(p_1+p_2)}) \otimes \ldots f_k(x_{\tau(p_1+\ldots p_{k-1}+1)} \ldots x_{\tau n})$$

élément de $\mathrm{F}_1 \otimes \mathrm{F}_2 \otimes \ldots \mathrm{F}_k$.

Exemple : Produit extérieur de n formes linéaires $\alpha_1, \alpha_2, \ldots \alpha_n$ sur E. Dans ce cas $\mathrm{T} \equiv \mathfrak{S}_n$, et on retrouve, en respectant la notation antérieure

$$\overline{\wedge} : (\alpha_1 \overline{\wedge} \alpha_2 \overline{\wedge} \ldots \alpha_n)(x_1, x_2, \ldots x_n) = \sum_{\tau \in \mathfrak{S}n} \varepsilon(\tau) \alpha_1(x_{\tau 1}) \alpha_2(x_{\tau 2}) \ldots \alpha_n(x_{\tau n}) = \mathrm{dét}|\alpha_i(x_j)|.$$

Proposition IV.8.A.

Le produit extérieur d'applications multilinéaires alternées est associatif.

Preuve : Soient f, g, h, des applications de E dans des espaces vectoriels F, G, H respectivement p, q et r-linéaires alternées et soit $n = p + q + r$. Nous allons montrer que $(f \wedge g) \wedge h$ est identique à $f \wedge g \wedge h$. On a :

$$((f \wedge g) \wedge h)(x_1, x_2, \ldots x_n) = \sum_{t \in \mathrm{T}_{p+q,r}} \varepsilon(t)(f \wedge g)(x_{t1}, x_{t2}, \ldots x_{t(p+q)}) \otimes h(x_{t(p+q+1)}, \ldots x_{tn}).$$

En posant $x_{tj} = y_j$, on a :

$$(f \wedge g)(y_1, \ldots y_{p+q}) = \sum_{\tau \in \mathrm{T}_{p,q}} \varepsilon(\tau) f(y_{\tau 1}, \ldots y_{\tau p}) \otimes g(y_{\tau(p+1)}, \ldots y_{\tau(p+q)}).$$

En prolongeant la permutation $\tau \in \mathrm{T}_{p,q} \subset \mathfrak{S}_{p+q}$ par l'identité sur $(p + q + 1), \ldots n)$, on plonge $\mathrm{T}_{p,q}$ dans $\mathfrak{S}_n$ et

$$(f \wedge g)(x_{t1}, \ldots x_{t(p+q)}) = \sum_{\tau \in \mathrm{T}_{p,q}} \varepsilon(\tau) f(x_{t\tau 1}, \ldots x_{t\tau p}) \otimes g(x_{t\tau(p+1)}, \ldots x_{t\tau(p+q)})$$

ce qui donne l'expression de $((f \wedge g) \wedge h)(x_1, x_2, \ldots x_n)$:

$$\sum_{t \in \mathrm{T}_{p+q,r}, \tau \in \mathrm{T}_{p,q}} \varepsilon(t\tau) f(x_{t\tau 1}, \ldots) \otimes g(x_{t\tau(p+1)}, \ldots) \otimes h(x_{t\tau(p+q+1)}, \ldots)$$

identique à l'expression de $(f \wedge g \wedge h)(x_1 x_2, \ldots x_n)$:

$$\sum_{\tau \in \mathrm{T}_{p,q,r}} \varepsilon(u) f(x_{u1}, \ldots) \otimes g(x_{u(p+1)}, \ldots) \Phi h(x_{u(p+q+1)}, \ldots)$$

puisque toute permutation $u \in \mathrm{T}_{p,q,r}$ peut s'écrire d'une façon et d'une seule sous la forme $u = t\tau$ avec $t \in \mathrm{T}_{p+q,r}$ et $\tau \in \mathrm{T}_{p,q}$. Il suffit en effet de placer les termes de la réunion des suites $(u_1 < u_2 < \ldots u_p)$ et $(u(p + 1) < \ldots < u(p + q))$ dans l'ordre naturel : $uj_1 < uj_2 < \ldots uj_{p+q}$ pour obtenir la suite qui doit obligatoirement s'identifier à $t_1 < t_2 < \ldots t_{(p+q)}$ ce qui détermine t.

La permutation τ telle que $u = t\tau$ est alors uniquement déterminée par les équations : $\tau j_1 = 1, \tau j_2 = 2, \ldots, \tau j_{p+q} = p + q$.

Vérifions que τ appartient bien à l'ensemble $\mathrm{T}_{p,q}$: si $j < j' \leqslant p$, on a $uj < uj'$ donc $t\tau j < t\tau j'$ ce qui entraîne $\tau j < \tau j'$. De la même façon, si $p < j < j'$, on a $uj < uj'$ donc $t\tau j < t\tau j'$ soit $\tau j < \tau j'$. De la même façon, on montre que $f \wedge (g \wedge h) = f \wedge g \wedge h$.

Proposition IV.8.B.

Le produit extérieur de formes multilinéaires alternées est anticommutatif : si f est une p-forme et g une q-forme sur E, $g \wedge f = (-1)^{pq} f \wedge g$.

Preuve : la permutation Π :

$$(1, 2, \ldots, p+q) \rightarrow (p+1, \ldots p+q, 1, 2, \ldots p)$$

a pour signature $(-1)^{pq}$ et détermine une correspondance biunivoque : $\tau' \in \mathrm{T}_{q,p} \rightarrow \tau = \tau'\Pi \in \mathrm{T}_{p,q}$. D'où :

$$\begin{aligned}
&(f \wedge g)(x_1, \ldots x_{p+q}) \\
&= \sum_{\tau \in \mathrm{T}p,q} \varepsilon(\tau) f(x_{\tau 1}, \ldots x_{\tau p}) g(x_{\tau(p+1)}, \ldots x_{\tau(p+q)}) \\
&= \sum_{\tau' \in \mathrm{T}q,p} \varepsilon(\tau) f(x_{\tau'\Pi 1}, \ldots x_{\tau'\Pi p}) g(x_{\tau'\Pi(p+1)} \cdots x_{\tau'\Pi(p+q)}) \\
&= (-1)^{pq} \sum_{\tau' \in \mathrm{T}q,p} \varepsilon(\tau') g(x_{\tau' 1}, \ldots x_{\tau' q}) f(x_{\tau'(p+1)}, \ldots x_{\tau'(p+q)}) \\
&= (-1)^{pq} (g \wedge f)(x_1, x_2, \ldots x_{p+q}).
\end{aligned}$$

Des considérations analogues permettent de montrer l'associativité et la commutativité du produit symétrique d'applications multilinéaires symétriques.

Il reste à vérifier que lorsqu'une p-forme u et une q-forme v sur E appartiennent aux sous-espaces $\wedge^p \mathrm{E}^*$ de $(\wedge^q \mathrm{E})^*$ et $\wedge^q \mathrm{E}^*$ de $(\wedge^q \mathrm{E})^*$ (c'est-à-dire sont « de type fini »), leurs produits extérieurs en tant qu'éléments des puissances extérieures ou en tant que formes coïncident. Il suffit de le vérifier pour des éléments décomposables $\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p$ et $\alpha_{p+1} \wedge \alpha_{p+2} \wedge \ldots \alpha_{p+q}$ où les $\alpha_j \in \mathrm{E}^*$, en montrant que le produit $\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_{p+q}$ est bien la $(p+q)$-forme sur E obtenue par symétrisation transverse du produit tensoriel de $\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge_p$ et $\alpha_{pn} \wedge \alpha_{p+2} \ldots \alpha_{p+q}$. Or c'est exactement ce qu'exprime la formule de Laplace du § IV.7.

IV.9 – Algèbres extérieures et symétriques

Soit E un espace vectoriel sur le corps K, de dimension quelconque. Convenons que $\wedge^\circ \mathrm{E} = \mathrm{K}$ et que $\wedge^1 \mathrm{E} = \mathrm{E}$, et notons $\wedge \mathrm{E}$ la somme directe de toutes les puissances extérieures de E : $\wedge \mathrm{E} = \bigoplus_{p=0}^{\infty} \wedge^p \mathrm{E} = \mathrm{K} \oplus \mathrm{E} \oplus \wedge^2 \mathrm{E} \oplus \ldots$

Un élément u de $\wedge \mathrm{E}$ s'écrit donc d'une façon et d'une seule comme une somme finie Σu_k où $u_k \in \wedge^k \mathrm{E}$ est appelé sa composante de degré k. u est homogène de degré p s'il n'a qu'une seule composante non nulle $u_p \in \wedge^p \mathrm{E}$. Le produit extérieur, qui a été défini entre éléments homogènes, s'étend par linéarité et détermine donc sur $\wedge \mathrm{E}$ une structure d'algèbre associative avec unité (celle de $\mathrm{K} = \wedge^\circ \mathrm{E}$) : si $u = \Sigma u_j$, $v = \Sigma v_k$, $u \wedge v = \sum_{j,k} u_j \wedge v_k$ $= \sum_m (\sum_{j+k=m} u_j \wedge v_k)$.

L'algèbre $\wedge \mathrm{E}$ est appelé *l'algèbre extérieure de* l'espace vectoriel E. C'est une algèbre graduée anticommutative.

On peut donner de $\wedge \mathrm{E}$ une définition globale, évitant la construction par les sous-espaces homogènes. En effet une construction analogue a déjà été faite pour l'algèbre tensorielle : $\otimes \mathrm{E} = \bigoplus_{p=0}^{\infty} (\otimes^p \mathrm{E})$, algèbre associative universelle sur E. On peut alors définir simplement l'algèbre extérieure de E comme le quotient $(\otimes \mathrm{E})/_\mathrm{L}$ de l'algèbre $\otimes \mathrm{E}$ par l'idéal bilatère L engendré par l'ensemble des éléments $\{x \otimes x ; \forall x \in \mathrm{E}\}$ (c'est un idéal homogène).

Nous allons prouver l'isomorphisme canonique de $\wedge \mathrm{E} = \oplus(\wedge^p \mathrm{E})$ et de $(\otimes \mathrm{E})/_\mathrm{L}$ en démontrant que ces deux algèbres sont solutions d'un même problème universel.

Convenons de dire qu'une application linéaire f de E dans une algèbre associative est *alternée* lorsque $f(x)^2 = 0$, $\forall x \in \mathrm{E}$.

Un couple (j, A) formé d'une algèbre associative unitaire A sur le corps K et d'une application linéaire alternée j de E dans A est dit *universel* pour les applications alternées si quelle que soit l'application linéaire alternée f de E dans une algèbre associative unitaire B, elle se factorise par j suivie d'un homomorphisme d'algèbres unitaires uniquement déterminé F, de A dans B :

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "j"'] & & B \\ & A \arrow[ur, "F"'] & \end{tikzcd} \qquad f = F \circ j \; ; \; F \text{ unique}$$

(La condition F unique peut être remplacée ici par : $j(\mathrm{E})$ et 1_{A} engrendrent A).

a) le couple formé de l'inclusion de $\mathrm{E} \equiv \wedge^1\mathrm{E}$ dans $\wedge\mathrm{E}$ possède cette propriété. En effet, f étant donnée, définissons F sur $\wedge^\circ\mathrm{E}$ par $\mathrm{F}(1_{\wedge\mathrm{E}}) = 1_{\mathrm{B}}$, sur $\wedge^1\mathrm{E}$ par $\mathrm{F} = f$. Sur $\wedge^p\mathrm{E}$, F est obligatoirement égal à $\overline{\wedge}^p f$ sur les éléments décomposables, donc uniquement déterminé. F respecte la multiplication sur les éléments décomposables et c'est donc un homomorphisme d'algèbres unitaires de $\wedge\mathrm{E}$ dans B.

b) le couple formé de l'application j composée de l'inclusion de E dans $\otimes\mathrm{E}$ et de la projection de $\otimes\mathrm{E}$ sur son quotient $(\otimes\mathrm{E})/\mathrm{L}$ possède également la propriété universelle. En effet f s'étend en un homomorphisme d'algèbre unitaire Φ de $\otimes\mathrm{E}$ dans B :

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "\text{inclusion } \subset"'] & & B \\ & \otimes E \arrow[ur, "\Phi"'] & \end{tikzcd} \qquad \Phi \text{ unique}$$

et puisque $\Phi(x \otimes x) = f(x) \cdot f(x) = 0$, Φ est nulle sur L donc passe au quotient pour définir F, c.q.f.d.

Remarque : on voit aisément que l'intersection de l'idéal bilatère L de $\otimes\mathrm{E}$ avec $\otimes^p\mathrm{E}$ n'est autre que le sous-espace d'antisymétrisation L_p. La propriété universelle de l'algèbre extérieure sur un espace vectoriel entraîne immédiatement les propriétés suivantes :

1) Une application linéaire f d'un espace vectoriel E dans un espace vectoriel F sur le même corps K détermine un homomorphisme d'algèbres unitaires graduées $f^\wedge$ de $\wedge\mathrm{E}$ dans $\wedge\mathrm{F}$. En effet, par composition de f et de l'inclusion de F dans $\wedge\mathrm{F}$ on obtient une application linéaire de E dans $\wedge\mathrm{E}$ telle que $f(x)\wedge f(x) = 0$, $\forall x \in \mathrm{E}$.

2) Si H est un troisième espace vectoriel sur K et si h est une application linéaire de F dans H, on a $(h \circ f)^{\vee} = h^{\vee} \circ f^{\vee}$. En particulier, si f est inversible, en prenant $\mathrm{H} = \mathrm{E}$ et $h = \overset{-1}{f}$, on obtient : $(\overset{-1}{f})^{\vee} = (\overset{-1}{f^{\vee}})$ autrement dit si f est un isomorphisme, $f^{\vee}$ l'est aussi.

Si F est un sous-espace de E, l'injection i de F dans E se prolonge en un homomorphisme $i^{\wedge}$ de $\wedge$F dans $\wedge$E. En prenant une base de E dont une partie est une base de F, on voit que $\hat{\imath}$ est une injection de l'algèbre extérieure de F : $\wedge$F dans $\wedge$E. Par une construction analogue, la somme directe des espaces de formes multilinéaires alternées sur E, ou p-formes : $\overset{\infty}{\underset{p=0}{\oplus}} (\wedge^p \mathrm{E})^*$ (où $(\wedge^{\circ}\mathrm{E})^* = \mathrm{K}$ et $(\wedge^1\mathrm{E})^* = \mathrm{E}^*$) est, avec le produit extérieur des p-formes étendu par linéarité, une algèbre associative avec unité, graduée, anticommutative. C'est l'*algèbre des formes multilinéaires alternées sur* E.

3) l'application f de E dans $\wedge$E qui a $x \in \mathrm{E}$ fait correspondre $f(x) = -x$ est alternée : $f(x) \wedge f(x) = 0$ donc se prolonge à un homomorphisme F d'algèbres unitaires de $\wedge$E dans elle-même. Mais puisque $f \circ f$ est l'identité de E, $\mathrm{F} \circ \mathrm{F}$ est l'identité de $\wedge$E. F est donc un *automorphisme involutif* qui transforme un élément u homogène de degré p en $(-1)^p u$, et qui est appelé *l'automorphisme principal* de $\wedge$E.

Les éléments invariants forment la sous-algèbre paire de $\wedge$E : $\overset{\infty}{\underset{p=0}{\oplus}} (\wedge^{2r}\mathrm{E})$.

Lorsque E est de dimension finie n, $\wedge^p\mathrm{E} = 0$ pour $p > n$ et $\dim \wedge^p \mathrm{E} = \binom{n}{p}$. L'algèbre extérieure $\wedge$E est alors de dimension finie égale au cardinal de l'ensemble des parties de $(1, 2, \dots n)$, soit $\sum_{p=0}^{n} \binom{n}{p} = 2^n$. L'espace vectoriel des p-formes alternées sur E s'identifie alors à la puissance extérieure p-ème du dual de E : $\wedge^p\mathrm{E}^*$ et l'algèbre des formes multilinéaires alternées sur E à l'algèbre extérieure $\wedge\mathrm{E}^*$.

L'analogie entre puissances extérieures et symétriques persiste dans la construction des algèbres. Dans la somme directe $\vee\mathrm{E} =$

$\overset{\infty}{\underset{p=0}{\oplus}}(\vee^p E)$, où $\vee^\circ E = K$ et $\vee^1 E = E$, le produit symétrique entre éléments homogènes s'étend par linéarité. Il y définit une structure d'algèbre associative avec unité, graduée, commutative : $\vee E$ est *l'algèbre symétrique sur* E.

On peut l'appeler aussi l'*algèbre des polynômes sur l'espace vectoriel* E.

Ici également, on peut donner de l'algèbre symétrique sur E une définition globale comme quotient de l'algèbre tensorielle $\otimes E$ par l'idéal bilatère M (homogène) engendré par les éléments $\{x \otimes y - y \otimes x;\ \forall x, y \in E\}$ qui assurent la symétrisation du quotient. L'isomorphisme canonique de $\vee E = \oplus(\vee^p E)$ et de $\otimes E/M$ peut être démontré par le fait que ces deux algèbres sont solutions du même problème universel.

Convenons de dire qu'une application linéaire f de E dans une algèbre associative est *commutative* lorsque $f(x)f(y) - f(y)f(x) = 0$, $\forall x, y \in E$, autrement dit lorsque les images des vecteurs de E commutent entre elles.

Un couple (j, S) formé d'une algèbre associative unitaire *commutative* sur le corps K et d'une application linéaire j de E dans S est dit *universel* pour les applications commutatives de E si quelle que soit l'application linéaire commutative f de E dans une algèbre associative unitaire B, elle se factorise par j suivi d'un homomorphisme d'algèbres unitaires uniquement déterminé F, de S dans B :

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "j"'] & & B \\ & S \arrow[ur, "F"'] & \end{tikzcd} \qquad f = F \circ j\,;\ F \text{ unique}$$

(F unique peut être remplacé ici par : 1_S et $j(E)$ engendrent S).

Des démonstrations analogues aux précédentes montrent alors que $\vee E$ et $\otimes E/M$ sont universelles pour les applications commutatives de E, donc canoniquement isomorphes.

On peut d'ailleurs vérifier que l'intersection de l'idéal M et de $\otimes^p E$ est identique au sous-espace de symétrisation Mp.

Une application linéaire f de E dans F détermine un homomorphisme d'algèbres unitaires $f^\vee$ de $\vee E$ dans $\vee F$ qui est un isomorphisme si f en est un.

Les formes multilinéaires symétriques peuvent également être rassemblées en un algèbre, associative, avec unité graduée et commutative, en formant la somme directe des espaces de formes p-linéaires symétriques sur E : $\bigoplus\limits_{p=0}^{\infty} (\vee^p \mathrm{E})^*$. Lorsque la dimension de E est finie, $\vee^p \mathrm{E}^*$ s'identifie à $(\vee^p \mathrm{E})^*$ et l'algèbre des formes multilinéaires symétriques à l'algèbre symétrique $\vee \mathrm{E}^*$ sur le dual de E.

Mais la grande différence entre $\wedge \mathrm{E}$ et $\vee \mathrm{E}$, c'est que l'algèbre symétrique est toujours de dimension infinie pour $\mathrm{E} \neq 0$.

IV.10 – Diverses manifestations de la dualité entre $\wedge \mathrm{E}$ et $\wedge \mathrm{E}^*$ lorsque E est de dimension finie. Produits intérieurs

Pour éviter toute hésitation dans l'écriture des diverses formules qui vont suivre, nous convenons chaque fois qu'apparaissent dans une expression des éléments de $\wedge \mathrm{E}$ et de $\wedge \mathrm{E}^*$, d'écrire à gauche ceux de $\wedge \mathrm{E}$, à droite ceux de $\wedge \mathrm{E}^*$. Pour représenter les produits intérieurs, il est d'usage d'employer les notations « équerre » : $\lrcorner$ et $\llcorner$. Si par exemple $u \in \wedge \mathrm{E}$ et $\varphi \in \wedge \mathrm{E}^*$, nous définissons ci-dessous $u \lrcorner \varphi$ et $u \llcorner \varphi$. Il est donc indispensable dans ces notations de pouvoir reconnaître lequel des éléments est considéré comme opérant sur l'autre. Un moyen mnémotechnique consiste à considérer la branche horizontale comme une flèche : $u \rightarrow\!\lrcorner\, \varphi$, u opère sur φ, $u \llcorner\!\leftarrow \varphi$ φ opère sur u.

Rappelons que E étant de dimension finie, $(\wedge^p \mathrm{E})^* = \wedge^p \mathrm{E}^*$, $\forall p$, d'où il résulte que $(\wedge \mathrm{E})^* = \wedge \mathrm{E}^*$.

Définition IV.10.A. Produits intérieurs

Soient E un espace vectoriel de dimension finie et $\langle\rangle$ le crochet de dualité entre $\wedge \mathrm{E}$ et $\wedge \mathrm{E}^*$. Soient $u \in \wedge \mathrm{E}$, $\varphi \in \wedge \mathrm{E}^*$. Le produit intérieur par u : $u \lrcorner$ dans $\wedge \mathrm{E}^*$ est l'opérateur linéaire de $\wedge \mathrm{E}^*$, opérant *à gauche* sur $\wedge \mathrm{E}^*$, transposé de la multiplication à droite par u dans $\wedge \mathrm{E}$:

$$\langle z \wedge u, \omega \rangle = \langle z, u \lrcorner\ \omega \rangle, \ \forall z \in \wedge \mathrm{E},\ \omega \in \wedge \mathrm{E}^*.$$

Le produit intérieur par φ : $\llcorner \varphi$ dans $\wedge \mathrm{E}$ est l'opérateur linéaire de $\wedge \mathrm{E}$, que l'on fait opérer à droite, sur $\wedge \mathrm{E}$, transposé de la multiplication à gauche par φ dans $\wedge \mathrm{E}^*$:

$$\langle z \llcorner \varphi, \omega \rangle = \langle z, \varphi \wedge \omega \rangle, \ \forall z \in \wedge \mathrm{E}, \ \omega \in \wedge \mathrm{E}^*.$$

Remarques :

1) on trouve parfois dans la littérature des interversions de facteurs dans les formules définissant les produits intérieurs, ce qui revient à les considérer comme les transposés des multiplications à gauche dans $\wedge \mathrm{E}$, où à droite dans $\wedge \mathrm{E}^*$. Ces transposés ont naturellement parfaitement le droit à l'existence, et de toutes façons lorsqu'on n'envisage que des éléments homogènes, ils ne diffèrent des autres que par des signes.

2) puisque les éléments décomposables engendrent linéairement $\wedge \mathrm{E}$ et $\wedge \mathrm{E}^*$, il suffit de prendre pour éléments quelconques z et ω dans les définitions des produits intérieurs des éléments décomposables, ou simplement homogènes.

3) On peut aussi observer que, puisque le crochet de dualité entre éléments homogènes de degrés différents de $\wedge \mathrm{E}$ et $\wedge \mathrm{E}^*$ est nul, le produit intérieur par des éléments homogènes u ou φ de degré r ou bien est nul ou bien diminue le degré des éléments homogènes sur lesquels ils opèrent de r. En effet si $u \lrcorner \omega$ où $z \llcorner \varphi$ n'est pas nul, on doit avoir d'après les égalités ci-dessus : $d^\circ z + d^\circ u = d^\circ \omega$ et $d^\circ(u \lrcorner \omega) = d^\circ z = d^\circ \omega - d^\circ u$; de même $d^\circ z = d^\circ \varphi + d^\circ \omega$ et $d^\circ(z \llcorner \varphi) = d^\circ \omega = d^\circ z - d^\circ \varphi$. Dès lors : $u \lrcorner (\wedge^p \mathrm{E}^*) = 0$ si $p < d^\circ u$ et $(\wedge^p \mathrm{E}) \llcorner \varphi = 0$ si $p < d^\circ \varphi$.

Les définitions abstraites des produits intérieurs ci-dessus ne nous renseignent pas immédiatement sur leur signification. Observons d'abord que $u \lrcorner \varphi$ et $u \llcorner \varphi$ sont bilinéaires en (u, φ), et que l'associativité des algèbres $\wedge \mathrm{E}$ et $\wedge \mathrm{E}^*$ se manifeste *grâce aux conventions choisies*, par :

$$\begin{aligned}
\langle z \wedge u \wedge v, \omega \rangle = \langle z \wedge u, v \lrcorner \omega \rangle &= \langle z, u \lrcorner (v \lrcorner \omega) \rangle \\
&= \langle z, (u \wedge v) \lrcorner \omega) \rangle \\
\langle z, \alpha \wedge \beta \wedge \omega \rangle = \langle z \llcorner \alpha, \beta \wedge \omega \rangle &= \langle (z \llcorner \alpha) \llcorner \beta, \omega \rangle \\
&= \langle z \llcorner (\alpha \wedge \beta), \omega \rangle
\end{aligned}$$

d'où : $\boxed{(u \wedge v) \lrcorner \omega = u \lrcorner (v \lrcorner \omega) \text{ et } z \llcorner (\alpha \wedge \beta) = (z \llcorner \alpha) \llcorner \beta}$.

Ces égalités s'interprètent de la façon suivante : le produit intérieur définit sur $\wedge E^*$ une structure de module à gauche sur l'algèbre $\wedge E$, et sur $\wedge E$ une structure de module à droite sur l'algèbre $\wedge E^*$.

Une conséquence est que le produit intérieur par un élément décomposable, par exemple $\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p \in \wedge^p E^*$ est le composé des produits intérieurs par les facteurs $\alpha_1, \alpha_2, \ldots \alpha_p \in E^*$.

Une autre, que le carré d'un produit intérieur $u \lrcorner\ (u \lrcorner\ \omega)$ ou $(z \llcorner \varphi) \llcorner \varphi$ par un élément u ou φ homogène et de degré impair est toujours nul, en particulier, pour $x \in E$ et $\alpha \in E^*$, $x \lrcorner\ x \lrcorner\ = 0$ et $\llcorner\, \alpha \llcorner\, \alpha = 0$.

Le calcul du produit intérieur d'un élément décomposable par un élément décomposable consiste simplement à écrire les développements de Laplace de la fin du § IV.6. sous une forme légèrement différente :

$$\begin{aligned}
&\langle x_1 \wedge x_2 \wedge \ldots x_{p+q};\ \alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_{p+q}\rangle = \text{dét}|\langle x_i; \alpha_j\rangle| \\
&\quad = \langle (x_1 \wedge x_2 \wedge \ldots x_{p+q}) \llcorner (\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p);\ \alpha_{p+1} \wedge \ldots \alpha_{p+q}\rangle \\
&\quad = \langle x_1 \wedge x_2 \wedge \ldots x_p;\ (x_{p+1} \wedge \ldots x_{p+q}) \lrcorner\ (\alpha_1 + \alpha_2 \wedge \ldots \alpha_{p+q})\rangle \\
&\quad = \langle (\sum_{\tau \in T_{p,q}} \varepsilon(\tau) \langle x_{\tau 1} \wedge \ldots x_{\tau p};\ \alpha_1 \wedge \ldots \alpha_p\rangle x_{\tau(p+1)} \wedge \ldots x_{\tau(p+q)});\\
&\alpha_{p+1} \wedge \ldots \alpha_{\tau p})\rangle \\
&\qquad = \langle x_1 \wedge x_2 \wedge \ldots x_p; (\sum_{\tau \in T_{p,q}} \varepsilon(\tau) \langle x_{p+1} \wedge \ldots x_{p+q}; \\
&\qquad\qquad \alpha_{\tau(p+1)} \wedge \ldots \alpha_{\tau(p+q)}\rangle \alpha_{\tau 1} \wedge \ldots \alpha_{\tau p})\rangle
\end{aligned}$$

et l'on obtient ainsi en notant $y_j = x_{p+j}$, $j = 1, 2, \ldots, q$:

$$\boxed{\begin{aligned}
&(x_1 \wedge x_2 \wedge \ldots x_{p+q}) \llcorner (\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p) \\
&= \sum_{\tau \in T_{p,q}} \varepsilon(\tau) \langle x_{\tau 1} \wedge \ldots x_{\tau p};\ \alpha_1 \wedge \ldots \alpha_p\rangle x_{\tau(p+1)} \wedge \ldots x_{\tau(p+q)} \\
&(y_1 \wedge y_2 \wedge \ldots y_p) \lrcorner\ (\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_{p+q}) \\
&\quad = \sum_{\tau \in T_{p,q}} \varepsilon(\tau) \langle y_1 \wedge \ldots y_q;\ \alpha_{\tau(p+1)} \wedge \ldots \alpha_{\tau_{p+q}}\rangle \alpha_{\tau 1} \wedge \ldots \alpha_{\tau p}
\end{aligned}},$$

qui montrent que le produit intérieur d'un élément homogène de degré $(p+q)$ par un élément homogène de degré p de l'algèbre duale revient à une contraction partielle sur p indices distribuée dans les $(p+q)$ indices par antisymétrisation.

Si $e = (e_1, e_2, \dots e_n)$ est une base de E, e^* la base duale de E^*, $\{e_\mathrm{I}\}$ et $\{e^{*\mathrm{I}}\}$ les bases associées de $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ où I parcourt l'ensemble des suites strictement croissantes de $(1, 2, \dots n)$, les formules précédentes, ou une application directe des définitions donnent les formules suivantes.

$$\begin{cases} e_\mathrm{I} \llcorner\, e^{*\mathrm{J}} = 0 \text{ si J n'est pas contenu dans I} \\ e_\mathrm{I} \llcorner\, e^{*\mathrm{J}} = \varepsilon(\mathrm{J}) e_{\mathrm{J}'} \text{ si} \\ \mathrm{J} = \{j_1 < j_2 < \dots j_p\} \subset \mathrm{I} = \{i_1 < i_2 < \dots i_{p+q}\} \end{cases}$$

où $\mathrm{J}' = \mathrm{I} - \mathrm{J}$ et $\varepsilon(\mathrm{J})$ est la signature de la permutation $\tau \in \mathrm{T}_{q,p}$ qui applique $(1, 2, \dots p)$ sur J et $(p+1, \dots p+q)$ sur la suite $\mathrm{J}' = \mathrm{I} - \mathrm{J}$ complémentaire de J dans I. Soit maintenant $\mathrm{J}' = \{k_1 < k_2 < \dots k_q\}$.

On a :

$$\begin{cases} e_{\mathrm{J}'} \lrcorner\; e^{*\mathrm{I}} = 0 \text{ si J}' \text{ n'est pas contenu dans I} \\ e_{\mathrm{J}'} \lrcorner\; e^{*\mathrm{I}} = \varepsilon(\mathrm{J}) e^{*\mathrm{J}} \text{ si } \mathrm{J}' \subset \mathrm{I} \text{ avec } \mathrm{J} = \mathrm{I} - \mathrm{J}' \end{cases}$$

La différence des signes apparaissent dans les formules qui expriment $e_\mathrm{I} \llcorner\, e^{*\mathrm{J}}$ et $e_{\mathrm{J}'} \lrcorner\; e^{*\mathrm{I}}$ provient naturellement du fait que les produits intérieurs $\llcorner$ et $\lrcorner$ sont les transposés, l'un d'une translation à gauche, l'autre d'une translation à droite.

Ces formules montrent également que le produit intérieur ($\llcorner$ ou $\lrcorner$) de deux éléments décomposables de $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ est un élément décomposable.

Cas particuliers :

1) *produit intérieur par un vecteur ou un covecteur*

Remarquons d'abord que le corps K, étant son propre dual, on a, si $x \in \mathrm{E}$ et $\alpha \in \mathrm{E}^*$:

$$\langle x \llcorner\, \alpha, 1\rangle = \langle x, \alpha\rangle \text{ d'où } x \llcorner\, \alpha = \langle x, \alpha\rangle 1 \in \wedge\mathrm{E}$$
$$\langle 1, x \lrcorner\; \alpha\rangle = \langle x, \alpha\rangle \text{ d'où } x \lrcorner\; \alpha = \langle x, \alpha\rangle 1 \in \wedge\mathrm{E}^*$$

Une permutation τ de $\mathrm{T}_{1,q}$ est entièrement déterminée par l'image $\tau 1 = m$, avec $\varepsilon(\tau) = (-1)^{m-1}$ tandis que la suite complémentaire est $(1, 2, \dots \widehat{m}, \dots q+1)$. En notant $i(\alpha)$ le produit intérieur par $\alpha \in \mathrm{E}^*$ on obtient :

$$i(\alpha)(x_1 \wedge x_2 \wedge \ldots x_{q+1}) = (x_1 \wedge x_2 \wedge \ldots x_{q+1}) \llcorner \alpha$$
$$\sum_{m=1}^{q+1} (-1)^{m-1} \langle x_m, \alpha \rangle x_1 \wedge \ldots \widehat{x}_m \wedge \ldots x_{q+1}$$

Nous verrons au § IV.11 que l'opérateur $i(\alpha)$ sur $\wedge\mathrm{E}$ est une dérivation de l'algèbre $\wedge\mathrm{E}$ (on dit aussi « antidérivation »).

C'est un opérateur de carré nul : $i_\alpha^2 = 0$, et en écrivant que $i(\alpha+\beta)^2 = 0$, on voit que $i(\beta)i(\alpha) = -i(\alpha)i(\beta)$.

De la même façon on calcule si $y \in \mathrm{E}$:

$$y \lrcorner (\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_{p+1})$$
$$= (-1)^p \sum_{m=1}^{p+1} (-1)^{m-1} \langle y, \alpha_m \rangle \alpha_1 \wedge \ldots \widehat{\alpha}_m \wedge \ldots \alpha_{p+1}$$

que l'on peut écrire : $(-1)^p i(y)(\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_{p+1})$.

2) *isomorphismes entre les espaces vectoriels* $\wedge^p\mathrm{E}$ et $\wedge^{n-p}\mathrm{E}^*$

Une base $e = (e_1, e_2, \ldots e_n)$ de E détermine une base $e_1 \wedge e_2 \wedge \ldots e_n$ de la droite $\wedge^n\mathrm{E}$ et une base $e^{*1} \wedge e^{*2} \wedge \ldots e^{*n}$ de la droite $\wedge^n\mathrm{E}^*$ *e étant fixé*, on définit une forme bilinéaire B sur $\wedge\mathrm{E}$ par :

$$\mathrm{B}(t,u) = \langle t \wedge u;\ e^{*1} \wedge e^{*2} \wedge \ldots e^{*n} \rangle = \langle t; u \lrcorner (e^{*1} \wedge e^{*2} \wedge \ldots e^{*n}) \rangle$$

dont l'application linéaire *associée à droite* que nous notons H, de $\wedge\mathrm{E}$ dans son dual $(\wedge\mathrm{E})^* = \wedge\mathrm{E}^*$, définie par $\mathrm{B}(t;u) = \langle t, \mathrm{H}(u) \rangle$, s'écrit donc :

$$\mathrm{H}(u) = u \lrcorner (e^{*1} \wedge e^{*2} \wedge \ldots e^{*n}).$$

Proposition IV.10.A.

H est un isomorphisme (B est non-dégénérée).

Preuve : Si u est homogène de degré p, $\mathrm{H}(u)$ est homogène de degré $(n-p)$. Pour chaque $p = 0, 1, 2, \ldots n$, la restriction de H à $\wedge^p\mathrm{E}$ est ainsi une application linéaire de $\wedge^p\mathrm{E}$ dans $\wedge^{n-p}\mathrm{E}^*$. Pour chaque partie $\mathrm{I} = \{i \leqslant i_1 < i_2 < \ldots i_p \leqslant n\}$ de p éléments de

$(1, 2, \dots n)$, de complémentaire I', la formule démontrée ci-dessus donne :

$$\mathrm{H}(e_\mathrm{I}) = e_\mathrm{I} \lrcorner\ (e^{*1} \wedge \dots e^{*n}) = \varepsilon(\mathrm{I}')e^{*\mathrm{I}'}.$$

H détermine donc une bijection aux signes près entre les éléments de la base $\wedge^p e$ de $\wedge^p \mathrm{E}$ et ceux de la base $\wedge^{n-p} e^*$ de $\wedge^{n-p}\mathrm{E}^*$, s'appuyant sur la bijection entre parties de p éléments et parties complémentaires de $(n-p)$ éléments de $(1, 2, \dots n)$. c.q.f.d.

On a : $e_{\mathrm{I}'} \wedge e_\mathrm{I} = \varepsilon(\mathrm{I}')e_1 \wedge e_2 \wedge \dots e_n$ d'où : $\mathrm{B}(e_{\mathrm{I}'}, e_\mathrm{I}) = \varepsilon(\mathrm{I}') = \langle e_{\mathrm{I}'}; e_\mathrm{I} \lrcorner\ (e^{*1} \wedge \dots e^{*n})\rangle$ et on retrouve : $\mathrm{H}(e_\mathrm{I}) = \varepsilon(\mathrm{I}')e^{*\mathrm{I}'}$

On peut définir d'une façon analogue la forme bilinéaire B^* sur $\wedge\mathrm{E}^*$:

$$\begin{aligned}\mathrm{B}^*(\varphi, \omega) = \langle e_1 \wedge e_2 \wedge \dots e_n; \varphi \wedge \omega\rangle &= \langle (e_1 \wedge e_2 \wedge \dots e_n) \llcorner\ \varphi; \omega\rangle \\ &= \widetilde{\mathrm{H}}(\varphi); \omega\rangle.\end{aligned}$$

L'application linéaire de $\wedge\mathrm{E}^*$ dans $\wedge\mathrm{E}$ *associée à gauche* de B^* est l'inverse $\overset{-1}{\mathrm{H}}$ de l'application H il suffit de le vérifier sur les vecteurs des bases de $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ associées à la base e de E.

Puisque $e^{*\mathrm{I}'} \wedge e^{*\mathrm{I}} = \varepsilon(\mathrm{I}')e^{*1} \wedge e^{*2} \wedge \dots e^{*n}$, on a : $\varepsilon(\mathrm{I}') = \langle e_1 \wedge e_2 \wedge \dots e_n;\ e^{*\mathrm{I}'} \wedge e^{*\mathrm{I}}\rangle = \langle \widetilde{\mathrm{H}}(e^{*\mathrm{I}'}; e^{*\mathrm{I}}\rangle$ d'où :

$$\widetilde{\mathrm{H}}(e^{*\mathrm{I}'}) = \varepsilon(\mathrm{I}')e_\mathrm{I} = \overset{-1}{\mathrm{H}}(e^{*\mathrm{I}'}).$$

Remarques :

1) Si dans l'égalité : $(t \wedge u) \lrcorner\ \omega = t \lrcorner\ (u \lrcorner\ \omega)$ on remplace ω par $e^{*1} \wedge e^{*2} \wedge \dots e^{*n}$, on obtient :

$$\mathrm{H}(t \wedge u) = t \lrcorner\ \mathrm{H}(u)$$

soit, si $\mathrm{H}(u) = \varphi : t \lrcorner\ \varphi = \mathrm{H}(t \wedge \overset{-1}{\mathrm{H}}(\varphi))$.

L'application H permet ainsi d'exprimer le produit intérieur à l'aide du produit extérieur.

D'une façon analogue on a : $t \lrcorner\ \varphi = \overset{-1}{\mathrm{H}}(\mathrm{H}(t) \wedge \varphi)$.

2) Puisque les produits intérieurs conservent le caractère « décomposable » (voir ci-dessus), il en résulte que tous les éléments

de $\wedge^{n-1}\mathrm{E}$, ou de $\wedge^{n-1}\mathrm{E}^*$ sont décomposables comme images des éléments de E^* ou de E.

3) Si l'on choisit une autre base de E : $e' = e \cdot \mathrm{S}$, on a $e'_1 \wedge e'_2 \wedge \dots e'_n = (\text{dét}\,\mathrm{S}) \cdot e_1 \wedge e_2 \wedge \dots e_n$ et $e'^{*1} \wedge \dots e'^{*n} = \overset{-1}{(\text{dét}}\ \mathrm{S}) e^{*1} \wedge \dots e^{*n}$ d'où $\mathrm{B}'(t,u) = \langle t \wedge u, e'^{*1} \wedge \dots e'^{*n} \rangle = \overset{-1}{(\text{dét}}\ \mathrm{S}), \mathrm{B}(t,u)$ et $\mathrm{H}'(u) = \overset{-1}{(\text{dét}}\ \mathrm{S})\mathrm{H}(u)$.

4) Si F est un sous-espace de dimension p de E, les formes linéaires nulles sur F forment le sous-espace orthogonal $\mathrm{F}^{\perp}$ de dimension $(n-p)$ de E^*. Toutes les bases de F définissent à un scalaire non nul près le même p-vecteur (Proposition IV.4. et remarque consécutive). La droite de $\wedge^p\mathrm{E}$ qui les porte a pour image par H la droite de $\wedge^{n-p}\mathrm{E}^*$ portant les $(n-p)$-covecteurs des bases de $\mathrm{F}^{\perp}$.

Nous allons maintenant mettre en évidence un autre aspect, cette fois purement tensoriel, des applications H. Développons pour cela leur égalité de définition sur un élément décomposable de $\wedge\mathrm{E}$:

$$
\begin{aligned}
&\langle x_1 \wedge \dots x_p \wedge x_{p+1} \wedge \dots x_n ; e^{*1} \wedge \dots e^{*n} \rangle \\
&\qquad\qquad = \langle x_1 \wedge \dots x_p ;\ \mathrm{H}(x_{p+1} \wedge \dots x_n) \rangle \\
&\quad = \text{dét}|x^i_j| = \langle x_1 \otimes \dots x_p \otimes x_{p+1} \otimes \dots x_n ;\ \mathrm{A}(e^{*1} \otimes \dots e^{*n}) \rangle \\
&\quad = \sum_{\sigma \in \mathfrak{S}_n} \prod_{j=1}^{n} \langle x_j ; e^{*\sigma j} \rangle = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{p} \langle x_j ; e^{*\sigma j} \rangle \prod_{k=p+1}^{n} \langle x_k ; e^{*\sigma k} \rangle) \\
&\quad = \langle x_1 \otimes x_2 \otimes \dots x_p ; \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)(e^{*\sigma 1} \otimes \dots e^{*\sigma p} \otimes e^{*\sigma(p+1)} \dots e^{*\sigma n}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \circledast (x_{p+1} \otimes \dots x_n) \rangle
\end{aligned}
$$

où l'on convient que le signe $\circledast$ signifie la contraction, en respectant l'ordre, des $(n-p)$ *derniers « indices »* de A $(e^{*1} \otimes e^{*2} \otimes \dots e^{*n})$ et de $x_{p+1} \otimes \dots x_n$ (il n'y a *malheureusement pas* de notation standard satisfaisante et universellement admise, pour les contractions !).

On a vu au § II.7 que la correspondance qui à chaque base e de E associe l'élément $\mathrm{A} \cdot (e^{*1} \otimes \dots e^{*n}) = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\sigma) e^{*\sigma 1} \otimes \dots e^{*\sigma n}$ de $\otimes^n\mathrm{E}^*$ définit le pseudotenseur covariant ε_* de Levi-Civita.

D'autre part, puisque les tenseurs décomposables engendrent toutes les puissances tensorielles, on peut étendre par linéarité à $\otimes^{n-p}\mathrm{E}$ l'égalité précédente et l'on obtient ainsi la :

Proposition IV.10.B.

Soit E un espace vectoriel de dimension n. Pour chaque entier $p : 0 \leqslant p \leqslant n$, il existe une application qui associe à chaque tenseur $t \in \otimes^{n-p}\mathrm{E}$ un pseudotenseur covariant antisymétrique d'ordre p, $\mathfrak{H}(t)$ *qui dépend du choix d'une base e de* E par sa variance scalaire (-1). Il est obtenu par contraction des $(n-p)$ indices de t avec les $(n-p)$ derniers indices du pseudotenseur ε_* de Levi-Civita.

$$\boxed{\mathfrak{H}(t)_{i_1 i_2 \dots i_p} = \sum \varepsilon_{i_1 i_2 \dots i_p j_1 j_2 \dots j_{n-p}} t^{j_1 j_2 \dots j_{n-p}}}.$$

Exemple : Le *produit vectoriel* de n vecteurs $x_1, x_2, \dots x_{n-1}$ de E est la pseudoforme linéaire déjà mentionnée au § III.7. :

$$(x_1 \times x_2 \times \dots x_{n-1})_j = \sum \varepsilon_{j i_1 i_2 \dots i_{n-1}} x^{i_1} x^{i_2} \dots x^{i_{n-1}}.$$

Pour $n = 3$, on obtient le produit vectoriel usuel de $x, y \in \mathrm{E}$:

$$\begin{aligned} x \times y &= (\textstyle\sum \varepsilon(\sigma) e^{*\sigma 1} \otimes e^{*\sigma 2} \otimes e^{*\sigma 3}) \circledast^* (x \otimes y) \\ &= e^{*1}(x^2 y^3 - x^3 y^2) + e^{*2}(x^3 y^1 - x^1 y^3) + e^{*3}(x^1 y^2 - x^2 y^1). \end{aligned}$$

Plus généralement, on a l'égalité :

$$\begin{aligned} \mathrm{dét}_e(z, x_1, \dots x_{n-1}) &= \textstyle\sum \varepsilon_{j i_1 i_2 \dots i_{n-1}} z^j x_1^{i_1} \dots x_{i_{n-1}, n-1} \\ &= \langle z ; x_1 \times x_2 \times \dots x_{n-1} \rangle. \end{aligned}$$

IV.11 – Extensions d'une forme bilinéaire aux puissances extérieures et symétriques

Soient E et F deux espaces vectoriels de dimension finie ou infinie sur le même corps K et b une forme bilinéaire sur $\mathrm{E} \times \mathrm{F}$. Rappelons (§ II. 11) que la puissance tensorielle p-ème de b, $\overline{\otimes}^p b$

est la forme bilinéaire sur $(\otimes^p \mathrm{E}) \times (\otimes^p \mathrm{F})$ définie sur les tenseurs décomposables par :

$$\begin{aligned}(\overline{\otimes}^p b)(x_1 \otimes x_2 \otimes \dots x_p;\ y_1 \otimes y_2 \otimes \dots y_p)& \\ = b(x_1, y_1) b(x_2, y_2) \dots b(x_p, y_p).&\end{aligned}$$

L'antisymétrisation, ou la symétrisation de l'un des arguments entraîne automatiquement, par bilinéarité l'antisymétrisation ou la symétrisation de l'autre :

1) Commençons par l'antisymétrisation :

$$\begin{aligned}(\overline{\otimes}^p b)(\mathrm{A} \cdot x_1 \otimes \dots x_n); y_1 \otimes \dots y_n)& \\ = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b(x_{\overset{-1}{\sigma} 1}; y_1) \dots b(x_{\overset{-1}{\sigma} p}, y_p)&\end{aligned}$$

ce qui s'écrit en rangeant les produits des $b(x_{\overset{-1}{\sigma} j}, y_j)$ suivant l'ordre croissant des indices des vecteurs x_k :

$$\begin{aligned}&= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b(x_1, y_{\sigma 1}) \dots b(x_n, y_{\sigma n}) \\ &= (\overline{\otimes}^p b)(x_1 \otimes \dots x_n;\ \mathrm{A}(y_1 \otimes \dots y_n)) = \text{dét}\, |b(x_i, y_i)|\end{aligned}$$

La fonction obtenue est donc alternée par rapport aux x_j comme par rapport aux y_k, et, s'annulant ainsi sur les sous-espaces d'antisymétrisation de $\otimes^p \mathrm{E}$ et $\otimes^p \mathrm{F}$, passe au quotient (théorème IV.4.B) en définissant une forme bilinéaire sur les espaces quotients $\wedge^p \mathrm{E}$ et $\wedge^p \mathrm{F}$ appelé *puissance extérieure p-ème de b*, notée $\overline{\wedge}^p b$.

La valeur de $\overline{\wedge}^p b$ sur un couple de p-vecteurs décomposables est donc :

$$(\overline{\wedge}^p b)(x_1 \wedge x_2 \wedge \dots x_p;\ y_1 \wedge y_2 \wedge \dots y_p) = \text{dét}\, |b(x_i; y_j)|.$$

Si ρ est l'application linéaire, associé à droite, de F dans E^*, et γ l'application linéaire associée à gauche de E dans F^*, définies par : $b(x; y) = \langle x, \rho(y) \rangle = \langle \gamma(x), y \rangle$, les puissances extérieures p-èmes $\overline{\wedge}^p \rho$ et $\overline{\wedge}^p \gamma$ sont définies sur les éléments décomposables par :

$$\begin{aligned}(\overline{\wedge}^p \rho)(y_1 \wedge \dots y_p) &= \rho(y_1) \wedge \dots \rho(y_p) \text{ et} \\ (\overline{\wedge}^p \gamma)(x_1 \wedge \dots x_p) &= \gamma(x_1) \wedge \dots \gamma(x_p)\end{aligned}$$

d'où il résulte, vue l'expression de $\overline{\wedge}^p b$ sur ces éléments, que $\overline{\wedge}^p \rho$ et $\overline{\wedge}^p \gamma$ sont les applications linéaires associées, à droite et à gauche de $\overline{\wedge}^p b$, à valeurs respectivement dans $\wedge^p \mathrm{E}^* \subset (\wedge^p \mathrm{E})^*$ et dans $\wedge^p \mathrm{F}^* \subset (\wedge^p \mathrm{F})^*$.

Supposons maintenant *finies* les dimensions de E et de F. La propriété pour b d'être une forme bilinéaire régulière (ou non-dégénérée) est équivalente à la propriété pour ρ (ou $\gamma =^t \rho$) d'être un isomorphisme. Mais si ρ est un isomorphisme, toutes ses puissances extérieures $\overline{\wedge}^p \rho$ sont aussi des isomorphismes. Il en résulte que toutes les puissances $\overline{\wedge}^p b$ sont régulières. De plus puisque $\overline{\wedge}^p(\overset{-1}{\rho}) = (\overset{-1}{\overline{\wedge}^{\rho}} \rho)$, on a $(\overset{-1}{\overline{\wedge}^{p}} b) = \overline{\wedge}^p(\overset{-1}{b})$.

Lorsque les dimensions de E et F sont finies, on peut considérer la forme bilinéaire b comme un élément de $\mathrm{E}^* \otimes \mathrm{F}^*$ et l'écrire

$$b = \sum_{k=1}^{m} \alpha_k \otimes \beta_k. \text{ Alors } b(x,y) = \sum_{k=1}^{m} \langle x; \alpha_k \rangle \langle \beta_k; y \rangle$$

$$\rho(y) = \sum_{k=1}^{m} \alpha_k \langle \beta_k, y \rangle \text{ et } \gamma(x) = \sum_{k=1}^{m} \langle x; \alpha_k \rangle \beta_k.$$

Le rang de b, qui est celui de ρ ou de γ est donc le plus grand entier r tel que $\overline{\wedge}^r b$, ou, ce qui revient au même, $\overline{\wedge}^r \rho$ ou $\overline{\wedge}\gamma$, soit non nul. Les expressions précédentes de ρ et de γ montrent, que le nombre m de termes $\alpha_k \otimes \beta_k$ nécessaires pour exprimer b est au moins égal à r.

Sur des p-vecteurs décomposables, on a :

$$\begin{aligned}(\overline{\wedge}^p b)(x_1 \wedge x_2 \wedge \dots x_p;\ y_1 \wedge y_2 \wedge \dots y_p) &= \text{dét} \left| \sum_{k=1}^{m} \langle x_i; \alpha_k \rangle \langle \beta_k; y_j \rangle \right| \\ &= \text{dét}\, \mathrm{AB}\end{aligned}$$

où A est la matrice à p lignes et m colonnes $|\langle x_i; \alpha_k \rangle|$ et B la matrice à m lignes et p colonnes $|\langle b_k; y_j \rangle|$. Si $\overline{\wedge}^p b \neq 0$, $p \leqslant r \leqslant m$.

La formule de Lagrange (§ IV.6) exprime le déterminant du produit des deux matrices A, B comme somme des produits des mineurs d'ordre p des matrices A et B, soit :

$$\begin{aligned}\text{dét}\, \mathrm{AB} = \sum_{\mathrm{I}} &\langle x_1 \wedge x_2 \wedge \dots x_p;\ \alpha_{i_1} \wedge \alpha_{i_2} \wedge \dots \alpha_{i_p} \rangle \\ &\langle \beta_{i_1} \wedge \beta_{i_2} \wedge \dots \beta_{i_p};\ y_1 \wedge y_2 \wedge \dots y_p \rangle\end{aligned}$$

sommation étendue aux suites $\mathrm{I} = \{1 \leqslant i_1 < i_2 < \dots i_p \leqslant m\}$.

On en déduit l'expression de $\overline{\wedge}^p b$ comme élément du produit tensoriel

$$(\wedge^p E^*) \otimes (\wedge^p F^*) : \overline{\wedge}^p b = \sum_{I} (\alpha_{i_1} \wedge \alpha_{i_2} \wedge \ldots \alpha_{i_p}) \otimes (\beta_{i_1} \wedge \beta_{i_2} \wedge \ldots \beta_{i_p}).$$

Si $e = (e_i)$ et $f = (f_j)$ sont des bases de E et F, la matrice de b relative à ces bases est $B = |b_{ij}| = |b(e_i, f_j)|$. La matrice de $\overline{\wedge}^p b$ relative aux bases $\wedge^p e$ et $\wedge^p f$ a pour éléments $(\overline{\wedge}^p b)_{I,J} = (\overline{\wedge}^p b)(e_I; f_J) = \underset{\substack{i_q \in I \\ j_r \in J}}{\text{dét}} |b(e_{i_q}; f_{j_r})| =$ mineur B_{IJ} de la matrice $|b_{ij}|$ où $I = \{1 \leqslant i_1 < i_2 < \ldots i_p \leqslant \dim E\}$ et $J = \{1 \leqslant j_1 < j_2 < \ldots j_p \leqslant \dim F\}$.

Les éléments de la matrice de $\overline{\wedge}^p b$ sont donc les mineurs d'ordre p de la matrice de b.

Le cas le plus important est naturellement celui d'une forme bilinéaire b sur un espace vectoriel E de dimension finie n.

Si $e = (e_1, e_2, \ldots e_n)$ est une base de E, $(\overline{\wedge}^n b)(e_1 \wedge \ldots e_n; e_1 \wedge \ldots e_n) = \text{dét}|b(e_i, e_j)|$ est le déterminant de b dans la base e et $\overline{\wedge}^n b = \text{dét}_e b \cdot (e^{*1} \wedge \ldots e^{*n} \otimes e^{*1} \wedge \ldots e^{*n})$.

Si e diagonalise b, les bases associées $\wedge^p e$ à e dans les $\wedge^p E$ diagonalisent $\overline{\wedge}^p b$.

D'autre part si b est symétrique, toutes ses puissances extérieures sont symétriques, si b est antisymétrique, les puissances extérieures impaires $\overline{\wedge}^{2r+1} b$ le sont aussi, alors que les puissances extérieures paires sont symétriques.

2) Si S est maintenant l'opérateur de symétrisation $\sum_{\sigma \in \mathfrak{S}_p} \sigma$, on a pour chaque p :

$$(\overline{\otimes} b)(S \cdot (x_1 \otimes \ldots x_p); y_1 \otimes \ldots y_p) = \sum_{\sigma \in \mathfrak{S}_p} b(x_{\sigma^{-1} 1}, y_1) \ldots b(x_{\sigma^{-1} p}, y_p)$$

$$= \sum_{\sigma \in \mathfrak{S}_p} b(x_1, y_{\sigma 1}) \ldots b(x_p, y_{\sigma p}) = \overline{\otimes}^p b(x_1 \otimes \ldots x_p; S \cdot (y_1 \otimes \ldots y_p))$$

$$= \text{per}|b(x_i, y_j)|.$$

Cette fonction étant symétrique aussi bien par rapport aux x_i que par rapport aux y_j s'annule sur les sous-espaces de symétrisation de $\otimes^p E$ et $\otimes^p F$. Elle passe donc au quotient d'après le théorème IV.4.B et définit une forme bilinéaire sur le couple $\vee^p E$, $\vee^p F$ des puissances symétriques p-èmes de E et de F appelée puissance symétrique p-ème de b et notée $\overline{\vee}^p b$. Les puissances symétriques p-ème $\overline{\vee}^p \rho$ et $\overline{\vee}^p \gamma$ des applications linéaires associées à droite et à gauche de b sont les applications associées de $\overline{\vee}^p b$ puisque par exemple :

$$\begin{aligned}\langle x_1 \vee \ldots x_p;\ (\overline{\vee}^p \rho)(y_1 \vee \ldots y_p)\rangle &= \langle x_1 \vee \ldots x_p;\ \rho(y_1) \vee \ldots \rho(y_p)\rangle \\ &= \mathrm{per}|\langle x_i; \rho(y_j)\rangle| = \mathrm{per}|b(x_i; y_j)|.\end{aligned}$$

IV.12 – Bivecteurs et Plaffiens

Dans tout ce paragraphe, E est un espace vectoriel de dimension finie n.

Proposition IV.12.

Pour tout bivecteur non nul u de E, on peut trouver $2p$ vecteurs linéairement indépendants : $a_1, a_2, \ldots, a_{2p}$ tels que :

1) $u = a_1 \wedge a_2 + a_3 \wedge a_4 + \ldots a_{2p-1} \wedge a_{2p}$. Une telle écriture de u est dite *réduite*.

2) l'entier p est caractérisé par la propriété d'être le plus grand entier tel que la puissance extérieure p-ème de u : u^p soit non nulle ; $2p$ est le rang de u.

3) u^p est un $2p$-vecteur décomposable et $a_1, a_2, \ldots a_{2p}$ forment une base du sous-espace F de E, de dimension $2p$, associé à u^p.

Preuve : dans une base quelconque e de E, u est une combinaison linéaire des $\{e_i \wedge e_j;\ i < j\}$ qui forment la base associée $\wedge^2 e$ de $\wedge^2$E. En renumérotant les vecteurs de e, on peut supposer que le coefficient de $e_1 \wedge e_2$ dans l'expression de u est non nul, et en remplaçant éventuellement e_1 par λe_1, que ce coefficient est égal à 1.

On peut alors écrire u :

$$\begin{aligned}u &= e_1 \wedge e_2 + e_1 \wedge (\sum_{j>3} \alpha_j e_j) + e_2 \wedge (\sum_{k>3} \beta_k e_k) + \sum_{p,q>3} \lambda_{pq} e_p \wedge e_q \\ &= (e_1 - \sum_{k>3} \beta_k e_k) \wedge (e_2 + \sum_{j>3} \alpha_j e_j) + \sum_{p,q>3} \mu_{pq} e_p \wedge e_q.\end{aligned}$$

Les vecteurs $\{a_1 = e_1 - \sum_{k>3} \beta_k e_k;\ a_2 = e_2 + \sum_{j>3} \alpha_j e_j;\ e_3 \dots e_n\}$ forment une base de e, et $u = a_1 \wedge a_2 + \sum_{p,q>3} \mu_{pq} e_p \wedge e_q$.

Si l'un au moins des coefficients $\mu_{p,q}$ n'est pas nul, le bivecteur non nul $u_1 = \sum_{p,q>3} \mu_{pq} e_p \wedge e_q$ appartient à la sous-algèbre $\wedge \mathrm{E}_1$ de $\wedge\mathrm{E}$ engendrée par le sous-espace E_1 de base $(e_3, e_4, \dots e_n)$, et on peut lui appliquer le même traitement qu'à u. En poursuivant ainsi, on obtient finalement une base de E : $\{a_1, a_2, \dots a_{2p}, e_{2p+1}, \dots e_n\}$ dans laquelle u s'écrit :

$$u = a_1 \wedge a_2 + a_3 \wedge a_4 + \dots a_{2p-1} \wedge a_{2p}.$$

Puisque les éléments de degré pair de $\wedge\mathrm{E}$ commutent et que le carré d'un multivecteur décomposable est nul, le calcul de la puissance extérieure p-ème de u donne : $u^p = p!\ a_1 \wedge a_2 \wedge a_3 \wedge \dots a_{2p-1} \wedge a_{2p}$ qui est non nul. On a alors $u^{p+1} = 0$ et l'affirmation 3) de la proposition résulte de la proposition IV.4.

Définition IV.12.A.

Soit E un espace vectoriel de dimension paire $n = 2r$. On appelle pfaffien d'un bivecteur u de E relativement à une base e, le scalaire $\mathrm{P}f_e(u)$ défini par l'égalité :

$$u^r = r!\mathrm{P}f_e(u) e_1 \wedge e_2 \wedge \dots e_{2r}$$

$\mathrm{P}f_e(u)$ est non nul si et seulement si u est de rang maximum $n = 2r$. Dans ce cas, en écrivant u : $u = a_1 \wedge a_2 + a_3 \wedge a_4 + \dots a_{2r-1} \wedge a_{2r}$ où $(a_1, a_2, \dots a_{2r})$ forment une base de E, on peut définir $\mathrm{P}f_e(u)$ par :

$$a_1 \wedge a_2 \wedge \dots \wedge a_{2r} = \mathrm{P}f_e(u) \cdot e_1 \wedge e_2 \wedge \dots e_{2r}$$

On a : $\mathrm{P}f_e(u) = \mathrm{dét}_e(a_1, a_2, \dots a_{2r})$ qui vérifie la condition ci-dessus.

Calculons $\mathrm{P}f_e(u)$ en fonction des composantes u_{ij} de u dans la base e. On a :

$$u = \sum_{i<j} u_{ij} e_i \wedge e_j$$

d'où $\quad u^r = \sum u_{i_1 j_1} u_{i_2 j_2} \dots u_{i_r j_r} e_{i_1} \wedge e_{j_1} \wedge \dots e_{i_r} \wedge e_{j_r}.$

Seuls sont non nuls dans la somme du second membre les termes pour lesquels l'ensemble des paires d'indices ordonnés $(i_k < j_k)$ qui y figurent forment une permutation de $(1, 2, \dots 2r)$. Soit $\mathfrak{P}$ l'ensemble dont les éléments sont les ensembles P de r paires d'indices ordonnés $(i < j)$ dont la réunion est une permutation de $(1, 2, \dots 2r)$. Si l'on range les paires $(i < j)$ appartenant à un même P en une suite, on obtient une permutation de $(1, 2, \dots 2r)$ dont la signature, que l'on peut noter $\varepsilon(\mathrm{P})$, ne *dépend pas* de l'ordre que l'on a choisi pour les paires $(i < j)$ dans la suite. En rétablissant l'ordre naturel des vecteurs de la base e dans l'expression de u^r, on obtient :

$$\boxed{\mathrm{P}f_e(u) = \sum_{\mathrm{P}\in\mathfrak{P}} \varepsilon(\mathrm{P}) u_{i_1 j_1} u_{i_2 j_2} \dots u_{i_r j_r}}$$

où P est la suite des paires $(i_1 j_1, \dots i_r j_r)$.

Cette expression ressemble à celle d'un déterminant, mais c'est en fait une sorte de « semi-déterminant ».

Par un changement de bases, $e' = e\mathrm{S}$, le pfaffien se transforme ainsi :

$$\begin{aligned} \mathrm{P}f_{e'}(u) = \text{dét}_{e'}(a_1, a_2, \dots a_{2r}) &= \overset{-1}{\text{dét}}\,\mathrm{S} \cdot \text{dét}_e(a_1, \dots a_{2r}) \\ &= \overset{-1}{\text{dét}}\,\mathrm{S} \cdot \mathrm{P}f_e(u) \end{aligned}$$

Si g est un automorphisme de E : $\mathrm{P}f_e(\overline{\wedge}^2 g \cdot u) = \text{dét}\, g \cdot \mathrm{P}f_e(u)$.
Examinons maintenant le point de vue dual. Soient E un espace vectoriel de dimension n et b une forme bilinéaire antisymétrique sur E. Si $b \neq 0$, il existe $x, y \in \mathrm{E}$ tels que $b(x, y) = \alpha \neq 0$. Soit P_1 le plan qu'ils engendrent : $a_1 = \dfrac{x}{\alpha}$ et $a_2 = y$ en forment une base telle que $b(a_1, a_2) = 1$ et E est somme directe orthogonale (relativement à b!) de P_1 et du sous-espace $\mathrm{E}_1 = \mathrm{P}_1^{\perp}$. Par récurrence on construit ainsi une suite d'un nombre pair $2p$ de vecteurs que l'on complète en une base $a = (a_1, a_2, \dots a_n)$ de E.

Désignons par $(\alpha^1, \alpha^2, \dots \alpha^n)$ la base de E^* duale de la base a. Dans cette base, l'élément $\widetilde{b}$ de $\otimes^2\mathrm{E}^*$ représentant b s'écrit donc sous la forme réduite :

$$\begin{aligned} \widetilde{b} = \alpha^1 \otimes \alpha^2 - \alpha^2 \otimes \alpha^1 + \alpha^3 \otimes \alpha^4 - \alpha^4 \otimes \alpha^3 \\ + \dots \alpha^{2p-1} \otimes \alpha^{2p} - \alpha^{2p} \otimes \alpha^{2p-1} \end{aligned}$$

$2p$ est le rang de b comme de $\widetilde{b}$.

Mais b, étant alternée, $\widetilde{b}$ est en fait une forme linéaire sur le quotient $\wedge^2\mathrm{E}$ de $\otimes^2\mathrm{E}$ et s'identifie au bi-covecteur β de $(\wedge^2\mathrm{E})^* = \wedge^2\mathrm{E}^*$ écrit sous forme réduite : $\beta = \alpha^1 \wedge \alpha^2 + \alpha^3 \wedge \alpha^4 + \dots \alpha^{2p-1} \wedge \alpha^{2p}$.

Dans une base quelconque e de E, $\widetilde{b}$ s'écrit :

$$\widetilde{b} = \sum_{(i,j)} b_{ij} e^{*i} \otimes e^{*j} \text{ avec } b_{ij} = -b_{ij} \text{ et } \beta = \sum_{(i<j)} b_{ij} e^{*i} \wedge e^{*j}.$$

Définition IV.12.B.

Soient E un espace vectoriel de dimension paire $n = 2r$, b une forme bilinéaire alternée sur E représentée par l'élément $\widetilde{b} = \sum_i b_{ij} e^{*i} \otimes e^{*j}$ de $\otimes^2\mathrm{E}^*$ dans une base e de E. *Le Pfaffien* $\mathrm{P}f_e(b)$ *de* b relativement à la base e est par définition le pfaffien du bi-covecteur $\beta = \sum_{i<j} b_{ij} e^{*i} \wedge e^{*j}$ associé à b, relativement à e. On l'appelle aussi *pfaffien de la matrice antisymétrique* $\mathrm{B} = |b_{ij}|$ des composantes $b_{ij} = b(e_i, e_j)$ de b dans $e \cdot \mathrm{P}f_e(b)$ n'est différent de zéro que si le rang de b est $n = 2r$.

Le calcul précédent donne, avec les notations définies ci-dessus :

$$\mathrm{P}f_e(b) = \mathrm{P}f(\mathrm{B}) = \sum_{\mathrm{P} \in \mathfrak{P}} \varepsilon(\mathrm{P}) b_{i_1 j_1} b_{i_2 j_2} \dots b_{i_r j_r}.$$

En supposant le rang de b égal à $n = 2r$, reprenons les écritures réduites de $\widetilde{b}$ et de β :

$$\begin{aligned}\widetilde{b} =& \alpha^1 \otimes \alpha^2 - \alpha^2 \otimes \alpha^1 + \alpha^3 \otimes \alpha^4 - \alpha^4 \otimes \alpha^3 \\ &+ \dots \alpha^{2r-1} \otimes \alpha^{2r} - \alpha^{2r} \otimes \alpha^{2r-1} \\ \beta =& \alpha^1 \wedge \alpha^2 + \alpha^3 \wedge \alpha^4 + \dots \alpha^{2r-1} \wedge \alpha^{2r}\end{aligned}$$

avec $\alpha^1 \wedge \alpha^2 \wedge \dots \alpha^{2r} = \mathrm{P}f_e(\beta) e^{*1} \wedge e^{*2} \wedge \dots e^{*2r}$.

Mais on a vu au paragraphe précédent IV.11 que la puissance extérieure $(2r)$-ème de b s'écrit comme élément de $(\wedge^{2r}\mathrm{E}^*) \otimes (\wedge^{2r}\mathrm{E}^*)$: $(\overline{\wedge}^{2r} b = (\alpha^1 \wedge \alpha^2 \wedge \dots \alpha^{2r}) \otimes (\alpha^2 \wedge (-\alpha^1) \wedge \alpha^4 \wedge (-\alpha^3) \wedge \dots \alpha^{2r} \wedge (\alpha^{2r-1})$ soit, puisque $\alpha^{2k} \wedge (-\alpha^{2k-1}) = \alpha^{2k-1} \wedge \alpha^{2k}$: $(\overline{\wedge}^{2r} b) = (\alpha^1 \wedge \alpha^2 \wedge \dots \alpha^{2r}) \otimes (\alpha^1 \wedge \alpha^2 \dots \alpha^{2r}) = \mathrm{P}f_e(b)^2 (e^{*1} \wedge e^{*2} \wedge$

$\dots e^{*2r}) \otimes (e^{*1} \wedge e^{*2} \wedge e^{*2r}) = \text{dét}_e\, be^{*1} \wedge \dots e^{*2r} \otimes e^{*1} \wedge \dots e^{*2r}$. D'où le

Théorème IV.12. Si B est une matrice antisymétrique d'ordre pair, son déterminant est le carré de son pfaffien :

$$\boxed{\text{dét}\, \mathrm{B} = \mathrm{P}f(\mathrm{B})^2}.$$

Exemples : $\text{dét} \begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix} = a^2$

$$\text{dét} \begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = (af - be + cd)^2.$$

IV.13 – Dérivations des algèbres tensorielles, extérieures et symétriques. Invariants. Cohomologie des algèbres de Lie

Nous allons d'abord expliciter la forme la plus générale de la notion de dérivation.

Nous n'aurons pas à l'utiliser dans toute sa généralité, mais elle servira de modèle pour la suite. On considère pour cela *deux* algèbres *quelconques* A et B (pas nécessairement associatives) sur le même corps K, *deux* homomorphismes φ et ψ de A dans B, et un espace K-vectoriel M sur lequel B opère à gauche *et* à droite (bimodule). Dans cette situation une *dérivation* D *de* A *dans* M associée au couple (φ, ψ) est une application linéaire de A dans M liée aux structures multiplicatives par la relation :

$$\mathrm{D}(xy) = \varphi(x) \cdot \mathrm{D}y + \mathrm{D}x \cdot \psi(y).$$

Exemples :

1) Si A et B possèdent des unités et si φ et ψ sont des homomorphismes d'algèbres unitaires, $\mathrm{D}(1.1) = \mathrm{D}1 = 2\mathrm{D}1 = 0$.

2) Le cas le plus simple est celui où $\mathrm{M} = \mathrm{B} = \mathrm{A}$, $\varphi = \psi =$ identité de A : une dérivation « ordinaire » de A est ainsi une

application linéaire de A dans elle-même telle que :

$$\mathrm{D}(ab) = a \cdot \mathrm{D}b + \mathrm{D}a \cdot b$$

Si A est associative, à tout élément a de A correspond une dérivation dite *intérieure*, notée $\operatorname{ad} a$, par : $\operatorname{ad} a(x) = ax - xa = [a, x]$ (crochet de a et de x).

En effet : $[a, xy] = axy - xya = (ax - xa)y + x(ay - ya)$. Observons que si D et D′ sont deux dérivations de A, l'application composée D′D n'est plus une dérivation. Par contre le crochet $[\mathrm{D}', \mathrm{D}] = \mathrm{D}'\mathrm{D} - \mathrm{D}\mathrm{D}'$ en est une, et cette loi de composition sur l'espace vectoriel $\mathfrak{D}(\mathrm{A})$ des dérivations de A en fait une algèbre de Lie (cf III. Définition III.11.B).

Les dérivations intérieures sont stables par le crochet avec D, $\forall \mathrm{D} \in \mathfrak{D}(\mathrm{A}) \cdot [\mathrm{D}, \operatorname{ad} a] = \operatorname{ad} \mathrm{D}a$. Elles forment un idéal de l'algèbre de Lie $\mathfrak{D}(\mathrm{A})$.

3) Une famille importante de dérivations s'obtient en prenant $\mathrm{B} = \mathrm{A}$, ψ = identité de A, φ = automorphisme α de A. L'identité de dérivation s'écrit alors :

$$\mathrm{D}(xy) = \mathrm{D}x \cdot y + \alpha(x) \cdot \mathrm{D}y.$$

Le cas particulier suivant, où $\mathrm{B} = \mathrm{M} = \mathrm{A}$ est une algèbre associative graduée permettra de cataloguer toutes les dérivations usuelles des algèbres extérieures :

Définition IV.13

Soient A une algèbre associative graduée par les entiers positifs, J l'automorphisme involutif principal qui consiste à multiplier les éléments homogènes de degré p par $(-1)^p$. Une J-dérivation de A de degré r est une dérivation homogène de degré r associée à la puissance r-ème de J : si x est homogène de degré p, $\mathrm{D}x$ est homogène de degré $p + r$, et quels que soient $x, y \in \mathrm{A}$, on a :

$$\mathrm{D}(xy) = \mathrm{D}x \cdot y + \mathrm{J}^r x \cdot \mathrm{D}y.$$

Si x est homogène, cette égalité s'écrit :

$$\mathrm{D}(xy) = \mathrm{D}x \cdot y + (-1)^{r \deg x} x \cdot \mathrm{D}y.$$

Si r est *pair*, une J-dérivation est donc tout simplement une dérivation homogène. Si r est *impair* une J-dérivation est parfois appelée une antidérivation, et l'on a alors pour x homogène :

$$\mathrm{D}(xy) = \mathrm{D}x \cdot y + (-1)^{\deg x} x \cdot \mathrm{D}y.$$

L'intérêt d'avoir une seule notion, de J-dérivation, rassemblant dérivations et antidérivations homogènes est la propriété suivante : si D et D$'$ sont deux J-dérivations de degrés r et r', le « crochet gradué » de D et D$'$: $\mathrm{DD}' - (-1)^{rr'}\mathrm{D}'\mathrm{D}$ est une J-dérivation de degré $(r + r')$.

4) Si $\mathrm{B} = \mathrm{A}$ et si M est un A-bimodule, tout élément m de M détermine une dérivation D_m, dite intérieure, de A dans M, par : $\mathrm{D}_m(x) = xm - mx$. On a bien : $\mathrm{D}_m(xy) = xym - mxy = x(ym - my) + (xm - mx)y$.

5) Si M$'$ est un autre A-bimodule et si f est un morphisme de M dans M$'$, c'est-à-dire une application linéaire telle que, quels que soient $a, b \in \mathrm{A}$, $f(a \cdot m \cdot b) = a \cdot f(m) \cdot b$, alors pour toute dérivation D de A dans M, $f\mathrm{D}$ est une dérivation de A dans M$'$:

$$(f\mathrm{D})(ab) = f(a \cdot \mathrm{D}b + \mathrm{D}a \cdot b) = a \cdot f\mathrm{D}d + f\mathrm{D}a \cdot b.$$

Si p est un homomorphisme d'une algèbre A$'$ dans A, l'application composée de p et d'une dérivation D de A dans M est une dérivation de A$'$ dans M, associée au couple $(\varphi \circ p, \psi \circ p)$. En effet, si $a, b \in \mathrm{A}'$:

$$\mathrm{D}p(ab) = \mathrm{D}\{(pa)(pb)\} = \varphi(pa)\mathrm{D}(pb) + \mathrm{D}(pa)\psi(pb).$$

Les dérivations peuvent donc se composer avec les morphismes.

6) Supposons maintenant l'algèbre A associative avec unité. La multiplication de A : $\mathrm{A} \times \mathrm{A} \to \mathrm{A}$ étant bilinéaire détermine une application linéaire $\mu : \mathrm{A} \otimes \mathrm{A} \to \mathrm{A}$ avec $\mu(x \otimes y) = xy$. Mais μ est un morphisme de bimodules : $\mu(a \cdot x \otimes y \cdot b) = \mu(ax \otimes yb) = a(xy)b = a\mu(x \otimes y)b$ et son noyau N est donc un sous-bimodule de $\mathrm{A} \otimes \mathrm{A}$.

L'application D_0 de A dans N :

$$x \in \mathrm{A} \to \mathrm{D}_0\mathrm{A} = x \otimes 1 - 1 \otimes x$$

est alors une dérivation puisque :

$$\mathrm{D}_0(xy) = xy \otimes 1 - 1 \otimes xy = x \cdot (y \otimes 1 - 1 \otimes y) + (x \otimes 1 - \otimes x)y$$

D_0 est une *dérivation universelle* pour les dérivations de A dans les A-bimodules : si D est une telle dérivation de A dans M, la restriction à N de l'application f de $A \otimes A$ dans M définie par $f(x \otimes y) = Dx \cdot y$ est un morphisme de bimodules et $D = f \cdot D_0$.

Lorsque les algèbres A et B sont *associatives* une récurrence immédiate donne la dérivée du produit de m éléments de A :

$$D(x_1 x_2 \ldots x_m) = \sum_{r=1}^{m} \varphi(x_1 x_2 \ldots x_{r-1}) Dx_r \cdot \psi(x_{r+1} \ldots x_m).$$

D'autre part, le problème de l'existence et de la construction des dérivations est simplifié par l'observation suivante.

Les matrices triangulaires :

$$\begin{vmatrix} b_1 & m \\ 0 & b_2 \end{vmatrix} \text{ avec } b_1, b_2 \in B \text{ et } m \in M$$

forment une algèbre associative $\widetilde{M}$, avec unité si B possède une unité. En effet, on a :

$$\begin{vmatrix} b_1 & m \\ 0 & b_2 \end{vmatrix} \begin{vmatrix} b_1' & m' \\ 0 & b_2' \end{vmatrix} = \begin{vmatrix} b_1 b_1' & b_1 m' + m b_2' \\ 0 & b_2 b_2' \end{vmatrix}$$

tandis que l'associativité résulte de l'égalité :

$$b_1 b_1' m'' + (b_1 m' + m b_2') b_2'' = b_1 (b_1' m'' + m' b_2'') + m b_2' b_2''.$$

Dès lors, il y a identité entre dérivations D dans M associées à (φ, ψ) et homomorphismes $\widetilde{D}$ de A dans $\widetilde{M}$ associés à (φ, ψ). Si

$$\widetilde{D}x = \begin{vmatrix} \varphi x & Dx \\ 0 & \psi^x \end{vmatrix}.$$

$\widetilde{D}$ est linéaire, et c'est un homomorphisme si et seulement si :

$$\widetilde{D}(xy) = \begin{vmatrix} \varphi(xy) & D(xy) \\ 0 & \psi(xy) \end{vmatrix} = \widetilde{D}x \cdot \widetilde{D}y = \begin{vmatrix} \varphi x & Dx \\ 0 & \psi x \end{vmatrix} \cdot \begin{vmatrix} \varphi y & Dy \\ 0 & \psi y \end{vmatrix}$$

soit si et seulement si $D(xy) = \varphi x \cdot Dy + Dx \cdot \psi y$, soit, si D est une dérivation de A dans M associée à (φ, ψ).

Supposons maintenant que A soit l'algèbre tensorielle $\otimes E$ sur l'espace vectoriel E. La propriété universelle de $\otimes E$ affirme que toute application linéaire de E dans une algèbre associative avec unité se prolonge de façon unique à $\otimes$E en un homomorphisme d'algèbres unitaires, tout automorphisme de $\otimes$E s'obtenant de cette façon.

Il n'y a donc aucune contrainte aux dérivations de $\otimes$E qui sont obtenues par prolongement d'une application linéaire arbitraire de E. En particulier tout opérateur linéaire u de E détermine une dérivation D_u de degré zéro des algèbres $\otimes$E et T(E).

Nous allons utiliser cette propriété pour construire les dérivations des algèbres quotients de $\otimes$E : $\wedge$E et $\vee$E, à l'aide de la remarque faite à l'exemple n° 5) ci-dessus.

Soient A une algèbre quotient de l'algèbre tensorielle $\otimes$E par un idéal bilatère N, et D une dérivation de A dans elle-même associée à un couple d'automorphismes (φ, ψ) de A. La composée de D et de la projection : $t \to \bar{t}$ de $\otimes$E sur A est une dérivation $\widetilde{\mathrm{D}}$ de $\otimes$E dans A nulle sur N, et $\widetilde{\mathrm{D}}(t \otimes u) = \varphi(\bar{t}) \cdot \mathrm{D}\overline{u} + \mathrm{D}\bar{t} \cdot \psi(\overline{u})$.

Réciproquement, pour construire les dérivations de A, il suffit de trouver celles des dérivations de $\otimes$E dans A qui s'annulent sur N.

1) *Dérivations de l'algèbre symétrique sur* E : $\vee\mathrm{E} = \otimes\mathrm{E}/\mathrm{M}$.

Pour qu'une dérivation D de $\otimes$E dans $\vee$E s'annule sur l'idéal M il est nécessaire qu'elle s'annule sur les générateurs de M : $\{x \otimes y - y \otimes x;\ \forall x, y \in \mathrm{E}\}$. Or, puisque $\vee$E est *commutative* $\mathrm{D}(x \otimes y - y \otimes x) = \{\varphi(\overline{x}) - \psi(\overline{x})\}\mathrm{D}\overline{y} + \{\psi(\overline{y}) - \varphi(\overline{y}\}\mathrm{D}\overline{x}$.

Il suffit donc que $\psi = \varphi$; et on vérifie alors immédiatement, vue la commutativité de $\vee$E, que D est nulle sur tous les éléments $\{t \otimes (x \otimes y - y \otimes x) \otimes u\}$ générateurs linéaires de M. En particulier, pour $\varphi = \psi$ = identité de $\vee$E, on obtient le :

Théorème IV.13.A. Une application linéaire quelconque D de l'espace vectoriel E dans son algèbre symétrique se prolonge de façon unique en une dérivation ordinaire de $\vee$E en elle-même.

Si $\{e_i;\ i \in \mathrm{E}\}$ est une base de E, $\vee$E s'identifie à l'algèbre des polynômes en les e_i : combinaisons linéaires finies de monômes $e_{r_1,i_1} e_{r_2,i_2} \ldots e_{r_k,i_k}$. Une application linéaire de E dans $\vee$E est entièrement déterminée par ses valeurs, qui peuvent être arbitraires, sur les vecteurs de base e_i.

La dérivation D_i de $\vee$E obtenue par prolongement de l'application linéaire de E dans $\vee$E : $\mathrm{D}_i e_i = 1$, $\mathrm{D}_i e_j = 0$ si $j \neq i$ est appelée la dérivation partielle par rapport à e_i. Si E est de dimension finie n

et si $\{e_1, e_2, \ldots e_n\}$ est une base de E, toute dérivation D de $\vee$E est une combinaison linéaire des dérivées partielles D_i à coefficients a_i dans $\vee$E : $\mathrm{D} = \sum_{i=1}^{n} a_i \mathrm{D}_i$.

2) *Dérivations de l'algèbre extérieure sur* E : $\wedge \mathrm{E} = \otimes \mathrm{E}/\mathrm{L}$. Pour qu'une dérivation D de $\otimes$E dans $\wedge$E s'annule sur l'idéal L, il est nécessaire qu'elle s'annule sur les générateurs de L : $\{x \otimes y + y \otimes x;\ \forall x, y \in \mathrm{E}\}$ soit que l'on ait : $\varphi(x) \wedge \mathrm{D}y + \mathrm{D}x \wedge \psi(y) + \varphi(y) \wedge \mathrm{D}x + \mathrm{D}y \wedge \psi(x) = 0$.

Supposons l'application D de E dans $\wedge$E homogène et de degré r soit : $d^\circ \mathrm{D}x = d^\circ x + r = r + 1$, $\forall x \in \mathrm{E}$. Supposons également que ψ soit l'identité de $\wedge$E. L'égalité précédente devient alors :

$$\{\varphi(x) + (-1)^{r+1}x\} \wedge \mathrm{D}y + \{\varphi(y) + (-1)^{r+1}y\} \wedge \mathrm{D}x = 0$$

qui impose : $\varphi(x) = (-1)^r x = \mathrm{J}^r(x)$ où J est l'automorphisme principal de $\wedge$E (voir définition IV.13 ci-dessus) :

L'anticommutativité de $\wedge$E montre alors que D est nulle sur tous les éléments homogènes $\{t \otimes (x \otimes y + y \otimes x) \otimes u\}$ de L, générateurs linéaires de L.

Nous avons donc démontré le :

Théorème IV.13.B. Toute application linéaire homogène de degré r, D de l'espace vectoriel E dans son algèbre extérieure, donc de E dans $\wedge^{r+1}\mathrm{E}$ se prolonge en une J-dérivation de degré r, de l'algèbre $\wedge$E en elle-même suivant la définition IV.13.

On a ainsi pour deux éléments quelconques de $\wedge$E : $\mathrm{D}(t \wedge u) = \mathrm{D}t \wedge u + \mathrm{J}^r(t) \wedge \mathrm{D}u$.

Les deux cas particuliers : $r = -1$, où l'application D de E dans $\wedge^\circ \mathrm{E} = \mathrm{K} \cdot 1$ est une forme linéaire α sur E, et $r = 0$, où l'application D de E dans $\wedge^1 = \mathrm{E} = \mathrm{E}$ est un opérateur linéaire de a de E, sont les plus importants dans les applications.

Dans le premier cas, notons $i(\alpha)$ la dérivation prolongeant α. Sur un p-vecteur décomposable, on a :

$$i(\alpha)(x_1 \wedge x_2 \wedge \ldots x_p) = \sum_{i=1}^{p} (-1)^{i-1} \langle \alpha, x_i \rangle x_1 \wedge x_2 \wedge \ldots \widehat{x}_i \wedge \ldots x_p$$

et $i(\alpha)$ est le produit intérieur par α : $\llcorner\, \alpha$, comme on l'a vu au § IV.10. On a $i(\alpha)^2 = 0$ puisque $(\llcorner\, \alpha)^2 = 0$.

Mais on peut également observer que le crochet gradué : $DD' - (-1)^{rr'}D'D$ de deux J-dérivations de degrés r et r' étant une J-dérivation de degré $r + r'$, $i(\alpha)^2 = \frac{1}{2}(i(\alpha)i(\alpha) - (-1)i(\alpha)i(\alpha))$ est une dérivation de degré (-2) donc nulle sur E, donc identiquement nulle.

De la même façon, chaque vecteur x de E étant une forme linéaire sur E^* définit une application $i(x) : E^* \to K$ par $i(x)\alpha = \langle x, \alpha \rangle$ et se prolonge en une dérivation de degré (-1) de l'algèbre $\wedge E^*$. On a pour une p-forme décomposable :

$$i(x) \cdot \alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p = \sum_{i=1}^{p} (-1)^{i-1} \langle x, \alpha_i \rangle \alpha_1 \wedge \alpha_2 \wedge \ldots \widehat{\alpha}_i \wedge \ldots \alpha_p.$$

Il faut bien prendre garde qu'avec les conventions choisies à la définition IV.10.A (voir le cas particulier 1 qui suit cette définition) on a sur $\wedge^p E^*$: $i(x) = (-1)^{p-1} x \lrcorner$.

L'opérateur $i(x)$ prend une forme simple si l'on interprète la p-forme $f = \alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p$ comme une fonction p-linéaire alternée sur E. En effet, si dans l'expression du crochet de dualité :

$$f(x_1, x_2, \ldots x_p) = \langle \alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p;\ x_1 \wedge x_2 \wedge \ldots x_p \rangle = \text{dét} |\langle \alpha_i, x_j \rangle|$$

on développe le déterminant par rapport à la première colonne (x_1), on obtient :

$$f(x_1, x_2, \ldots x_p) = \sum_{i=1}^{p} (-1)^{i-1} \langle \alpha_i, x_1 \rangle \langle \alpha_1 \wedge \ldots \widehat{\alpha}_i \wedge \alpha_p; x_2 \wedge \ldots x_p \rangle$$

ce qui, comparé à la définition précédente, montre que :

$$\boxed{(i(x)f)(x_1, x_2, \ldots x_{p-1}) = f(x, x_1, x_2, \ldots x_{p-1})}.$$

L'égalité précédente s'écrit aussi :

$$\langle \alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_p;\ x_1 \wedge x_2 \wedge \ldots x_p \rangle = \langle i(x_1) \alpha_1 \wedge \alpha_2 \ldots \alpha_p;\ x_2 \wedge \ldots x_p \rangle$$

et rappelle que $i(x_1)$ est le transposé dans $\wedge E^*$ de l'opérateur $\mu(x_1)$ de multiplication à gauche par x_1 dans $\wedge E$:

$$\boxed{i(x) =^t \mu(x)}$$

Dans le second cas, la J-dérivation D_a de $\wedge\mathrm{E}$, de degré zéro, associée à l'opérateur linéaire a de E est une dérivation ordinaire de l'algèbre $\wedge\mathrm{E}$. Sur un p-vecteur décomposable :

$$\mathrm{D}_a(x_1 \wedge x_2 \wedge \ldots x_p) = \sum_{i=1}^{p} x_1 \wedge \ldots x_{i-1} \wedge ax_i \wedge x_{i+1} \ldots x_p.$$

C'est l'extension de Lie de l'opérateur a à $\wedge^p\mathrm{E}$ (cf § III.14).
Si E est de dimension finie n, la restriction de D_a à $\wedge^n\mathrm{E}$ est l'homothétie ayant pour rapport la trace de a :

$$\begin{aligned}\langle \mathrm{D}_a(e_1 \wedge \ldots e_n); e^{*1} \wedge \ldots e^{*n}\rangle \\ &= \langle \sum_{i=1}^{n} e_1 \wedge \ldots ae_i \wedge \ldots e_n;\ e^{*1} \wedge \ldots e^{*n}\rangle \\ &= \sum_{i=1}^{n} \langle ae_i, e^{*i}\rangle = \sum a_i^i = \mathrm{Tr}\, a.\end{aligned}$$

Si $a, b \in \mathfrak{L}(\mathrm{E})$, le crochet $[\mathrm{D}_a, \mathrm{D}_b] = \mathrm{D}_a\mathrm{D}_b - \mathrm{D}_b\mathrm{D}_a$ est une dérivation de degré zéro de $\wedge\mathrm{E}$, déterminée par sa restriction à E qui est l'opérateur linéaire $[a, b] = ab - ba$. On a donc : $[\mathrm{D}_a, \mathrm{D}_b] = \mathrm{D}_{[a,b]}$. L'application qui à chaque $a \in \mathfrak{L}(\mathrm{E})$ fait correspondre $\mathrm{D}_a \in \mathfrak{L}(\wedge\mathrm{E})$ est donc une représentation de l'algèbre de Lie $\mathfrak{gl}(\mathrm{E})$ des opérateurs linéaires de E dans l'espace vectoriel $\wedge\mathrm{E}$ et plus précisément dans chaque sous-espace homogène $\wedge^p\mathrm{E}$ (cf § IV.3). En utilisant des opérateurs diagonaux dans une base de E, on prouve aisément que les représentations de $\mathfrak{gl}(\mathrm{E})$ dans les $\wedge^p\mathrm{E}$ sont irréductibles.

Application : Invariants des algèbres de Lie

Nous allons examiner très rapidement les algèbres extérieures $\wedge\mathfrak{g}$ et $\wedge\mathfrak{g}^*$ associées à une algèbre de Lie $\mathfrak{g}$, en raison de leur rôle important en algèbre, en géométrie et en physique. Soit $\vee$ un $\mathfrak{g}$-module : espace vectoriel sur le corps K de $\mathfrak{g}$ muni d'une représentation $\rho : \mathfrak{g} \longrightarrow \mathfrak{L}(\vee)$. Nous supposerons $\mathfrak{g}$ et $\vee$ de dimension finie. L'espace vectoriel $\mathrm{C}^p(\mathfrak{g}; \vee) = \mathfrak{L}(\wedge^p\mathfrak{g}; \vee)$ des applications linéaires de $\wedge^p\mathfrak{g}$ dans $\vee$ est appelé l'espace des *cochaînes de degré p de* $\mathfrak{g}$ *à valeurs dans* $\vee$. On peut aussi le définir comme l'espace vectoriel des applications p-linéaires alternées de $\mathfrak{g}$ dans $\vee \cdot \mathrm{C}^p(g; \vee) \equiv \vee \otimes \wedge^p\mathfrak{g}^*$.
En particulier, si $\vee = \mathrm{K}$ et $\rho = 0$, $\mathrm{C}^p(\mathfrak{g}; \mathrm{K}) \equiv \wedge^p\mathfrak{g}^*$. En définissant $\mathrm{C}^\circ(\mathfrak{g}; \vee) = \vee$, la somme directe $\oplus\mathrm{C}^p(\mathfrak{g}; \vee) = \mathrm{C}(\mathfrak{g}; \vee)$

est *l'espace vectoriel gradué des cochaînes de* $\mathfrak{g}$ *à valeurs dans* $\vee$. $\wedge\mathfrak{g}^*$ possède en plus une multiplication : c'est une algèbre associative unitaire.

La représentation ρ de $\mathfrak{g}$ dans $\vee$ et l'extension de la représentation adjointe de $\mathfrak{g}$ dans $\wedge^p\mathfrak{g}$ (ou dans $\wedge^p\mathfrak{g}^*$) déterminent une *représentation* θ de $\mathfrak{g}$ dans $\mathfrak{L}(\wedge^p\mathfrak{g};\vee) \equiv \vee \otimes \wedge^p\mathfrak{g}^*$. Si $a \in \mathfrak{g}$ et $f \in \mathrm{C}^p(\mathfrak{g};\vee)$, $\theta(a)$ est l'opérateur de degré zéro :

$$(\theta(a)f)(x_1, x_2, \dots x_p)$$
$$= \rho(a)\cdot f(x_1, x_2, \dots x_p) - \sum_{j=1}^{p} f(x_1, \dots |a, x_j|, \dots x_p)$$

$\mathfrak{g}$ opère également dans $\mathrm{C}(\mathfrak{g};\vee)$ par les opérateurs $i(a)$ définis ci-dessus, de degré (-1) (donc nuls pour $p = 0$) :

$$(i(a)f)(x_1, x_2, \dots x_{p-1}) = f(a, x_1, x_2, \dots x_{p-1}).$$

Sur $\mathrm{C}(\mathfrak{g};\vee)$, on définit un opérateur homogène de degré $(+1)$, noté d, directement inspiré de l'expression, en fonction du crochet des champs de vecteurs, de la différentielle extérieure, qui sera vue au paragraphe suivant (cf Théorème IV.14.A). d est appelé pour cela *différentielle*. Si $f \in \mathrm{C}^p(\mathfrak{g};\vee)$.

$$\boxed{\begin{aligned}&(df(x_1, x_2, \dots x_{p-1}) = \sum(-1)^{i-1}\rho(x_i)\cdot f(x_1, \dots \widehat{x}_i, \dots x_{p+1})\\&+ \sum_{i<j}(-1)^{i+j} f(|x_i, x_j|, x_1, \dots \widehat{x}_i, \dots \widehat{x}_j, x_{p+1}\end{aligned}}.$$

La différentielle est liée aux opérateurs $\theta(a)$, $i(a)$ par un certain nombre de relations dont la démonstration (par récurrence sur le degré par exemple) est aisée :

$$\begin{cases}\theta(x) = [i(x), d] = i(x)d + di(x)\\ i([x;y]) = \theta(x)i(y) - i(y)\theta(x)\\ d\theta(x) = \theta(x)d\end{cases}$$

Montrons par récurrence que d est de carré nul. Tout d'abord, pour $f \in \mathrm{C}^\circ(\mathfrak{g};\vee) \equiv \vee$, $ddf(x,y) = \rho(x)df(y) - \rho(y)df(x) - df([x,y]) = \rho(x)\rho(y)f - \rho(y)\rho(x)f - \rho([x,y]f = 0$.

Supposons maintenant prouvé $ddf = 0$ pour tout $f \in \mathrm{C}^{p-1}(\mathfrak{g};\vee)$, et soit $f \in \mathrm{C}^p(\mathfrak{g};\vee)$. $i(x)$ commute avec dd en vertu des relations précédentes : $i(x)ddf = ddi(x)f$, qui est nul d'après l'hypothèse de récurrence, et $i(x)ddf = 0$ quel que soit x entraîne $ddf = 0$, c.q.f.d.

Les éléments de l'image de d sont appelés *cobords*. Les éléments du noyau de d sont appelés *cocycles*. Si Z^p est le sous-espace de C^p que forment les cocycles de degré p, il contient $d\mathrm{C}^{p-1}$, puisque $dd = 0$. Le quotient $\mathrm{H}^p(\mathfrak{g}; \mathrm{V}) = \mathrm{Z}^p / d\mathrm{C}^{p-1}$ est l'*espace vectoriel de cohomologie de degré p de* $\mathfrak{g}$ *à coefficients dans* V.

On a $\theta(x)du = d\theta(x)u$: le sous-espace des cobords est stable par la représentation θ de $\mathfrak{g}$. Mais si $dz = 0$, on a $\theta(x)z = di(x)z$: le sous-espace des cocycles est appliqué dans celui des cobords par θ. Dans les quotients H^p, la représentation quotient $\widetilde{\theta}$ de $\mathfrak{g}$ est nulle : les $\mathrm{H}^p(\mathfrak{g}; \mathrm{V})$ sont des invariants naturels.

Considérons plus particulièrement le cas ou $\mathrm{V} = \mathrm{K}$ et $\mathrm{C}(\mathfrak{g}; \mathrm{K}) \equiv \wedge\mathfrak{g}^*$. On peut alors écrire la différentielle :

$$\begin{aligned} &df(x_1, x_2, \dots x_{p+1}) \\ &\quad = \frac{1}{2} \sum_{i,j} (-1)^{i+j} f([x_i, x_j], x_1, \dots \widehat{x}_i, \dots \widehat{x}_j, \dots x_{p+1}) \\ &\quad = \frac{1}{2} \sum_{i} (-1)^{i-1} (\theta(x_i) f(x_1, \dots \widehat{x}_i, \dots x_{p+1}). \end{aligned}$$

Si l'on prend une base $(e_1, \dots e_n)$ de l'espace vectoriel $\mathfrak{g}$ et la base duale $(e^{*1}, \dots e^{*n})$ de $\mathfrak{g}^*$, $x_i = \sum \langle e^{*k}, x_i \rangle e_k$.

$$\begin{aligned} &df(x_1, x_2, \dots x_{p+1}) \\ &\qquad = \frac{1}{2} \sum_{i;k} (-1)^{i-1} \langle e^{*k}, x_i \rangle (\theta(e_k) f)(x_1, \dots \widehat{x}_i, \dots x_{p+1}). \end{aligned}$$

Or, le produit extérieur d'une forme linéaire α et d'une p-forme $g : \alpha \wedge g = \mu(\alpha) g$ s'écrit :

$$(\alpha \wedge g)(y_1, y_2, \dots y_{p+1}) = \sum (-1) i \langle \alpha, y_i \rangle g(y_1, \dots \widehat{y}_i, \dots y_{p+1}).$$

On en déduit une expression de d en fonction des dérivations $\theta(x)$ et des multiplications extérieures $\mu(x)$:

$$df(x_1, x_2, \dots x_{p+1}) = \frac{1}{2} (\sum \mu(e^{*k}) \theta(e_k) f)(x_1 x_2, \dots x_{p+1})$$

$$\boxed{d = \frac{1}{2} \sum \mu(e^{*k}) \theta(e_k)}\,.$$

Cette expression de d permet de démontrer immédiatement les propriétés suivantes :

a) d est une dérivation de degré $+1$ de l'algèbre $\wedge\mathfrak{g}^*$. Si u et v sont homogènes :

$$d(u \wedge v) = \frac{1}{2}\sum e^{*k} \wedge (\theta(e_k)u) \wedge v + u \wedge \theta(e_k)v)$$
$$= \frac{1}{2}\{\sum(e^{*k} \wedge \theta(e_k)u) \wedge v + (-1)^{\deg u} u \wedge \sum(e^{*k} \wedge \theta(e_k)v)\}$$
$$= du \wedge v + (-1)^{\deg u} u \wedge dv.$$

Il en résulte que le produit de deux cocycles est un cocycle, le produit d'un cocycle par un cobord est un cobord.

La multiplication des cocycles passe au quotient et fait de l'espace vectoriel de cohomologie $\mathrm{H}^*(\mathfrak{g}; \mathrm{K})$ une algèbre associative unitaire anticommutative.

b) les invariants de $\wedge\mathfrak{g}^*$ sont des cocycles : si $\theta(x)f = 0$ quel que soit $x \in \mathfrak{g}$, $df = 0$.

Lorsque l'algèbre de Lie $\mathfrak{g}$ est semi-simple, on démontre que toute représentation de dimension finie de $\mathfrak{g}$ est semi-simple.

Cette propriété entraîne que pour chaque p, le $\mathfrak{g}$-module des cocycles Z^p est somme directe du sous-module des bords $d\mathrm{C}^{p-1}$ et d'un sous-espace I^p : $\mathrm{Z}^p = d\mathrm{C}^{p-1} \oplus \mathrm{I}^p$.

Puisque $\theta(x)\mathrm{Z}^p \subset d\mathrm{C}^{p-1}$, $\forall x \in \mathfrak{g}$, $\theta(x)\mathrm{I}^p = 0$, I^p est un sous-espace d'invariants de $\wedge^p\mathfrak{g}^*$, évidemment isomorphe à $\mathrm{H}^p(\mathfrak{g}) = \mathrm{Z}^p/d\mathrm{C}^{p-1}$. Or, tout invariant de $\wedge^p\mathfrak{g}^*$ est un cocycle, donc dans Z^p. Nous allons prouver qu'il est nécessairement dans I^p. En effet, $\mathfrak{g}$ étant semi-simple, C^{p-1} est somme directe : $\mathrm{C}^{p-1} = \mathrm{Z}^{p-1} \oplus \mathrm{B}^{p-1}$ de deux sous $\mathfrak{g}$-modules avec B^{p-1} isomorphe, en tant que $\mathfrak{g}$-module, à $d\mathrm{C}^{p-1}$.

Un invariant de $d\mathrm{C}^{p-1}$ est l'image d'un invariant de B^{p-1}, donc nul puisque tous les invariants de C^{p-1} sont dans Z^{p-1}. Nous avons prouvé le

Théorème IV.13.C. Si $\mathfrak{g}$ est une algèbre de Lie semi-simple de dimension finie, l'algèbre unitaire des invariants de $\wedge\mathfrak{g}^*$ est isomorphe à l'algèbre de cohomologie de $\mathfrak{g}$: $\mathrm{H}^*(\mathfrak{g}; \mathrm{K})$.

La topologie algébrique et la géométrie différentielle (théorie des classes caractéristiques des espaces fibrés) ont mis en évidence une relation étroite entre les invariants antisymétriques de $\mathfrak{g}$ dans

$\wedge\mathfrak{g}^*$, et symétriques, dans $\vee\mathfrak{g}^*$. Or la détermination directe de ces derniers est généralement facile. L'expression de cette relation utilise le produit tensoriel des algèbres extérieure et symétrique sur $\mathfrak{g}^*$, soit $\mathrm{W} = \wedge\mathfrak{g}^* \otimes \vee\mathfrak{g}^*$. C'est une algèbre associative unitaire. Si l'on se réfère à la situation géométrique qui en est l'origine, $\wedge\mathfrak{g}^*$ représente les fibres (formes différentielles extérieures invariantes à gauche sur le groupe de Lie G) et $\vee\mathfrak{g}^*$ la base, (formes différentielles extérieures de degré *pair* sur la base) des espaces fibrés différentiables G-principaux). On définit sur W les structures naturelles suivantes :

1) une *graduation* : $\wedge\mathfrak{g}^*$ et $\vee\mathfrak{g}^*$ sont des algèbres graduées. Sur leur produit tensoriel, on choisit comme bidegré le couple $(p, 2q)$ déterminé par le degré p sur $\wedge\mathfrak{g}^*\otimes 1$ et le degré de $1\otimes\vee\mathfrak{g}^*$ *multiplié par deux*. Le degré total sur W est $m = p + 2q$. W est engendrée par son unité $1 = 1 \otimes 1$, par ses éléments de degré un : $\mathfrak{g}^* \otimes 1$, et de degré deux : $1 \otimes \mathfrak{g}^*$.

2) des *dérivations* de W dans elle-même. Leur construction et la démonstration de leur existence résultent du schéma indiqué au paragraphe précédent : une J-dérivation D de $\mathrm{A}\otimes\mathrm{B}$ est équivalente à un homomorphisme $\widetilde{\mathrm{D}}$ d'algèbres unitaires :

$$\begin{matrix}\widetilde{\mathrm{D}} \\ u \rightarrow\end{matrix}\begin{vmatrix} \mathrm{J}u & \mathrm{D}\mathrm{U} \\ 0 & u \end{vmatrix} \text{ de } \mathrm{A}\otimes\mathrm{B} \text{ dans l'algèbre } \begin{vmatrix} \mathrm{A}\otimes\mathrm{B} & \mathrm{A}\otimes\mathrm{B} \\ 0 & \mathrm{A}\otimes\mathrm{B} \end{vmatrix}$$

et de la propriété universelle du produit tensoriel $\mathrm{A} \otimes \mathrm{B}$ (§ III.13). Ces dérivations sont homogènes et sont entièrement déterminées par leur valeur sur les générateurs :

– dérivations $i(x)$, $\forall x \in \mathfrak{g}$, de degré (-1), nulles sur $1 \otimes \vee\mathfrak{g}^*$:

$$i(x) \cdot \alpha \otimes 1 = (i(x) \cdot \alpha) \otimes 1 = \langle x, \alpha\rangle \otimes 1; \ i(x)(1 \otimes \alpha) = 0, \ \forall\alpha \in \mathfrak{g}^*$$

– dérivations $\theta(x)$, $\forall x \in \mathfrak{g}$, de degré zéro, extension des représentations naturelles θ de $\mathfrak{g}$ dans $\wedge\mathfrak{g}^* \otimes 1$ et $1 \otimes \vee\mathfrak{g}^*$:

$$\theta(x) \cdot \alpha \otimes 1 = (\theta(x)\alpha) \otimes 1 \text{ et } \theta(x) \cdot (1 \otimes \alpha) = 1 \otimes (\theta(x)\alpha), \ \forall\alpha \in \mathfrak{g}^*$$

3) une *différentielle d de degré* $+1$, analogue à la différentielle définie plus haut sur l'algèbre $\wedge\mathfrak{g}^*$, que des raisons géométriques conduisent à définir comme suit sur les générateurs de W :

– $d(\alpha \otimes 1) = (d\alpha) \otimes 1 + 1 \otimes \alpha$

– $d(1 \otimes \alpha)$ déterminé par la condition $i(x)d(1 \otimes \alpha) = 1 \otimes \theta(x)\alpha$ ce qui s'explicite en prenant une base e_j de $\mathfrak{g}$ et la base duale (e^{*j}) de $\mathfrak{g}^*$ par :

$$d(1 \otimes \alpha) = \sum e^{*k} \otimes \theta(e_k)\alpha$$

d se prolonge en une dérivation de degré $+1$ de W, et puisque les $\theta(x)$ sont des dérivations, de $\vee\mathfrak{g}^*$, on vérifie immédiatement que quel que soit $u \in \vee\mathfrak{g}^*$, on a toujours :

$$d(1 \otimes u) = \sum e^{*k} \otimes \theta(e_k)u.$$

Les propriétés suivantes se vérifient sur les générateurs de W :

$$dd = 0;\ \theta(x) = i(x)d + di(x);\ d\theta(x) = \theta(x)d,\ \forall x \in \mathfrak{g}.$$

Une algèbre associative munie d'une différentielle d, dérivation de carré nul : $dd = 0$, est appelée une *algèbre différentielle*.

Tout élément de W s'écrit $\Sigma u_j \otimes 1 + \Sigma v_k \otimes s_k$ avec $d^\circ s_k > 0$, W est ainsi la somme directe de la sous-algèbre $\wedge\mathfrak{g}^* \otimes 1$ et de l'idéal bilatère S engendré par les éléments de degré > 0 de $1 \otimes \mathfrak{g}^*$. Le quotient W/S s'identifie à $\wedge\mathfrak{g}^*$ et on vérifie immédiatement que l'application $\Pi : \mathrm{W} \to \wedge\mathfrak{g}^* \equiv \mathrm{W/S}$ commute avec tous les opérateurs : i, θ, d.

Soient $\mathrm{I}a, \mathrm{I}s, \mathrm{I}w$ les algèbres unitaires que forment les invariants (annulés par tous les $\theta(x)$, $x \in \mathfrak{g}$) de $\wedge\mathfrak{g}^*, \vee\mathfrak{g}^*$ et W. On peut identifier I_a à $\mathrm{I}_a \otimes 1$, I_s à $1 \otimes \mathrm{I}_s$ en multipliant les degrés par 2, et considérer ainsi I_a et I_s comme des sous-algèbres de I_W. L'expression de la différentielle sur $1 \otimes \vee\mathfrak{g}^*$ montre qu'un élément $1 \otimes a$ est un cocycle si et seulement si a est un invariant : $1 \otimes \mathrm{I}_s \equiv \mathrm{Z}(1 \otimes \vee\mathfrak{g}^*)$.

Puisque $d\theta(x) = \theta(x)d$, $\forall x \in \mathfrak{g}$, $\mathrm{I}_a, \mathrm{I}_s, \mathrm{I}_\mathrm{W}$ sont des algèbres différentielles graduées. La propriété fondamentale des algèbres différentielles W ou de I_W est que leur cohomologie est triviale : tout cocycle de W ou de I_W, de degré > 0, est le cobord : $z = dw$ d'un élément de cette algèbre défini à un cocycle près.

Ce résultat se démontre par récurrence descendante sur le deuxième degré partiel. Si h est la dérivation de degré (-1) de W, nulle sur $\wedge\mathfrak{g}^* \otimes 1$ qui prolonge l'application : $h(1 \otimes \alpha) = \alpha \otimes 1$ de $1 \otimes \mathfrak{g}^*$, sur $\mathfrak{g}^* \otimes 1$, $dh + hd$ est une dérivation de degré zéro. Si u est homogène de degré $m > 0$, et si q est le maximum du second degré partiel de ses termes non nuls, ce maximum est $\leqslant q-1$ dans $u - \dfrac{1}{m-q}(hd+dh)$. En itérant le procédé, on obtient finalement zéro, et si u est un cocycle, on l'égale ainsi à un cobord dw.

Le passage des invariants symétriques de $\vee\mathfrak{g}^*$ identifiés, avec un degré doublé, à ceux de $1 \otimes \vee\mathfrak{g}^*$ dans W, aux invariants antisymétriques se fait alors de la façon suivante : si $z \in 1 \otimes \mathrm{I}_s$, c'est que $dz = 0$, et $z = dw$ où $w \in \mathrm{I}_{\mathrm{W}}$ n'est défini qu'à un cocycle donc au cobord d'un invariant près : $z = dw = d(w + da)$. Si Π est la projection de W sur $\wedge\mathfrak{g}^*$, $\Pi(w+da) = \Pi w + d(\Pi a)$. Mais Πa est un invariant de $\wedge\mathfrak{g}^*$, donc un cocycle, et $d(\Pi a) = 0$. On obtient ainsi une application canonique Π de l'algèbre des invariants symétriques dans celle des invariants antisymétriques. En reprenant leurs graduations naturelles, on a :

$$\mathrm{I}_s^p \xrightarrow{\Pi} \mathrm{I}_a^{2p-1} \text{ pour}, p \geqslant 1.$$

Soit $\mathrm{P} = \mathrm{Im}\,\Pi$. On peut choisir une base de P formée d'éléments homogènes, de degré impairs et en prendre des images réciproques dans I_s. Soit T le sous-espace de I_s dont ils forment une base. On a alors le résultat remarquable suivant, dont la démonstration dépasse les limites de cet ouvrage :

Théorème IV.14.D. Si $\mathfrak{g}$ est semi-simple, $\mathrm{I}_a = \wedge\mathrm{P}$ est l'algèbre extérieure de P, $\mathrm{I}_s = \vee\mathrm{T}$ est l'algèbre des polynômes sur T. La dimension commune de P et de T est le rang de $\mathfrak{g}$.

Exemples :

1) L'algèbre de Lie $\mathfrak{U}(n)$ du groupe unitaire $\mathrm{U}(n)$ est formée des matrices complexes $n \times n$ antihermitiennes. Si $\mathrm{X} \in \mathbb{C}(n)$ les coefficients du polynôme en t : $\mathrm{dét}(t\mathrm{I}_n - \mathrm{X}) = t^n - c_1(\mathrm{X})t^{n-1} + \ldots (-1)^n c_n(\mathrm{X})$ avec $c_n(\mathrm{X}) = \mathrm{dét}\,\mathrm{X}$ sont les traces des représentations dans les puissances extérieures de l'espace $\mathbb{C}^n$ de départ (Proposition IV.6). Si X est antihermitienne, les $c_j(\mathrm{X})$ sont des fonctions polynômiales sur $\mathfrak{U}(n)$ invariantes par la représentation adjointe du groupe $\mathrm{U}(n)$ donc aussi par $\mathfrak{U}(n)$. Elles sont algébriquement indépendantes et engendrent l'algèbre des fonctions polynômiales sur $\mathfrak{U}(n)$ invariantes par la représentation adjointe de $\mathfrak{U}(n)$ et par celle de $\mathrm{U}(n)$).

2) L'algèbre de Lie $\mathfrak{O}(n)$ du groupe orthogonal $\mathrm{O}(n)$ est formée des matrices antisymétriques réelles $n \times n$. De même que dans 1), si $\mathrm{X} \in \mathfrak{O}(n)$, les coefficients du polynôme en t :

$$\text{dét}(t\mathrm{I}_n - \mathrm{X}) = t^n + p_1(\mathrm{X})t^{n-2} + \dots$$

sont, si $n = 2r$ ou $2r+1$, r fonctions polynômiales sur $\mathfrak{O}(n)$. Si $n = 2r$, $p_r(\mathrm{X}) = \text{dét}\,\mathrm{X}$ est le carré du pfaffien $\mathrm{P}f$ (§ IV.12) de X.

Les r fonctions polynômiales : $p_1, \dots p_r$ si $n = 2r+1$, $p_1, \dots p_{r-1}$, $\mathrm{P}f$ si $n = 2r$ sont algébriquement indépendantes et engrendrent l'algèbre des fonctions polynômiales sur $\mathfrak{O}(n)$ invariantes par la représentation adjointe. Elles sont invariantes par $\mathrm{SO}(n)$. Le groupe $\mathrm{O}(n)$ les laisse invariantes à l'exception de $\mathrm{P}f$ qui change de signe par les éléments de $\mathrm{O}(n)_-$.

IV.14 – Champs de tenseurs, opérateurs différentiels et formes différentielles extérieures sur un espace vectoriel réel de dimension finie nu (sans métrique). Gradient, divergence et rotationnel

Soit U un domaine d'un espace vectoriel *réel* E de dimension finie. Soit F un espace vectoriel fixe, par exemple $\mathbb{R}$, E lui-même ou plus généralement un espace de tenseurs sur E : $\otimes^{(v)}\mathrm{E}$.

Une fonction f sur U, à valeurs dans F, est ainsi appelée, suivant les cas, une fonction scalaire, un champ de vecteurs, un champ de tenseurs, etc. ... sur U. Rappelons (§ II.2) que f est dite différentiable en un point x_0 s'il existe une application linéaire $\mathrm{D}f$ de E dans F telle que :

$$f(x_0 + h) - f(x_0) = \mathrm{D}f_{x_0}(h) + ||h||g(x_0; h)$$

où $g(x_0; h)$ tend vers zéro avec h.

Si $(e_1, e_2, \dots e_n)$ est une base de E, il en résulte que :

$$\lim_{t \to 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = \left(\frac{\partial f}{\partial x^j}\right)_{x_0} = \mathrm{D}f_{x_0}(e_j)$$

$$\mathrm{D}f_{x_0}(h) = \mathrm{D}f_{x_0}(\Sigma e_j h^j) = \Sigma \mathrm{D}f_{x_0}(e_j)h^j = \Sigma\langle h, e^{*j}\rangle \mathrm{D}f_{x_0}(e_j).$$

On peut ainsi écrire la différentielle $\mathrm{D}f_{x_0}$ comme l'élément de l'espace $\mathrm{E}^* \otimes \mathrm{F}$:

$$\mathrm{D}f_{x_0} = \Sigma e^{*j} \otimes \mathrm{D}f_{x_0}(e_j) = \Sigma e^{*j} \otimes \left(\frac{\partial f}{\partial x^j}\right)_{x_0}.$$

Il est ici d'usage d'identifier l'espace $\mathfrak{L}(\mathrm{E};\mathrm{F})$ des applications linéaires de E dans F au produit tensoriel $\mathrm{E}^* \otimes \mathrm{F}$, la valeur de la différentielle $\mathrm{D}f_{x_0}$ pour un vecteur h s'obtenant par contraction *à gauche* de $\mathrm{D}f_{x_0}$ par h. Ainsi, le contracté $e_i \cdot \mathrm{D}f_{x_0}$ est la dérivée $\left(\dfrac{\partial f}{\partial x^i}\right)_{x_0}$.

La contraction de $\mathrm{D}f_{x_0}$ et d'un vecteur X se note $\mathrm{X} \cdot f$ et est appelé *dérivée de f par rapport à* X. Le vecteur X est considéré dans ce cas comme un *opérateur de dérivation* et le fait que cette dérivation ait lieu en un point x_0 particulier, conduit à attacher le vecteur X à ce point lorsqu'on lui attribue cette fonction et à lui donner le nom de vecteur tangent en x_0 (cf § II.2).

Soient X, Y deux *champs de vecteurs* sur U, f une fonction sur U à valeurs dans $\mathrm{F}, \mathrm{X}, \mathrm{Y}$ et f étant supposés de classe $\mathfrak{C}_2$. On a :

$$\begin{aligned}\mathrm{X} \cdot (\mathrm{Y} \cdot f) &= \sum \mathrm{X}^i \frac{\partial}{\partial x^j}\left(\sum \mathrm{Y}^j \frac{\partial f}{\partial x^j}\right) \\ &= \sum \mathrm{X}^i \frac{\partial \mathrm{Y}_j}{\partial x^i}\frac{\partial f}{\partial x^j} + \sum \mathrm{X}^i \mathrm{Y}^j \frac{\partial^2 f}{\partial x^j \partial x^i}\end{aligned}$$

d'où : $\mathrm{X} \cdot (\mathrm{Y} \cdot f) - \mathrm{Y}(\mathrm{X} \cdot f) = \sum\limits_j (\sum\limits_i (\mathrm{X}^i \dfrac{\partial \mathrm{Y}^j}{\partial x^i} - \mathrm{Y}^i \dfrac{\partial \mathrm{X}^j}{\partial x^i}))\dfrac{\partial f}{\partial x^j} = \mathrm{Z} \cdot f.$

L'opérateur $f \to \mathrm{Z} \cdot f = \mathrm{X} \cdot (\mathrm{Y} \cdot f) - \mathrm{YY} \cdot (\mathrm{X} \cdot f)$ est donc, en chaque point de U, un vecteur tangent, et sur U un champ de vecteurs, de composantes $\mathrm{Z}^j = \sum\limits_i \left(\mathrm{X}^i \dfrac{\partial \mathrm{Y}^j}{\partial x^i} - \mathrm{Y}^i \dfrac{\partial \mathrm{X}^j}{\partial x^i}\right)$. Ce champ de vecteurs Z est égal à : $\mathrm{Z} = \mathrm{DY}(\mathrm{X}) - \mathrm{DX}(\mathrm{Y})$.

Définition IV.14.A.

Si X et Y sont deux champs de vecteurs de classe $\mathfrak{C}_r$ sur un domaine U de E, leur *crochet* noté $[\mathrm{X}, \mathrm{Y}]$ est le champ de vecteurs de classe $\mathfrak{C}_{r-1}$:

$$[\mathrm{X}, \mathrm{Y}] = \mathrm{DY}(\mathrm{X}) - \mathrm{DX}(\mathrm{Y}).$$

On vérifie aisément les propriétés suivantes du crochet :

1) antisymétrique : $[X, Y] = -[X, Y]$.
2) identité de Jacobi : $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.
3) Si $X \cdot f = 0$ et $Y \cdot f = 0$, alors $[X, Y] \cdot f = 0$.
4) Les champs de vecteurs-dérivations par rapport aux vecteurs d'une base « commutent » :

$$\left[\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right] = 0, \forall j, k.$$

Définition IV.14.B.

Si X est un champ de vecteurs continûment différentiable dans un domaine U de E, l'opérateur L_x qui associe à chaque champ de vecteurs Y continûment différentiable sur U le crochet $[X, Y] = L_X Y$ est appelé *dérivée de Lie* de Y par rapport à X. On verra au chapitre IX.§ 2, la signification géométrique et l'extension de l'opérateur L_X aux champs de tenseurs.

Munis du crochet, les champs de classe $\mathfrak{C}_\infty$ sur U forment ainsi une algèbre de Lie dont les translations à gauche L_X sont des dérivations : $L_X \cdot [Y, Z] = [L_X Y, Z] + [Y, L_X Z]$ (identité de Jacobi).

L'opérateur de différentiation s'écrit dans une base e :

$$D = \sum e^{*k} \otimes \frac{\partial}{\partial x^k}.$$

En mathématiques appliquées, on note ∇ l'opérateur de dérivation ce qui se lit del, ou parfois « nabla » ! et on le considère comme un « vecteur covariant » de composantes $\dfrac{\partial}{\partial x^1}, \ldots, \dfrac{\partial}{\partial x^n}$. Cela présente un avantage mnémotechnique dans l'énoncé de certaines formules. Son usage est d'ailleurs essentiellement limité à l'espace euclidient E_3.

Si F est un champ de tenseurs d'ordre p, DF est un champ de tenseurs d'ordre $(p + 1)$ une variance covariante étant ajoutée à celles de F.

Exemples :

1) La différentielle d'une fonction à valeurs réelles f différentiable est un *champ de covecteurs*, le *gradient de* f : $\operatorname{grad} f = \mathrm{D}f = \sum \dfrac{\partial f}{\partial x^k} e^{*k}$, qui s'écrit aussi ∇f.

2) Si v est un champ différentiable de *vecteurs*, $v = \sum v^j e_j$, la différentielle $\mathrm{D}v$ est un champ de tenseurs mixtes (c'est-à-dire d'opérateurs linéaires) :

$$\mathrm{D}v = \sum e^{*k} \otimes \frac{\partial v}{\partial x^k} = \sum \frac{\partial v^j}{\partial x^k} e^{*k} \otimes e_j$$

dont les traces forment une fonction scalaire appelée *divergence* du champ v : $\operatorname{div} v = \sum \dfrac{\partial v^j}{\partial x^j}$. L'opérateur ∇ permet d'écrire la divergence comme un produit scalaire $\operatorname{div} v = \nabla \cdot v$.

3) Si a est un champ différentiable de *covecteurs* $a = \sum a_j e^{*j}$ la différentielle $\mathrm{D}a$ est un champ de formes bilinéaires :

$$\mathrm{D}a = \sum e^{*k} \otimes \frac{\partial a}{\partial x^k} = \sum \frac{\partial a_j}{\partial x^k} e^{*k} \otimes e^{*j}$$

dont l'antisymétrisé est appelé le rotationnel de a que l'on retrouvera ci-dessous :

$$\operatorname{rot} a = \sum_{j<k} \left(\frac{\partial a_k}{\partial x^j} - \frac{\partial a_j}{\partial x^k} \right) (e^{*j} \otimes e^{*k} - e^{*k} \otimes e^{*j}).$$

On a : $\operatorname{rot} \operatorname{grad} f \equiv 0$.

On a déjà indiqué au § II.2 la notation adoptée en géométrie différentielle pour désigner les formes linéaires coordonnées (e^{*1}, $e^{*2} \dots e^{*n}$) d'une base e de E : on les représente par $(dx^1, dx^2, \dots dx^n)$, différentielles des formes linéaires coordonnées $(x \to x^j)$ auxquelles elles sont identiques. Avec cette notation, un champ de covecteurs prend alors la forme : $a = \sum a_j dx^j$ et le nom de forme différentielle. Plus généralement nous allons considérer des « formes différentielles extérieures » de tous les degrés p compris entre 0 (fonctions numériques) et n, dimension de E. Ce sont les expressions qui apparaissent dans les intégrales multiples :

– curvilignes : $\int \sum a_j dx^j$, que l'on intègre sur un arc de courbe;

– de surface : $\int\int\sum a_{ij}dx^i dx^j$, que l'on intègre sur un morceau de surface etc... Les règles du calcul intégral sont celles du calcul différentiel extérieur. Dans le plan $\mathbb{R}^2$ par exemple, rapporté à des coordonnées (x, y), un changement de variables conservant l'orientation : $x = x(u, v)$, $y = y(u, v)$ se traduit dans une intégrale double par :

$$\int\int a(x,y)dxdy = \int\int a(x(u,v), y(u,v))\frac{\mathrm{D}(x;y)}{\mathrm{D}(u,v)}dudv$$

De l'égalité : $dx\,dy = \text{dét}\begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} du\ dv$ il résulte que $dydx = -dxdy$. Le produit des différentielles $dxdy$ est bien un produit extérieur. Nous l'écrivons le plus souvent $dx \wedge dy$ pour éviter toute ambiguïté, mais l'usage en calcul intégral est de sous-entendre le signe $\wedge$ du produit extérieur.

Définition IV.14.C.

Un champ de p-formes extérieures dans un domaine U de l'espace vectoriel réel E est appelé une *p-forme différentielle extérieure* sur U. Dans une base e de E, elle s'écrit :

$$\alpha(x) = \sum_{(i_1<i_2<\ldots i_p)} a_{i_1 i_2 \ldots i_p}(x)dx^{i_1} \wedge dx^{i_2} \wedge \ldots dx^{i_p}.$$

D'après les propositions IV.1.B. et IV.1.C. il correspond canoniquement à une telle forme différentielle extérieure α un champ de tenseurs covariants antisymétriques a qui s'écrit dans la base e :

$$a(x) = \sum_{(i_1<i_2<\ldots i_p)} a_{i_1 i_2 \ldots i_p}(x) \sum_{\sigma\in\mathfrak{S}_p} \varepsilon(\sigma)e^{*^{i}\sigma 1} \otimes e^{*^{i}\sigma 2} \otimes \ldots e^{*^{i}\sigma p}$$

I désignant une suite strictement croissante $(1 \leqslant i_1 < i_2 < \ldots i_p \leqslant n)$ de p indices, et A_p l'opérateur d'antisymétrisation $\sum_{\sigma\in\mathfrak{S}_p} \varepsilon(\sigma)$:

$$\alpha(x) = \sum a_{\mathrm{I}}(x)dx^{\mathrm{I}} \text{ et } a(x) = \sum a_{\mathrm{I}}(x)\mathrm{A}_p(\otimes^{\mathrm{I}} e^*)$$

Si $X_1, X_2, \dots X_p$ sont p vecteurs tangents au point x :

$$\begin{aligned}\alpha(x)(X_1, X_2, \dots X_p) &= \langle \alpha(x); X_1 \wedge X_2 \wedge \dots X_p \rangle \\ &= \sum a_I(x) \text{dét}|X_{k,i}| = a(x)(X_1, X_2, \dots X_p)\end{aligned}$$

est le contracté de $\alpha(x)$, ou de $a(x)$ et des p vecteurs. Si maintenant $X_1, X_2, \dots X_p$ sont p champs de vecteurs sur U, la forme α leur fait correspondre la fonction numérique sur U qui à chaque x associe le scalaire $\alpha(x)(X_{1x}, X_{2x}, \dots X_{px})$. α détermine ainsi une application $\widetilde{\alpha}$ de $X_1 \times X_2 \times X_p$ à valeurs dans les fonctions numériques sur U. Quelles sont les propriétés de cette application $\widetilde{\alpha}$? Supposons les champs de vecteurs et les formes r-fois continûment différentiables et soit $\mathfrak{F}_U$ l'algèbre des fonctions numériques r fois continûment différentiables sur U.

Si $f_1, f_2, \dots f_p \in \mathfrak{F}_U$, on a $\widetilde{\alpha}(f_1 X_1, \dots f_p X_p) = f_1 \dots f_p \widetilde{\alpha}(X_1, \dots X_p)$ $\widetilde{\alpha}$ est donc $\mathfrak{F}_U$-p-linéaire alternée. Réciproquement, ce sera une conséquence du théorème très général de tensorialité (théorème IX.2) qu'une telle application $\widetilde{\alpha}$ détermine une forme différentielle extérieure α. Cette identification, fort importante dans les applications, permet donc d'écrire une forme différentielle extérieure sur $U \subset E$ de deux manières :

1) comme une fonction α sur U à valeurs dans $\wedge^p E^* \equiv \mathfrak{L}(\wedge^p E; \mathbb{R}) \equiv \mathrm{Alt}_p(E; \mathbb{R})$ (cf théorème IV.1.A)

2) comme une fonction $\widetilde{\alpha}$ $\mathfrak{F}_U$-p-linéaire alternée des champs de vecteurs sur U.

Si maintenant a est le champ de tenseurs covariants antisymétriques correspondant à la p-forme différentielle extérieure α, sa différentielle Da est un champ de tenseurs $(p+1)$-fois covariants mais qui *ne sont plus* antisymétriques. On a en effet :

$$Da = \sum \frac{\partial a_I}{\partial x^k} e^{*k} \otimes A_p(\otimes^I e^*).$$

Le procédé canonique pour rendre Da antisymétrique est l'antisymétrisation partielle $A_{1,p}$ du premier facteur e^{*k} avec les p autres : $A_{p+1} = A_{1,p} \circ A_p$ (cf § III.6 et 8). Le résultat est la différentielle antisymétrisée $\widetilde{D}a$:

$$\begin{aligned}\widetilde{D}a &= \sum \frac{\partial a_I}{\partial x^k} A_{1,p}(e^{*k} \otimes A_p(\otimes^I e^*)) \\ &= \sum \frac{\partial a_I}{\partial x^k} A_{p+1}(e^{*k} \otimes (\otimes^I e^*)).\end{aligned}$$

La forme différentielle extérieure qui lui correspond canoniquement est appelée la *différentielle extérieure* $d\alpha$ de α :

$$\boxed{\text{Si } \alpha = \sum a_{\mathrm{I}}(x) dx^{\mathrm{I}},\ d\alpha = \sum \frac{\partial a_{\mathrm{I}}}{\partial x^k} dx^k \wedge dx^{\mathrm{I}} = \sum da_{\mathrm{I}} \wedge dx^{\mathrm{I}}}.$$

Si l'on considère α comme une fonction sur U à valeurs dans $\wedge^p \mathrm{E}^*$, la différentielle ordinaire de cette fonction est :

$$\mathrm{D}\alpha = \sum dx^k \otimes \frac{\partial \alpha}{\partial x^k} = \sum \frac{\partial a_{\mathrm{I}}}{\partial x^k} dx^k \otimes dx^{\mathrm{I}} = \sum da_{\mathrm{I}} \otimes dx^{\mathrm{I}}.$$

La différentielle extérieure de la forme α en est l'antisymétrisée par l'antisymétrisation partielle $\mathrm{A}_{1,p}$ dans le sens suivant : quels que soient les vecteurs $\mathrm{X}_0, \mathrm{X}_1, \ldots \mathrm{X}_p$, on a :

$$\boxed{d\alpha(\mathrm{X}_0, \mathrm{X}_1, \ldots \mathrm{X}_p) = \sum_{i=0}^{p} (-1)^i \mathrm{D}\alpha(\mathrm{X}_i; \mathrm{X}_0, \ldots \widehat{\mathrm{X}}_i, \ldots \mathrm{X}_p)}.$$

Si α est une forme monôme, par exemple $\alpha = a dx^1 \wedge \ldots \wedge dx^p$, l'égalité ci-dessus n'exprime rien d'autre que le développement du déterminant : $(da \wedge dx^1 \wedge \ldots \wedge dx^p)(\mathrm{X}_0, \mathrm{X}_1, \ldots \mathrm{X}_p)$ par rapport à la première ligne :

$$\boxed{\sum_{i=0}^{p} (-1)^i \langle \mathrm{D}a, \mathrm{X}_i \rangle \cdot dx^{\mathrm{I}}(\mathrm{X}_i; \mathrm{X}_0, \ldots \widehat{\mathrm{X}}_i, \ldots \mathrm{X}_p)}.$$

La propriété remarquable est que l'expression de la différentielle extérieure reste la même quelle que soit la base de E, et, plus encore, quel que soit le choix de coordonnées « curvilignes » dans U, telles qu'elles seront définies au § VII.9. On peut naturellement vérifier cette propriété par un calcul direct, permettant de constater que l'antisymétrisation élimine les dérivées secondes d'un changement de coordonnées curvilignes. Mais on peut aussi remplacer le calcul par un raisonnement basé sur les propriétés de la différentielle extérieure d, que l'on peut démontrer en utilisant son expression dans *un* système de coordonnées particulières, soient :

1) d est $\mathbb{R}$-linéaire : $d(\lambda\alpha + \mu\beta) = \lambda d\alpha + \mu d\beta$

2) si f est une fonction scalaire différentiable, considérée comme une forme différentielle de degré zéro, $df = \mathrm{D}f = \sum \frac{\partial f}{\partial x^j} dx^j$

3) si α et β sont des formes différentielles extérieures de degrés respectifs p et q, leur produit extérieur $\alpha \wedge \beta$ est une forme différentielle de degré $(p+q)$ et $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \wedge d\beta$.

Par linéarité, il suffit de prouver cette égalité sur des formes monômes : $\alpha = adx^{\mathrm{I}}$, $\beta = bdx^{\mathrm{J}}$, d'où $\alpha \wedge \beta = ab\, dx^{\mathrm{I}} \wedge dx^{\mathrm{J}}$. On a :

$$\begin{aligned} d(\alpha \wedge \beta) &= \sum \left(\frac{\partial a}{\partial x^k} b + a \frac{\partial b}{\partial x^k} \right) dx^k \wedge dx^{\mathrm{I}} \wedge dx^{\mathrm{J}} \\ &= \left(\sum \frac{\partial a}{\partial x^k} dx^k \wedge dx^{\mathrm{I}} \right) \wedge bdx^{\mathrm{J}} + (-1)^p adx^{\mathrm{I}} \wedge \\ &\qquad \left(\sum \frac{\partial b}{\partial x^k} dx^k \wedge dx^{\mathrm{J}} \right) \text{ c.q.f.d.} \end{aligned}$$

4) $d(d\alpha) = 0$ $(d^2 \equiv 0)$ pour α de classe $\mathfrak{C}_2$.

Là encore il suffit de prouver cette égalité sur une forme monôme $\alpha = adx^{\mathrm{I}}$.

$$\begin{aligned} d(d\alpha) &= \sum_i \sum_k \frac{\partial^2 a}{\partial x^k \partial x^j} dx^j \wedge dx^k \wedge dx^{\mathrm{I}} \\ &= \sum_{j<k} \left(\frac{\partial^2 a}{\partial x^k \partial x^j} - \frac{\partial^2 a}{\partial x^j \partial x^k} \right) dx^j \wedge dx^k \wedge dx^{\mathrm{I}} = 0. \end{aligned}$$

Cette propriété a naturellement pour origine le fait que le carré de l'opérateur de différentiation :

$$\mathrm{D}^2 = \sum \left(e^{*j} \otimes \frac{\partial}{\partial x^j} \right) \circ \sum \left(e^{*k} \otimes \frac{\partial}{\partial x^k} \right) = \sum e^{*j} \otimes e^{*k} \otimes \frac{\partial^2}{\partial x^k \partial x^j}$$

est symétrique et donne zéro par antisymétrisation.

Si la démonstration de ces propriétés a fait intervenir un système de coordonnées particulier, les propriétés elles-mêmes sont indépendantes d'un tel choix et déterminent la différentielle extérieure complètement. Si une forme différentielle extérieure s'écrit : $\alpha = \sum a_{\mathrm{I}}(u) du^{i_1} \wedge du^{i_2} \wedge \ldots du^{i_p}$ où $u^1, u^2, \ldots u^n$ sont des fonctions différentiables de $x^1, x^2, \ldots x^n$, les propriétés précédentes donnent immédiatement : $d\alpha = \sum da_{\mathrm{I}}(u) \wedge du^{\mathrm{I}} = \sum_{\mathrm{I},k} \frac{\partial a_{\mathrm{I}}}{\partial u^k} du^k \wedge du^{\mathrm{I}}$.

Théorème IV.14.A. Si $X_0, X_1, \dots X_p$ sont $(p+1)$ champs de vecteurs continûment différentiables, on peut exprimer la différentielle extérieure d'une p-forme α dans l'interprétation des formes différentielles comme fonctions multilinéaires alternées de champs de vecteurs, à l'aide d'une somme de deux termes :

$$\boxed{\begin{aligned} d\alpha(X_0, X_1, \dots X_p) = &\sum_{i=0}^{p}(-1)^i X_i \cdot \alpha(X_0, X_1, \dots \widehat{X}_i, \dots X_p) \\ &+ \sum_{i<j}(-1)^{i+j}\alpha([X_i, X_j], (X_0, \dots \widehat{X}_i, \dots, \widehat{X}_j, \dots X_p) \end{aligned}}\,.$$

En particulier pour $p = 2$:

$$\boxed{d\alpha(X, Y) = X \cdot \alpha(Y) - Y \cdot \alpha(X) - \alpha([X, Y])}\,,$$

pour $p = 3$:

$$\begin{aligned} d\alpha(X, Y, Z) = X \cdot \alpha(Y, Z) + Y \cdot \alpha(Z, X) + Z \cdot \alpha(X, Y) \\ -\alpha([X, Y], Z) - \alpha([Y, Z], X) - \alpha([Z, X], Y) \end{aligned}$$

Preuve : Calculons la différentielle de la fonction scalaire : $x \to f(x) = \alpha(x)(X_1(x), X_2(x), \dots X_p(x))$ et prenons la valeur de cette différentielle pour le vecteur $X_0(x)$

$$\begin{aligned} X_0 \cdot f = X_0 \cdot \alpha(X_1, X_2, \dots X_p) &= D\alpha(X_0; X_1, \dots X_p) \\ &+\alpha(DX_1(X_0), X_2, \dots X_p) + \alpha(X_1, DX_2(X_0), \dots X_p) \\ &+ \dots \alpha(X_1, \dots DX_p(X_0)) \\ &= D\alpha(X_0; X_1, \dots X_p) + \sum(-1)^j \alpha(DX_j(X_0), X_1, \dots \widehat{X}_j, \dots X_p) \end{aligned}$$

En écrivant de façon analogue la valeur de la différentielle de la fonction $\alpha(X_0, X_1, \dots \widehat{X}_i, \dots X_p)$ pour X_i, on obtient en antisymétrisant :

$$\begin{aligned} \sum_{i=0}^{p}(-1)^i X_i \cdot \alpha(X_0, X_1, \dots \widehat{X}_i, \dots X_p) &= d\alpha(X_0; X_1, \dots X_p) \\ &+ \sum_{j<i}(-1)^{i+j}\alpha(DX_j(X_i), X_0, \dots \widehat{X}_i, \dots \widehat{X}_j, \dots X_p) \\ &+ \sum_{i<j}(-1)^{i+j-1}\alpha(DX_j(X_i), X_0, \dots \widehat{X}_i, \dots \widehat{X}_j, \dots X_p) \end{aligned}$$

et puisque $DX_j(X_i) - DX_i(X_j) = [X_i, X_j]$ en rassemblant les termes correspondant au même couple (i, j) dans les deux dernières sommes, on obtient la formule annoncée.

Exemple : le rotationnel d'un champ de covecteurs $a = \Sigma a_j e^{*j}$ a été défini comme l'antisymétrisé de sa différentielle. En termes de formes différentielles, a est une 1-forme : $a = \Sigma a_j dx^j$ et le rotationnel est la 2-forme $da = \Sigma \left(\frac{\partial a_k}{\partial x^j} - \frac{\partial a_j}{\partial x^k}\right) dx^j \wedge dx^k$.

Définition IV.14.D.

Une forme différentielle extérieure a dont la différentielle extérieure da est nulle est dite *fermée*.

Une forme a qui est la différentielle d'une autre : $a = db$ est souvent appelé un « *bord* » (terminologie un peu laxiste tenant de la topologie algébrique).

Tout bord est naturellement fermé. Réciproquement.

Théorème de Poincaré IV.14.B. Si a est une forme différentielle extérieure fermée, de classe $\mathfrak{C}_p$, $p \geqslant 2$, sur un ouvert U de $\mathbb{R}^n$ difféomorphe à une boule, il existe sur U, une forme b, de classe $\mathfrak{C}_{p-1}$ dont a est la différentielle extérieure.

Preuve : elle est classique mais assez longue : on intègre sur les « rayons » issus d'un point de U.

Plus brièvement, toute forme fermée est localement un bord. En particulier, si le rotationnel d'un champ de covecteurs a est nul, a est localement la différentielle d'une fonction scalaire.

Nous reviendrons au chapitre V, sur les formes différentielles extérieures, cette fois en présente d'une métrique.

EXERCICES SUR LE CHAPITRE IV

EX. IV.1 :

Si l'espace vectoriel E est somme directe de deux sous-espaces F et H, montrer que l'on a un isomorphisme naturel de $\wedge^m E$ sur la somme directe des $(m+1)$ espaces vectoriels : $\wedge^p F \otimes \wedge^{m-p} H$ où $p = 0, 1, 2, \ldots m$.

EX. IV.2 :

Le rang d'un bivecteur b est le plus petit entier r tel que b soit somme de r bivecteurs décomposables : $b = x_1 \wedge y_1 + \dots x_r \wedge y_r$. Montrer qu'alors les $2r$ vecteurs $x_1, y_1, \dots x_r, y_r$ sont linéairement indépendants. Réciproque.

EX. IV.3 :

Trouver un système d'équations entre les composantes $b_{ij} (i < j)$ d'un bivecteur b exprimant la condition nécessaire et suffisante pour que b soit décomposable (de rang un).

EX. IV.4 :

Soit E un espace vectoriel de dimension $n > 1$. Montrer qu'il n'existe pas d'isomorphisme canonique entre $\wedge^p \mathrm{E}$ et $\wedge^{n-p} \mathrm{E}$ $(p \neq n - p)$ c'est-à-dire d'isomorphisme φ tel que pour tout $a \in \mathcal{L}(\mathrm{E})$, on ait $\varphi \circ \overline{\wedge}^p a = \overline{\wedge}^{n-p} a \circ \varphi$.

EX. IV.5 :

Déterminer les éléments inversibles de l'algèbre $\wedge \mathrm{E}$.

EX. IV.6 :

En utilisant le développement de Laplace montrer que si $a \in \mathcal{L}(\mathrm{E})$, on a : $\mathrm{dét}(\overline{\wedge}^p a) \cdot \mathrm{dét}(\overline{\wedge}^{n-p} a) = (\mathrm{dét}\, a)^{\binom{n}{p}}$.

EX. IV.7 :

Soit $u \in \wedge \mathrm{E}$. Montrer que les covecteurs $\alpha \in \mathrm{E}^*$ tels que $u \llcorner \alpha = 0$ forment un sous-espace Φ de E^* possédant la propriété suivante : l'orthogonal F de Φ dans E est le plus petit sous-espace de E tel que $u \in \wedge \mathrm{F}$.

EX. IV.8 :

Montrer que l'intégrale curviligne suivante prise sur un chemin différentiable de $\mathbb{R}^3$ ne dépend que des extrêmités P et Q de ce chemin :

$$\int_{\mathrm{P}}^{\mathrm{Q}} (x^2 - yz)dx + (y^2 - zx)dy + (z^2 - xy)dz$$

EX. IV.9 :

On a vu que tout élément de $\wedge^{n-1}\mathrm{E}$ (avec $\dim \mathrm{E} = n$) est décomposable. On peut se poser un problème analogue (*local*) pour une $(n-1)$ forme différentielle extérieure.

Dans $\mathbb{R}^3$, soit $\omega = \mathrm{A}\, dy \wedge dz + \mathrm{B} dz \wedge dx + \mathrm{C} dx \wedge dy$ une 2-forme de classe $\mathfrak{C}^2$ définie dans un voisinage U d'un point M_0 où $\mathrm{A}, \mathrm{B}, \mathrm{C}$ sont non nulles. Montrer qu'on peut trouver deux fonctions f, g au voisinage de x_0 telles que : $df \wedge \omega = 0$; $dg \wedge \omega = 0$; $df \wedge dg \neq 0$ et qu'il existe alors une fonction h telle que $\omega = h\, df \wedge dg$. Montrer que l'on peut avoir $h \equiv 1$ si et seulement si $d\omega = 0$.

EX. IV.10 :

Déterminer une fonction $f(y)$ telle que la 2-forme de $\mathbb{R}^3$: $\omega = x\, dy \wedge dz - 2z\, f(y) dx \wedge dy + y\, f(y) dz \wedge dx$ soit telle que $d\omega = dx \wedge dy \wedge dz$.

CHAPITRE V

Les tenseurs restreints par un produit scalaire symétrique

Tous les espaces vectoriels considérés dans ce chapitre sont de dimension finie.

Les espaces vectoriels rencontrés dans les applications sont généralement munis d'un produit scalaire naturel et par conséquent ont une « structure géométrique » : euclidienne, minkovskienne, symplectique, hermitienne, etc...

Un produit scalaire est un tenseur deux fois covariant, de rang maximum, qui prend le nom de *tenseur fondamental* ou de *tenseur métrique.*

Un tel tenseur définit un isomorphisme de l'espace sur son dual et permet ainsi de les identifier. On peut alors identifier tous les espaces tensoriels de même ordre : il n'y a plus qu'un seul type de tenseur d'ordre p dont on peut choisir les composantes avec la variance que l'on veut. Avec cette identification, les tenseurs prennent le nom de tenseurs restreints : euclidiens, minkovskiens, etc... Leur groupe de structure est le groupe des opérateurs linéaires de l'espace qui laissent invariant le tenseur fondamental, donc les identifications. Ces groupes sont suivant les cas, le groupe orthogonal, de Lorentz, symplectique, unitaire, etc...

Dans ce chapitre, on suppose simplement l'existence d'un produit scalaire bilinéaire *symétrique*. L'étude des tenseurs euclidiens pseudoeuclidiens et symplectiques fera l'objet des chapitres VII et VIII.

V.4 / Extension de la métrique aux espaces tensoriels et pseudotensoriels.
V.5 / Dualité dans l'algèbre extérieure d'un espace métrique $(\mathrm{E}; g)$.
V.6 / Représentations du groupe $\mathrm{SO}(\mathrm{E}; g)$ dans les puissances extérieures de E.

V.1 – Produit scalaire bilinéaire symétrique. Tenseurs restreints

Soient E un espace vectoriel de dimension finie sur le corps K et g une forme bilinéaire symétrique non-dégénérée sur E, qui prend le nom de produit scalaire. Nous noterons $q(x) = g(x, x)$, la forme quadratique associée.

La notation simplifiée $(x \mid y)$ pour désigner $g(x, y)$ sera souvent utilisée. C'est le même isomorphisme ρ de E sur son dual E^* qui sert d'application associée à droite comme à gauche de g :

$$g(x, y) = (x \mid y) = \langle x, \rho y \rangle = \langle \rho x, y \rangle.$$

Rappelons que par les identifications de $\mathcal{L}(\mathrm{E}, \mathrm{E}^*) \equiv \mathrm{E}^* \otimes \mathrm{E}^*$ et $\mathrm{Bil}(\mathrm{E}) \equiv \mathrm{E}^* \otimes \mathrm{E}^*$, ρ et g sont deux aspects d'un même tenseur de $\mathrm{E}^* \otimes \mathrm{E}^* \cdot \rho$ permet de définir le produit scalaire $\overset{-1}{g}$ sur E^* : $\overset{-1}{g}(\alpha, \beta) = g(\overset{-1}{\rho}\,\alpha, \overset{-1}{\rho}\,\beta)$ (cf. §II.10). Chaque vecteur x de E est identifié à la forme linéaire $\rho x = (x|$, produit scalaire par x, chaque covecteur $\alpha \in \mathrm{E}^*$ au vecteur $\overset{-1}{\rho}\,\alpha$ qui représente α comme un produit scalaire : $\langle \alpha, y \rangle = (\overset{-1}{\rho}\,\alpha \mid y)$.

L'isomorphisme linéaire ρ est aussi un isomorphisme pour la structure que définissent g sur E et $\overset{-1}{g}$ sur E^* puisque :

$$\overset{-1}{g}(\rho x, \rho y) = g(x, y).$$

Soit a un opérateur linéaire quelconque de E. De :

$$(ax \mid y) = \langle ax, \rho y \rangle = \langle x, {}^t a \rho y \rangle) = (x \mid \overset{-1}{\rho}\,{}^t a \rho y)$$

il résulte qu'il lui correspond un opérateur linéaire $a^* = \overset{-1}{\rho}\,{}^t a \rho$, dit *adjoint* défini par la condition :

$(ax \mid y) = (x \mid a^*y)$, $\forall x, y \in \mathrm{E}$. L'adjonction $*$ est une antiinvolution de l'algèbre $\mathfrak{L}(\mathrm{E})$: $a^{**} = a$ et $(ab)^* = b^*a^*$.

Soit $\mathrm{O}(\mathrm{E}; g)$ le groupe que forment les opérateurs linéaires de E qui conservent le produit scalaire, ou *groupe orthogonal* de $(\mathrm{E}; g)$:

$$\begin{aligned} a \in \mathrm{O}(\mathrm{E}; g) &\Leftrightarrow (ax \mid ay) = (x \mid a^*ay) = (x \mid y),\ \forall x, y \in \mathrm{E} \\ &\Leftrightarrow a^*a = \mathrm{Id} \text{ ou } a^* = \overset{-1}{a} \\ &\Leftrightarrow \overset{-1}{\rho}\, a\rho a = \mathrm{I} \text{ soit } \rho a =^t \overset{-1}{a}\, \rho = \check{a}, \rho \end{aligned}$$

D'où, pour l'opérateur contragrédient de a : $\check{a} = \rho a \overset{-1}{\rho}$

Les opérateurs orthogonaux de $(\mathrm{E}; g)$ sont donc exactement les opérateurs dont l'image par ρ dans E^* est leur contragrédient $\check{a} = {}^t\overset{-1}{a}$, c'est-à-dire ceux dont le contragrédient laisse invariant le produit scalaire $\overset{-1}{g}$ sur E^*. Il en résulte que ρ est aussi un isomorphisme de $\mathrm{O}(\mathrm{E}; g)$-modules, le groupe $\mathrm{O}(\mathrm{E}; g)$ opérant directement sur E, et par la représentation contragrédiente sur E^*.

Soient $e = (e_1, e_2, \ldots, e_n)$ une base de E, e^* sa duale. Le tenseur métrique g s'écrit dans e : $g = \sum g_{ij} e^{*i} \otimes e^{*j}$ avec $g_{ij} = g(e_i, e_j) = (e_i \mid e_j)$. Un vecteur x de E a deux ensembles de composantes : les ordinaires x^j ou *contravariantes*, et ses composantes en tant que forme linéaire $\rho x = (x \mid$, ou *covariantes*, que l'on écrit avec des indices inférieurs; x_i, et qui sont obtenues par contraction des x^j avec le tenseur métrique :
$x_i = (\rho x)_i = \langle \rho x; e_i \rangle = (x \mid e_i) = (\sum x^j e_j \mid e_i) = \sum g_{ij} x^j$.

Le produit scalaire de x et de y : $(x \mid y) = \sum g_{ij} x^i y^j$ peut donc s'écrire $(x \mid y) = \sum x^i y_i = \sum x_j y^j$. La matrice de ρ relative aux bases e, e^* est celle de g soit $|g_{ij}|$. La matrice de $\overset{-1}{g}$ est celle de $\overset{-1}{\rho}$, matrice inverse de $|g_{ij}|$. On la note $|g^{ij}|$ avec des indices supérieurs. Les g^{ij} sont les composantes du tenseur deux fois contravariant $\overset{-1}{g} = \sum g^{ij} e_i \otimes e_j$. Avec les identifications qui vont être faites ci-après, nous verrons qu'il ne s'agit pas là d'un tenseur distinct mais simplement de la forme deux fois contravariantes de g, les g^{ij} étant ainsi les composantes deux fois contravariantes du tenseur métrique.

Les x^j s'obtiennent de façon symétrique à partir des x_i par contraction avec le tenseur $\overset{-1}{g}$. On a $\rho x = \sum x_i e^{*i}$ et $x = \sum x_i \overset{-1}{\rho} e^{*i}$ d'où $x^j = \sum x_i \langle \overset{-1}{\rho} e^{*i}, e^{*j} \rangle = \sum x_i g^{ij}$.

On voit apparaître dans ce calcul l'image réciproque $f = \overset{-1}{\rho} e^*$ de la base duale. C'est une base de E dite *base supplémentaire de e* caractérisée par les équations : $(e_i \mid f^j) = \delta_{j,i}$ puisque $(e_i \mid f^j) = \langle e_i, \rho f^j \rangle = \langle e_i, e^{*j} \rangle$.

Les composantes covariantes du vecteur x dans la base e sont donc ses composantes relativement à la base supplémentaire f de e. La matrice $\mathrm{G} = |g_{ij}|$ est la matrice de passage de f à e : $e = f\mathrm{G}$. En effet, $g_{ij} = (e_i \mid e_j) = \langle e_i, \rho e_j \rangle$ d'où $\rho e_j = \sum e^{*i} g_{ij}$ et $e_j = \sum \overset{-1}{\rho} e^{*i} g_{ij} = \sum f^i g_{ij}$. Inversement $f = e \overset{-1}{\mathrm{G}}$, et $\overset{-1}{g} = \sum e_j \otimes f^j$.

g étant symétrique, il existe toujours des bases e de E diagonalisant g : $g_{ij} = 0 \; \forall (i \neq j)$, ou bases orthogonales.

Une telle base est caractérisée par le fait que sa base supplémentaire lui est « colinéaire » : $f^i = g^{ii} e_i$, $\forall i$.

Pour qu'il existe une base e pour laquelle il y ait identité entre composantes contravariantes et covariantes, il est nécessaire et suffisant que $f = e$, soit que la matrice G soit la matrice unité dans cette base : le produit scalaire est « de type euclidien » $(x \mid y) = \sum\limits_i x^i y^i$ et la base e est orthonormale.

Etant données deux espaces tensoriels sur E de même ordre mais de variances différentes : $\otimes^{(u)}\mathrm{E}$ et $\otimes^{(v)}\mathrm{E}$ avec $|u| = |v| = p$, soient pour chaque rang $k = 1, 2, \ldots, p$, ρ_k celle des trois applications : l'identité, ρ ou $\overset{-1}{\rho}$ qui applique (isomorphiquement) le k-ème facteur de $\otimes^{(u)}\mathrm{E}$ sur le k-ème facteur de $\otimes^{(u)}\mathrm{E}$, et $\rho_{(v)(u)}$ le produit tensoriel de ces applications : $\rho_1 \bar{\otimes} \rho_2 \bar{\otimes} \ldots \rho_p$. Par exemple :

$$\mathrm{E}^{(u)} = \mathrm{E} \otimes \mathrm{E}^* \otimes \mathrm{E} \longrightarrow \mathrm{E}^{(v)} = \mathrm{E} \otimes \mathrm{E} \otimes \mathrm{E}^*$$

$$\rho_{(v)(u)} = \mathrm{Id} \bar{\otimes} \overset{-1}{\rho} \otimes \rho$$

$\rho_{(v)(u)}$ permet d'identifier $\otimes^{(u)}\mathrm{E}$ et $\otimes^{(v)}\mathrm{E}$ exactement comme ρ et $\overset{-1}{\rho}$ ont permis d'identifier E et E*. Si $\otimes^{(w)}\mathrm{E}$ est un troisième espace tensoriel de même ordre, on a évidemment $\rho_{(w)(v)} \circ \rho_{(v)(u)} = \rho_{(w)(u)}$ tandis que $\rho_{(u)(v)} = \overset{-1}{\rho}_{(v)(u)}$.

Ces formules montrent que les isomorphismes $\rho_{(v)(u)}$ déterminent une relation d'équivalence dans la réunion des espaces tensoriels sur E de même ordre p. L'ensemble quotient, en bijection avec l'un quelconque des espaces tensoriels d'ordre p est un espace vectoriel appelé *espace des tenseurs restreints d'ordre p de* (E; g). Nous le noterons $\mathrm{T}_p(\mathrm{E})$. Un tenseur affine, élément de $\otimes^{(v)}\mathrm{E}$ avec $|v| = p$ détermine un tenseur restreint t. Mais à un tenseur restreint t correspond 2^p tenseurs affines, un pour chacune des variances d'ordre p, qui sont appelés les *représentants de t dans les diverses variances*. Dans une base quelconque e de E, *il faut*, pour déterminer le tenseur restreint t, choisir une variance (v), qui permet d'écrire *les composantes de variance* (v) *de t dans la base e*. Si l'on connait les composantes de t d'une certaine variance (v), on obtient ses composantes d'une autre variance (w) en contractant avec g chaque indice qui de contravariant devient covariant, avec $\overset{-1}{g}$ chaque indice qui de covariant devient contravariant.

Il en est ainsi, en effet, pour les tenseurs décomposables, puisque, si $x \in \mathrm{E}$, $\rho x = g \cdot x$ et si $\alpha \in \mathrm{E}^*$, $\overset{-1}{\rho}\,\alpha = \overset{-1}{g} \cdot \alpha$. Cette opération de changement de variance est souvent décrite comme *descente ou montée des indices*.

Par exemple, les composantes de variance $(* * .)$: t_{ij}^{k} d'un tenseur restreint t dont on connait les composantes de variance $(. * *)$, sont :

$$t_{ij}^{k} = \sum g_{il} g^{km} t_{jm}^{l}.$$

Lorsque la base est orthogonale, le passage des composantes d'un tenseur d'une variance à une autre consiste en de simples multiplications par des scalaires. Par exemple :

$$t_{ij}^{k} = g_{ii} g^{kk} t_{jk}^{i} \quad \text{(pas de sommation !).}$$

Si le tenseur métrique est de type euclidien et la base e orthonormale, et *seulement dans ce cas* les composantes dans e d'un tenseur restreint t sont les mêmes quelle que soit la variance, puisque les matrices G et $\overset{-1}{\mathrm{G}}$ sont les matrices unités.

Proposition V.1

Le tenseur métrique g, son inverse $\overset{-1}{g}$, les tenseurs $\mathrm{I} \in \mathrm{E}\otimes\mathrm{E}^*$ et $\mathrm{I}^* \in \mathrm{E}^* \otimes \mathrm{E}$ sont les quatre représentants de diverses variances du même tenseur restreint, usuellement appelé tenseur fondamental.

Preuve : Les identifications $\rho_{(v)(u)}$ s'expriment par les équivalences :

$$\begin{aligned}\mathrm{I} = \sum e_j \otimes e^{*j} &\equiv \sum \rho e_j \otimes e^{*j} = g = \sum g_{ij} e^{*i} \otimes e^{*j} \\ &\equiv \sum e_j \otimes \overset{-1}{\rho}\, e^{*j} = \overset{-1}{g} = \sum gij e_i \otimes e_j.\end{aligned}$$

Si X et Y sont les matrices colonnes des composantes de x et de y dans la base e : $g(x,y) = \sum x^i g_{ij} y^j$ s'écrit matriciellement ${}^t\mathrm{XGY}$. g est invariant par l'opérateur a, de matrice A dans e, si et seulement si ${}^t\mathrm{X}{}^t\mathrm{AGAY} = {}^t\mathrm{XGY}$ quels que soient x et y, soient ${}^t\mathrm{AGA} = \mathrm{G}$.

Il en résulte $(\text{dét}\, a)^2 = (\text{dét A})^2 = 1$, soit $\text{dét}\, a = \pm 1$.

Les opérateurs orthogonaux de déterminant $+1$ forment un sous-groupe, dit *spécial orthogonal* : $\mathrm{SO}(\mathrm{E}; g)$, de $\mathrm{O}(\mathrm{E}; g)$.

V.2 – Opérations tensorielles sur les tenseurs restreints

1) Permutation des facteurs (ou des indices). Une remarque essentielle est que l'ordre de tous les indices des composantes d'un tenseur restreint dans une variance donnée doit être rigoureusement préservé ; contrairement au cas des tenseurs affines où l'on ne pouvait passer d'une variance à une autre. Ainsi, par exemple, il faut noter $a^i_{\ j}$, avec des indices décalés, les composantes d'un opérateur linéaire $a = \sum a^i_{\ j} e_i \otimes e^{*j}$, où $a^i_{\ j} = a(e_j)^i$, celles de son transposé ${}^t a$, qui composent la matrice transposée, étant notées $a_j^{\ i} = a^i_{\ j}$, soit : ${}^t a = \sum a_j^{\ i} e^{*j} \otimes e_i$. Les tenseurs restreints qu'ils définissent sont *distincts*. Leurs représentant deux fois covariants sont des formes bilinéaires transposées l'une de l'autre b et ${}^t b$ avec $b_{kj} = \sum g_{ki} \cdot a^i_{\ j}$

et $({}^tb)k_j = \sum a_k^i g_{ij}$ d'où : $({}^tb)_{jk} = b_{jk}$. En effet :

$$\begin{aligned} a &= \sum a_j^i e_i \otimes e^{*j} \equiv \sum a_j^i \rho e_i \otimes e^{*j} = \sum a_j^i g_{ik} e^{*k} \otimes e^{*j} = b \\ {}^t a &= \sum a_j{}^i e^{*j} \otimes e_i \equiv \sum a_j{}^i e^{*j} \otimes \rho e_i = \sum a_j{}^i g_{ik} e^{*j} \otimes e^{*k} = {}^t b. \end{aligned}$$

Proposition V.2 :

Soit a un opérateur linéaire de E. Son transposé ${}^ta \in \mathfrak{L}(\mathrm{E}^*)$ et son adjoint a^* sont deux représentants du même tenseur restreint, et $a_l^{*k} = \sum g^{kj} a_j^i g_{il}$.

Preuve : On a : $a^* = \overset{-1}{\rho}\,{}^ta\rho$, double contracté $\overset{-1}{g} \cdot {}^ta \cdot g$ de ${}^ta$. Entre les matrices des composantes dans une base, A et A*, on a donc la relation $\mathrm{A}^* = \overset{-1}{\mathrm{G}}\,{}^t\mathrm{AG}$, ce que l'on peut calculer directement :

$$\begin{aligned} (ax \mid y) &= (\textstyle\sum a_j^i x^j e_i \mid y) = \sum a_j^i x^j y_j = \sum a_j^i x^j y_i \\ &= \textstyle\sum a_j^i g^{jk} x_k g_{il} y^l \\ (x \mid a^* y) &= (x \mid \textstyle\sum a^{*k}{}_l y^l e_k) = \sum a^{*k}{}_l x_k y^l \text{ d'où la relation.} \end{aligned}$$

Soient maintenant $t \in \otimes^{(v)}\mathrm{E}$ et $t' = \rho_{(v)(u)} t \in \otimes^{(v)}\mathrm{E}$ des représentants d'un tenseur restreint $r \in \mathrm{T}_p(\mathrm{E})$. Une permutation $\sigma \in \mathfrak{S}_p$ de $(1, 2, \ldots, p)$ opère sur chacun des symboles de variances (u) et (v), qui deviennent (σu) et (σv) suivant le principe exposé au §III.3 (σ remplace les indices par $(\overset{-1}{\sigma}(1), \overset{-1}{\sigma}(2), \ldots, \overset{-1}{\sigma}(p))$, et sur les facteurs des produits tensoriels de la même façon.

Les tenseurs $\sigma t'$ et σt sont donc reliés par :
$\sigma t' = \rho_{(\sigma v)(\sigma u)} \sigma t$ et sont ainsi les représentants d'un même tenseur restreint que l'on note σr.

Le groupe symétrique $\mathfrak{S}_p$ et son algèbre $\mathrm{K}(\mathfrak{S}_p)$ opèrent donc de façon naturelle dans l'espace vectoriel $\mathrm{T}_p(\mathrm{E})$.

Le quotient de $\mathrm{T}_p(\mathrm{E})$ par le sous-espace d'antisymétrisation est la puissance extérieure $\wedge^p \mathrm{T}_1(\mathrm{E})$ qui s'identifie à $\wedge^p\mathrm{E}$ et à $\wedge^p\mathrm{E}^*$, l'identification entre ces derniers résultant des isomorphismes $\overline{\wedge}^p \rho$ et $\overline{\wedge}^p \overset{-1}{\rho}$.

De même, le quotient de $\mathrm{T}_p(\mathrm{E})$ par le sous-espace de symétrisation s'identifie à $\vee^p\mathrm{E}$ et à $\vee^p\mathrm{E}^*$, les puissances symétriques p-ème de ρ et $\overset{-1}{\rho}$: $\overline{\vee}^p\rho$ et $\overline{\vee}^p\,\overset{-1}{\rho}$ identifiant ces deux espaces.

2) *Action du groupe orthogonal sur les tenseurs restreints.* Un opérateur linéaire inversible a de E opère sur une puissance tensorielle $\otimes^{(v)}\mathrm{E}$ par l'action de a sur les facteurs E et celle du contragrédient $\breve{a}$ sur les facteurs E^*. On a vu que a est orthogonal si et seulement s'il commute avec ρ : $\rho a = \breve{a}\rho$, soit aussi $a\,\overset{-1}{\rho} = \overset{-1}{\rho}\,\breve{a}$. Il en résulte que les opérations du groupe orthogonal $\mathrm{O}(\mathrm{E}; g)$ commutent avec les isomorphismes $\rho_{(v)(u)}$ et définissent ainsi les opérations de $\mathrm{O}(\mathrm{E}; g)$ sur les classes d'équivalence des tenseurs d'un même ordre p par les $\rho_{(v)(u)}$, c'est-à-dire sur l'espace vectoriel des tenseurs restreints $\mathrm{T}_p(\mathrm{E})$. On peut aussi exprimer cette propriété en disant que puisque les représentations de $\mathrm{O}(\mathrm{E}; g)$ dans E et dans E^* sont isomorphes, toutes les représentations de $\mathrm{O}(\mathrm{E}; g)$ dans les $\otimes^{(v)}\mathrm{E}$ d'un même ordre $|v| = p$ sont isomorphes, d'où par passage au quotient, la représentation naturelle de $\mathrm{O}(\mathrm{E}; g)$ dans $\mathrm{T}_p(\mathrm{E})$.

$\mathrm{O}(\mathrm{E}; g)$ *est le groupe de structure des tenseurs restreints de* $(\mathrm{E}; g)$.

L'action de $\mathrm{O}(\mathrm{E}; g)$ sur $\mathrm{T}_p(\mathrm{E})$ commute avec celle du groupe symétrique $\mathfrak{S}_p$ et de son algèbre $\mathrm{K}(\mathfrak{S}_p)$. Les sous-espaces des tenseurs symétriques et antisymétriques de $\mathrm{T}_p(\mathrm{E})$ sont donc stables par $\mathrm{O}(\mathrm{E}; g)$: ce sont des sous-$(\mathrm{O}(\mathrm{E}; g))$-modules de $\mathrm{T}(\mathrm{E})$. Le tenseur g est invariant par $\mathrm{O}(\mathrm{E}; g)$.

3) La *multiplication tensorielle* commute avec les isomorphismes $\rho_{(v)(u)}$:

$$\begin{tikzcd}
\otimes^{(u)}\mathrm{E} \arrow[d, "\rho_{(v)(u)}"'] & \times & \otimes^{(u')}\mathrm{E} \arrow[d, "\rho_{(v')(u')}"] \arrow[rr] & & \otimes^{(uu')}\mathrm{E} \arrow[d, "\rho_{(vv')(uu')}"] \\
\otimes^{(v)}\mathrm{E} & \times & \otimes^{(v')}\mathrm{E} \arrow[rr] & & \otimes^{(vv')}\mathrm{E}
\end{tikzcd}$$

$$\rho_{(v)(u)}\,t \otimes \rho_{(v')(u')}\,t' = \rho_{(vv')(uu')}\,t \otimes t'$$

ainsi qu'avec les opérations de $\mathrm{O}(\mathrm{E}; g)$ et détermine une multiplication tensorielle entre tenseurs restreints

$$\mathrm{T}_p(\mathrm{E}) \times \mathrm{T}_{p'}(\mathrm{E}) \to \mathrm{T}_{p+p'}(\mathrm{E})$$

qui commute avec les opérations de $\mathrm{O}(\mathrm{E};g)$ (multiplication de $\mathrm{O}(\mathrm{E};g)$-modules).

4) la *contraction* « sur deux indices » quelconques, ou entre deux facteurs quelconques d'un tenseur restreint s'exprime sur un représentant de variance (v) par le crochet de dualité si les indices sont des variances opposés, par le produit scalaire si les variances des indices sont les mêmes.

En effet, on a la diagramme commutatif suivant où les flèches verticales sont les contractions :

$$\begin{array}{ccccccc}
\mathrm{E}\otimes\mathrm{E} & \xrightarrow{p\otimes \mathrm{Id}} & \mathrm{E}^*\otimes\mathrm{E} & \xrightarrow{p^{-1}\otimes \mathrm{Id}} & ,\mathrm{E}\otimes\mathrm{E}^* & \xrightarrow{p\otimes \mathrm{Id}} & \mathrm{E}^*\otimes\mathrm{E}^* \\
x\otimes y & \equiv & \rho x\otimes y & \equiv & x\otimes\rho y & \equiv & \rho x\otimes\rho y \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
(x|y) & = & \langle\rho x, y\rangle & = & \langle x,\rho y\rangle & = & (\rho x|\rho y)
\end{array}$$

Si par exemple un tenseur restreint t d'ordre p est donné par ses composantes de variance (v), le tenseur restreint τ, d'ordre $(p-2)$, contracté de t sur les deux premiers indices, a pour composantes de variances $(w) = (v$ moins les deux premiers signes) suivant le cas :

$$\tau^{\cdot\cdot}_{\cdot\cdot} = \sum t^{i\cdot\cdot}_{i\cdot\cdot} = \sum t_{i\cdot\cdot}{}^{i\cdot\cdot} = \sum g_{ij}t^{ij\cdots}{}_{\cdots} = \sum g^{ij}t^{\cdots}{}_{ij\cdots}$$

Autrement dit, la contraction sur deux indices de même variance s'effectue par la contraction de cette paire d'indices avec le tenseur métrique.

Le contracté de g :

$$\sum g^{ij}(e_i \mid e_j) = \mathrm{Tr}\,\overset{-1}{\mathrm{G}}\,\mathrm{G} = n = \dim\mathrm{E}.$$

Dans le cas d'un tenseur d'ordre deux, ses représentants dans $\mathrm{E}\otimes\mathrm{E}^* = \mathcal{L}(\mathrm{E})$ et dans $\mathrm{E}^*\otimes\mathrm{E} = \mathcal{L}(\mathrm{E}^*)$ sont des opérateurs linéaires a et ${}^t a$ de composantes dans une base a^i_j et $a_j{}^i$, et la contraction est alors la *trace* de ces opérateurs $\mathrm{Tr}\,a = \mathrm{Tr}\,{}^t a = \sum a^i_i$.

C'est pourquoi même pour les tenseurs d'ordre supérieur, les contractions sur les diverses paires d'indices sont souvent appelées des traces.

Le diagramme précédent montre que ces contractions ou traces commutent avec les opérations du groupe orthogonal.

Chacune d'elles est donc un *homomorphisme de* $\mathrm{O}(\mathrm{E}; g)$*-modules de* $\mathrm{T}_p(\mathrm{E})$ *dans* $\mathrm{T}_{p-2}(\mathrm{E})$, dont le noyau est un sous-$\mathrm{O}(\mathrm{E}; g)$-module de $\mathrm{T}_p(\mathrm{E})$. (cf. §IV.4).

5) Nous terminerons ce paragraphe par une remarque évidente mais importante. Lors d'un changement de base quelconque : $e' = e\mathrm{S}$, les formules de changement des composantes de variance (v) de la base e à la base e' restent évidemment *les mêmes* que s'il s'agissait d'un tenseur affine. Rien n'est changé ! L'observation essentielle est que les identifications $\rho_{(\bar{v})(u)}$ entre représentants de diverses variances d'un tenseur restreint, sont définies intrinsèquement, indépendamment du choix de toute base. Elles commutent avec les transformations de composantes déterminées par un changement de base, qui peuvent être effectués sur les composantes de n'importe quelle variance au choix.

Exemple : vérifier ce qui vient d'être dit pour les composantes contravariantes et covariantes d'un vecteur.

Il faut cependant noter, et c'est cela qui peut être à l'origine de certaines ambiguïtés, que pour les tenseurs affines, les opérations du groupe linéaire complet $\mathrm{Gl}(\mathrm{E})$ et les changements de base de E jouaient des rôles équivalents. La situation est différente pour les tenseurs restreints : les changements de bases restent arbitraires tandis que seules les opérations du groupe de structure $\mathrm{O}(\mathrm{E}; g)$ ont un sens sur les tenseurs restreints.

On a ainsi trois catégories de formules : passage d'une base à une autre, d'une variance à une autre, transformation par un opérateur orthogonal. En mathématique appliquée, en mécanique par exemple, on se garde bien de compliquer la situation en prenant des bases quelconques. On se limite aux bases orthonormales donc aux changements de base obtenus par des opérateurs orthogonaux. Cela rétablit l'équivalence entre changements de base et action du groupe de structure mais ce n'est sûrement pas de nature à clarifier les concepts.

V.3 – Pseudoscalaires et pseudotenseurs restreints

Ainsi qu'on l'a vu à la fin du §IV.3, si $\dim \mathrm{E} = n$, les droites $\wedge^n \mathrm{E}$ et $\wedge^n \mathrm{E}^*$ s'identifient aux droites des pseudoscalaires contravariants et covariants introduites au §III.6.

L'isomorphisme ρ de E sur E*, défini par la métrique, détermine un isomorphisme $\overline{\wedge}^n\rho$ qui identifie les pseudoscalaires contravariants et covariants en un seul type de *pseudoscalaire restreint* : p, qui a une composante contravariante p_e et une composante covariante p_{e^*} dans chaque base e. On passe de l'une à l'autre par :
$p = p_e e_1 \wedge e_2 \wedge \dots e_n \equiv p_e(\rho e_1 \wedge \dots \rho e_n) = p_e \cdot \overline{g} \cdot e^{*1} \wedge e^{*2} \wedge \dots e^{*n}$
où $\overline{g} = \text{dét G} = \text{dét } |g_{ij}| = \text{dét}_e{}^{e^*} |\rho|$, soit :

$$\boxed{p_{e^*} = p_e \cdot \overline{g}}.$$

L'effet d'un changement de base $e' = e\text{S}$ sur chacune de ces composantes est, comme on l'a vu : $p_e e_1 \wedge e_2 \wedge \dots e_n = p_{e'} \wedge e'_2 \wedge \dots e'_n = p_e \text{ dét S} \cdot e_1 \wedge e_2 \wedge \dots e_n$ soit $p_{e'} = p_e(\text{dét}^{-1}\text{S})$ et $p^*_{e'} = p^*_e(\text{dét S})$.

On a $p^*_{e'} = p_{e'}\overline{g} = p_e(\text{dét}^{-1}\text{S})\overline{g}'$, tandis que la relation $\text{G}' ={}^t \text{SGS}$ donne $\overline{g}' = (\text{dét S})^2 \cdot \overline{g}$.

Enfin, si a est un opérateur linéaire de E, et $\check{a} = {}^t \overset{-1}{a}$ son contragrédient, on a $\text{dét}\,\check{a} = (\text{dét}^{-1} a)$. Si a est orthogonal, $\text{dét}\, a = \pm 1$ d'où il résulte que $\text{dét } \check{a} = \text{dét } a$. Cela revient au fait que les représentations du groupe orthogonal $\text{O}(\text{E}; g)$ dans $\wedge^n\text{E}$ et $\wedge^n\text{E}^*$ sont isomorphes. Les éléments de $\text{SO}(\text{E}; g)$ opèrent comme l'identité, ceux de déterminant (-1) par retournement.

Plus généralement, un *pseudoscalaire restreint d'ordre* r est, suivant la définition III.6 un élément de la droite $\text{D}^r = \otimes^r(\wedge^n\text{E})$ identifiée par l'isomorphisme $\bar{\otimes}^r(\overline{\wedge}^n\rho)$ à la droite $\text{D}^{*r} = \otimes^r(\wedge^n\text{E}^*)$. Un tel élément q possède donc une composante contravariante q_e et une composante covariante q_{e^*} dans chaque base e. Si l'on note $\overline{e} = e_1 \wedge e_2 \wedge \dots e_n$ et $\overline{e}^* = e^{*1} \wedge e^{*2} \wedge \dots e^{*n}$ les bases de $\wedge^n\text{E}$ et $\wedge^n\text{E}^*$ associées à e, avec, comme ci-dessus $\overline{\wedge}^n\rho \cdot \overline{e} = \overline{g} \cdot \overline{e}^*$, on a $\bar{\otimes}^r(\overline{\wedge}^n\rho)\overline{e} = \otimes^r\overline{g}\overline{e}^{*r}$ d'où la relation entre composantes de q : $q_{e^*} = \overline{g}^r q_e$.

Par un changement de base $e' = e\text{S}$ on a vu au §III.6 que $q_{e'} = q_e(\text{dét S})^{-r}$ et $q_{e'}{}^* = q_{e^*}(\text{dét S})^r$ tandis que $\overline{g}'^r = (\text{dét S})^{2r}\overline{g}^r$.

Nous noterons Q^r la droite des pseudoscalaires restreints d'ordre r. D'après ce qui a été dit plus haut le groupe $\text{O}(\text{E}; g)$ opère sur Q^r par la représentation : $u \longrightarrow (\text{dét } u)^r$, donc par la représentation identique si r est pair.

L'identification des tenseurs de même ordre et de variances diverses par les isomorphismes $\rho_{(v)(u)}$ et celle des pseudoscalaires de même ordre r par les $\bar{\otimes}^r(\overline{\wedge}^n\rho)$ définit, suivant la définition III.7, les pseudotenseurs restreints (par g) dont on peut choisir les représentants donc les composantes de la variance scalaire que l'on veut.

Définition V.3.A :

L'espace des *pseudotenseurs restreints* d'ordre p, d'ordre pseudoscalaire r est le produit tensoriel $\mathrm{T}_p(\mathrm{E}) \otimes \mathrm{Q}^r$ de l'espace des tenseurs restreints d'ordre p par la droite des pseudoscalaires restreints d'ordre r.

Le groupe $\mathrm{O}(\mathrm{E}; g)$ y opère par le produit tensoriel de ses représentations sur les facteurs. Cette représentation est identique à la représentation de $\mathrm{O}(\mathrm{E}; g)$ dans $\mathrm{T}_p(\mathrm{E})$ si r est pair. Si r est impair, c'est la représentation associée suivant la définition suivante :

Définition V.3.B :

Si μ est une représentation linéaire du groupe orthogonal $\mathrm{O}(\mathrm{E}; g)$ dans un espace vectoriel de dimension finie F, il lui correspond une autre représentation μ' dans F, dite *associée* ainsi définie : $\mu'(a) = (\text{dét } a) \cdot \mu(a)$, μ' est donc le produit tensoriel de μ et de la représentation de $\mathrm{O}(\mathrm{E}; g)$ dans K : $a \to \text{dét } a$.

Exemple : les pseudotenseurs ε et ε^* de Lévi-Civita définies au §III.7 sont les représentants d'*un même pseudotenseur restreint*.

Les formules de passage des composantes de l'un aux composantes de l'autre (descente ou montée des indices) sont simplement des expressions des développements des déterminants $\overline{g} = \text{dét}|g_{ij}|$ et $\overset{-1}{\overline{g}} = \text{dét}\,|g^{ij}|$:

$$\varepsilon_{i_1 i_2 \ldots i_n} = \overset{-1}{\overline{g}} \cdot \sum g_{i_1 j_1} g_{i_2 j_2} \cdots g_{i_n j_n} \cdots \varepsilon^{j_1 j_2 \cdots j_n}$$
$$\varepsilon^{i_1 i_2 \ldots i_n} = \overline{g} \cdot \sum g^{i_1 j_1} g^{i_2 j_2} \ldots g^{i_n j_n} \varepsilon_{j_1 j_2 \cdots j_n}.$$

Le choix d'une unité π des pseudoscalaires détermine une bijection entre pseudotenseurs restreints et tenseurs restreints de même ordre p :

$$\mathrm{T}_p(\mathrm{E}) \otimes \mathrm{Q}^r \leftrightarrow \mathrm{T}_p(\mathrm{E}) \quad \text{par} \quad t \otimes \pi^r \leftrightarrow t.$$

Cette bijection est donc déterminée à un scalaire près. Elle peut l'être de façon beaucoup plus précise au signe près dépendant de l'orientation lorsque E est un espace vectoriel *réel*, dont la métrique g est euclidienne ou pseudoeuclidienne, cette dernière *d'indice* (r, s).

Définition V.3.C :

Sur un espace euclidien ou pseudoeuclidien (E, g) orienté, *l'unité de volume orienté*, qui peut également être considérée comme *unité des scalaires orientés*, est l'élément ω de $\wedge^n \mathrm{E} \equiv \wedge^n \mathrm{E}^*$ défini dans toute base orienté positivement e de E par :

$$\begin{aligned}\omega &= |\overline{g}|^{-1/2}\overline{e} = |\overline{g}|^{-1/2} e_1 \wedge e_2 \wedge \dots e_n \equiv |\overline{g}|^{1/2}\overline{e}^* \\ &= |\overline{g}|^{1/2} e^{*1} \wedge e^{*2} \wedge \dots e^{*n}\end{aligned}$$

où $\overline{g} = \text{dét}|g_{ij}|$ et où le signe $\equiv$ indique que les éléments sont images l'un de l'autre par l'isomorphisme $\wedge^n \rho$, puisque $(\wedge^n \rho)\overline{e} = \overline{ge}^*$; ω est de carré scalaire égal à $(-1)^3$: $(\omega \mid \omega) = |\overline{g}|^{-1}(\overline{e} \mid \overline{e}) = |\overline{g}|^{-1}\overline{g}$ $=$ signe de $\overline{g} = (-1)^s$. Pour toute base e orthonormale, positive $\omega = \overline{e} = e_1 \wedge e_2 \wedge \dots e_n$.

ω est invariant par tout changement de base $e' = e\mathrm{S}$ de déterminant positif, et est transformée en son opposé si $\text{dét}\,\mathrm{S} < 0$. En effet $\overline{g}' = (\text{dét}\,\mathrm{S})^2\overline{g}$, et $\overline{e}' = \text{dét}\,\mathrm{S} \cdot \overline{e}$ d'où $|\overline{g}'|^{-1/2}\overline{e}' = |\overline{g}|^{-1/2}|\text{dét}\,\mathrm{S}|^{-1} \cdot \text{dét}\,\mathrm{S}\overline{e} = \text{signe}\,|\text{dét}\,\mathrm{S})|\overline{g}|^{-1/2}\overline{e}$. E étant orienté (§III.8) on peut choisir l'élément ω comme *base canonique* de $\wedge^n \mathrm{E} \equiv \wedge^n \mathrm{E}^*$, c'est-à-dire comme base des pseudoscalaires restreints. Dans cette base les composantes contravariante et covariante d'un pseudoscalaire restreint sont égales.

ω peut être exprimée à l'aide des pseudotenseurs ε de Levi-Civita (§III.7). Les composantes de ω dans une base e sont :

– composantes covariantes : $\omega_{i_1 i_2 \dots i_n} = |\overline{g}|^{1/2}\varepsilon_{i_1 i_2 \dots i_n}$

– composantes contravariantes : $\omega^{i_1 i_2 \dots i_n} = |\overline{g}|^{-1/2}\varepsilon^{i_1 i_2 \dots i_n}$

$\omega \otimes \omega$ est un invariant absolu : $\omega \otimes \omega = \overline{ge}^* \otimes \overline{e}^* = \overset{-1}{\overline{g}}\,\overline{e} \otimes \overline{e}$.

Rappelons à titre d'exemple les formules de changement de base pour un pseudovecteur. Soit $x \in \mathrm{T}_1(\mathrm{E}) \otimes \mathrm{Q}^r$. Ses composantes x^j contravariantes et de variance scalaire contravariante dans une base e de E sont définies par : $x = x^j e_j \otimes \overline{e}^r$. On peut les rassembler en une matrice colonne X, soit $x =$

$e \otimes \overline{e}^r \cdot \mathrm{X}$, et dans une autre base $e' = e\mathrm{S}$, $x = e' \otimes \overline{e}'^r \cdot \mathrm{X}'$. Puisqu'alors $\overline{e}'^r = (\text{dét S})^r \overline{e}^r$, on a : $x = e' \otimes \overline{e}'^r \cdot \mathrm{X}' = e \otimes \overline{e}^r \cdot \mathrm{SX}'(\text{dét S})^r$ d'où, ${}^tX' = \check{\mathrm{S}}{}^t\mathrm{X}(\text{dét}\,\mathrm{S})^{-r}$ qui s'écrit :

$$x'^k = (\text{dét}\,\mathrm{S})^{-r} \sum \check{s}_{k,j} x^j.$$

Pour la même contravariance scalaire, les composantes covariantes de x dans e sont $x_j = \sum g_{jk} x^k$.

Si l'on choisit maintenant des composantes scalairement covariantes x_*^j ou x_{*j}, puisque $\overline{e}^r = \overline{g}^r \cdot \overline{e}_*^r$ on les obtient simplement par $x_*^j = \overline{g}^r x^j$ et $x_{*j} = \overline{g}^r x^j$.

V.4 – Extension de la métrique aux espaces tensoriels et pseudo-tensoriels

Suivant les §II.11 et IV.11, la forme bilinéaire symétrique non-dégénérée g sur E se prolonge en les formes bilinéaires symétriques non-dégénérées $\bar{\otimes}^p g$, $\bar{\otimes}^p \overset{-1}{g} \bar{\otimes}^{(v)} g$, $\overline{\wedge}^p g$, $\overline{\vee}^p g$, qui sont des produits scalaires sur chacun des espaces concernés. Par exemple, sur des éléments décomposables, on a :

$$\begin{aligned}(x_1 \otimes x_2 \otimes \dots x_p \mid y_1 \otimes y_2 \otimes \dots y_p)&\\ &= (x_1 \mid y_1)(x_2 \mid y_2)\dots(x_p \mid y_p)\\ &= \langle x_1; \rho y_1\rangle\langle x_2; \rho y_2\rangle \dots \langle x_p; \rho y_p\rangle\\ &= \langle x_1 \otimes x_2 \otimes \dots x_p; (\bar{\otimes}^p \rho) \cdot y_1 \otimes y_2 \otimes \dots y_p\rangle\end{aligned}$$

dont l'application associée de $\otimes^p\mathrm{E}$ dans $\otimes^p\mathrm{E}^*$ est donc $\bar{\otimes}^p \rho$.

$$\begin{aligned}(x_1 \wedge x_2 \wedge \dots x_p \mid y_1 \wedge y_2 \wedge \dots y_p) &= \text{dét}|(x_i \mid y_i)| = \text{dét}|\langle x_i; \rho y_j\rangle)|\\ &= \langle x_1 \wedge x_2 \wedge \dots x_p; \overline{\wedge}^p \rho \cdot (y_1 \wedge y_2 \wedge \dots y_p)\rangle\end{aligned}$$

dont l'application associée de $\wedge^p\mathrm{E}$ dans $\wedge^p\mathrm{E}^*$ est donc $\overline{\wedge}^p \rho$.

En particulier, si $e = (e_1, e_2, \dots, e_n)$ est une base de E :

$$(e_1 \wedge e_2 \wedge \dots e_n \mid e_1 \wedge e_2 \wedge \dots e_n) = \text{dét}|(e_i \mid e_j)| = \text{dét}|g_{ij}| = \overline{g}$$

Puisque ρ et $\overset{-1}{\rho}$ sont des isomorphismes entre les espaces métriques (E, g), $(\mathrm{E}^*, \overset{-1}{g})$, les $\rho_{(v)(u)}$ entre $\otimes^{(u)}\mathrm{E}$ et $\otimes^{(v)}\mathrm{E}$, $\overline{\wedge}^p\rho$ entre $\wedge^p\mathrm{E}$ et $\wedge^p\mathrm{E}^*$, $\overline{\vee}^p\rho$ entre $\vee^p\mathrm{E}$ et $\vee^p\mathrm{E}^*$ sont également des isomorphismes d'espaces métriques, ce qui définit le produit scalaire et la structure d'espace métrique sur les espaces de tenseurs restreints $\mathrm{T}_p(\mathrm{E})$, $\wedge^p\mathrm{T}_1(\mathrm{E})$, $\vee^p\mathrm{T}_1(\mathrm{E})$.

Le produit scalaire s'étend également aux espaces de pseudotenseurs $(\otimes^{(v)}\mathrm{E})\otimes(\otimes^r(\wedge^n\mathrm{E}))$ et $(\otimes^{(v)}\mathrm{E})\otimes(\otimes^r(\wedge^n\mathrm{E}^*))$ et par le même procédé aux espaces de pseudotenseurs restreints $\mathrm{T}_p(\mathrm{E})\otimes\mathrm{Q}^r$. Ecrivons par exemple le produit scalaire de deux *pseudovecteurs* à l'aide de leurs composantes contravariantes et scalairement contravariantes :

$$\begin{aligned}(x \mid y) &= \textstyle\sum x^j y^k (e_j \otimes \overline{e}^r \mid e_k \otimes \overline{e}^r) = \overline{g}^r \sum g_{jk} x^j y^k \\ &= \overline{g}^r . {}^t\mathrm{XGY}.\end{aligned}$$

On peut en vérifier le caractère intrinsèque par son invariance par un changement de bases $e' = e\mathrm{S}$. On a vu au paragraphe précédent que $\mathrm{X} = \mathrm{SX}'(\det \mathrm{S})^r$, $\mathrm{G}' ={}^t \mathrm{SGS}$ et $\overline{g}' = (\det \mathrm{S})^2\overline{g}$, d'où :

$$\begin{aligned}(x \mid y) &= (\text{dét}\,\mathrm{S})^{-2r}\overline{g}'^r(\text{dét}\,\mathrm{S})^{rt}\mathrm{X}'^t\mathrm{SGSY}'(\text{dét}\,\mathrm{S})^r \\ &= \overline{g}'^{rt}\mathrm{X}'\mathrm{G}'\mathrm{Y}', \text{ cqfd.}\end{aligned}$$

Remarques

1) L'expression de ces produits tensoriels et leur caractère multilinéaire entraîne, par exemple, que si $t, t' \in \otimes^p\mathrm{E}$ et $u, u' \in \otimes^q\mathrm{E}$, on a $(t\otimes u \mid t'\otimes u') = (t \mid t')(u \mid u')$ ce qui permet de définir le produit scalaire partiel, qui n'est autre qu'une contraction, de $t \otimes u$ par t', égal à $(t \mid t')u$. Ce procédé a déjà été utilisé dans le cas de l'algèbre extérieure au §IV.

2) On a vu aux §III.7 et IV.10 la définition du produit vectoriel de $(n-1)$vecteurs : $x_1, x_2, \dots, x_{n-1}$ de E. C'est le pseudocovecteur z obtenu par contraction des $(n-1)$ derniers indices du pseudotenseur ε de Levi-Civita par $(x_1, x_2, \dots, x_{n-1})$: $z_j = \sum \varepsilon_{ji_1i_2\dots i_{n-1}} x_1^{i_1} x_2^{i_2} \dots x_{n-1}^{i_{n-1}}$.

Si l'espace E est muni de la métrique g, il correspond à z un vecteur de E définir à un scalaire près. La droite portant ce vecteur qui est, elle, intrinsèquement déterminée, est alors orthogonale à $x_1, x_2, \dots$ et x_{n-1}.

En effet, pour chaque $k = 1, 2, \dots, n-1$, le produit scalaire de x_k et de z, $\sum x_k^j$ est le déterminant dans la base e de n vecteurs $(x_k, x_1, x_2, \dots, x_{n-1})$ dont deux sont égaux, et est donc nul.

Proposition V.4.A :

Les opérations du groupe orthogonal $\mathrm{O}(\mathrm{E}; g)$ sur les espaces de tenseurs et de pseudotenseurs sur E, conservent la métrique, extension de g, sur ces espaces.

Preuve : Le produit scalaire de deux tenseurs ou pseudotenseurs défini par l'extension de g est une somme de produits de produits scalaires de vecteurs de E qui sont invariants par les opérations de $\mathrm{O}(\mathrm{E}; g)$.

Proposition V.4.B :

Les opérations du groupe symétrique $\mathfrak{S}_p$ sur les espaces tensoriels $\otimes^p\mathrm{E}$, $\otimes^p\mathrm{E}^*$, et $\mathrm{T}_p(\mathrm{E})$ conservent le produit scalaire extension de g.

Preuve : Il suffit de vérifier que sur des tenseurs décomposables t, u on a quel que soit $\sigma \in \mathfrak{S}_p$: $(\sigma t \mid \sigma u) = (t \mid u)$, soit :

$$\begin{aligned}(\sigma \cdot x_1 \otimes \dots x_p \mid \sigma \cdot y_1 \otimes \dots y_p) &= (x_{\overset{-1}{\sigma}1} \otimes \dots x_{\overset{-1}{\sigma}p} \mid y_{\overset{-1}{\sigma}1} \otimes \dots y_{\overset{-1}{\sigma}p}) \\ &= (x_{\overset{-1}{\sigma}1} \mid y_{\overset{-1}{\sigma}p}) \dots (x_{\overset{-1}{\sigma}1} \mid y_{\overset{-1}{\sigma}p}) \\ &= (x_1 \mid y_1) \dots (x_p \mid y_p) \quad \text{cqfd.}\end{aligned}$$

Corollaire V.4 : Soit τ_{ij} l'opérateur de permutation des facteurs de rangs i et j de $\mathrm{T}_p(\mathrm{E})$. Tout tenseur (ij)-symétrique est orthogonal à tout tenseur (ij)-antisymétrique. Il en résulte que le sous-espace des tenseurs antisymétriques $\mathrm{A}_p(\mathrm{E})$ est orthogonal au sous-espace engendré par les tenseurs présentant une symétrie partielle. De même le sous-espace des tenseurs symétriques $\mathrm{S}_p(\mathrm{E})$ est orthogonal au sous-espace engendré par les tenseurs présentant une antisymétrie partielle.

Preuve : Puisque le produit scalaire est invariant par permutation des facteurs, $(\tau_{ij}s \mid \tau_{ij}t) = (s \mid t)$.

Si $\tau_{ij}s = s$ et $\tau_{ij}t = -t$ on a donc $(s \mid t) = 0$ cqfd.

Proposition V.4.C :

Le produit scalaire $(x \mid y)$ de deux vecteurs x, y de E, qui est, dans $\mathrm{T}_2(\mathrm{E})$, le contracté de $x \otimes y$, est aussi égal au produit scalaire de $x \otimes y$ et du tenseur métrique g, ou au produit tensoriel doublement contracté de $x \otimes y$ et de g.

Preuve : $(x \otimes y \mid \sum g^{ij} e_i \otimes e_j) = \sum g^{ij}(x \mid e_i)(y \mid e_j)$

$$= \mathrm{C}_{13}\mathrm{C}_{24}(x \otimes y \otimes \textstyle\sum (g^{ij} e_i \otimes e_j)) = \sum g^{ij} x_i y_i = (x \mid y).$$

Proposition V.4.D :

La contraction C_{ij} sur deux indices (i, j) de $\mathrm{T}_p(\mathrm{E})$ dans $\mathrm{T}_{p-2}(\mathrm{E})$ est identique à la contraction de ce couple d'indices par le tenseur métrique g et revient au « produit scalaire partiel » par g. Réciproquement l'injection γ_{12} de $\mathrm{T}_{p-2}(\mathrm{E})$ dans $\mathrm{T}_p(\mathrm{E})$: $u \longrightarrow \dfrac{1}{n}\, g \otimes u$ ($n = \dim \mathrm{E}$) est un inverse à droite de la contraction C_{12} et est une injection de $\mathrm{O}(\mathrm{E}; g)$-modules.

Preuve : La première partie est une application directe de la proposition précédente. La seconde résulte du fait que le contracté de $g = \sum g^{ij} e_i \otimes e_j$ est égal à $n = \dim \mathrm{E}$, d'où $\mathrm{C}_{12}\gamma_{12}(u) = \mathrm{C}_{12}\Big(\dfrac{1}{n}\, g \otimes u\Big) = u$. Si $a \in \mathrm{O}(\mathrm{E}; g)$, $\gamma_{12}(a \cdot u) = \dfrac{1}{n}\, g \otimes a \cdot u = a \cdot \gamma_{12}(u)$ car $a \cdot g = g$.

Naturellement il existe pour chaque contraction c_{ij} du couple d'indices (i, j) une application inverse γ_{ij} de $\mathrm{T}_{p-2}(\mathrm{E})$ dans $\mathrm{T}_p(\mathrm{E})$ analogue à celle qui vient d'être indiquée pour C_{12}. L'écrire de façon intrinsèque par une multiplication tensorielle s'insérant dans les deux indices (i, j) n'est pas faisable avec les notations traditionnelles. On doit donc faire choix d'une base e de E. Dès lors, pour chaque couple d'indices (i, j), et chaque $t \in \mathrm{T}_p(\mathrm{E})$ écrit, par exemple, en composantes contravariantes, on a :

$$\boxed{(\mathrm{C}_{ij}t)^{r_1 \ldots \widehat{r_1} \ldots \widehat{r^j} \ldots r_p} = \sum g_{r_i r_j} t^{r_1 \ldots r_p}}$$

et l'application inverse à droite γ_{ij} de $\mathrm{T}_{p-2}(\mathrm{E})$ dans $\mathrm{T}_p(\mathrm{E})$ est définie par

$$\boxed{(\gamma_{ij}u)^{r_1\ldots r_p} = \frac{1}{n}\, g^{r_i r_j} u^{r_1\ldots\widehat{r_i}\ldots\widehat{r_j}\ldots r_p}}.$$

Définition V.4 :

On appelle *espace des tenseurs de traces nulles d'ordre p* de l'espace métrique $(\mathrm{E}; g)$ le sous-espace $\mathrm{T}^0_p(\mathrm{E})$ de $\mathrm{T}_p(\mathrm{E})$ formé par les éléments t de $\mathrm{T}_p(\mathrm{E})$ annulés par toutes les contractions c_{ij} sur les couples d'indices (i,j) de $(1,2,\ldots,p)$. $\mathrm{T}_{0,p}(\mathrm{E})$ est stable par les opérations du groupe orthogonal $\mathrm{O}(\mathrm{E}; g)$ et du groupe symétrique de permutation des facteurs $\mathfrak{S}_p$.

Proposition V.4.E :

Quels que soient $t \in \mathrm{T}_p(\mathrm{E})$, $u \in \mathrm{T}_{p-2}(\mathrm{E})$, on a $(t \mid \gamma_{ij}u) = \frac{1}{n}\,(c_{ij}t \mid u)$. Il en résulte que le sous-espace des tenseurs de traces nulles $\mathrm{T}^0_p(\mathrm{E})$ est l'orthogonal du sous-espace engendré par les images de $\mathrm{T}_{p-2}(\mathrm{E})$ par les applications γ_{ij}.

Preuve : Ecrivons les produits scalaires en utilisant des composantes covariantes pour t, contravariantes pour u et g :

$$\begin{aligned}(t \mid \gamma_{ij}u) &= \frac{1}{n}\,\sum t_{r_1\ldots r_p} g^{r_i r_j} u^{r_1\ldots\widehat{r_i}\ldots\widehat{r_j}\ldots r_p}\\ &= \frac{1}{n}\,\sum \Big(\sum t_{r_1\ldots r_p} g^{r_i r_j}\Big) u^{r_1\ldots\widehat{r_i}\ldots\widehat{r_j}\ldots r_p}\\ &= \frac{1}{n}\,(\mathrm{C}_{ij}t \mid u).\end{aligned}$$

Exemples :

1) Si un tenseur t est antisymétrique relativement au couple d'indices (i,j) : $\tau_{ij}t = -t$, son contracté par la contraction c_{ij} est nul, puisque cette dernière est symétrique en i et j : $\mathrm{C}_{ij}(\tau_{ij}t) = C_{ij}t = \mathrm{C}_{ij}(-t) = 0$. Le sous-espace $\mathrm{A}_p(\mathrm{E})$ des tenseurs antisymétriques de $\mathrm{T}_p(\mathrm{E})$, qui est stable par $\mathrm{O}(\mathrm{E}; g)$ est donc un sous-espace de $\mathrm{T}_{0,2}(\mathrm{E})$.

2) Pour $p = 2$, $\gamma_{12}(1) = \dfrac{1}{n}\, g$ et $\mathrm{T}_2(\mathrm{E}) = \mathrm{K}\cdot g \oplus \mathrm{T}_2^0(\mathrm{E})$ (somme directe orthogonale).

D'après le corollaire V.4 : $\mathrm{T}_2^0(\mathrm{E}) = \mathrm{A}_2(\mathrm{E}) \oplus \mathrm{S}_2^0(\mathrm{E})$ où $\mathrm{S}_2^0(\mathrm{E})$ est le sous-espace des tenseurs symétriques d'ordre deux de trace nulle. $\mathrm{T}_2(\mathrm{E})$ est de cette façon, décomposé en la somme directe orthogonale de trois représentations irréductibles du groupe orthogonal : $\mathrm{K}\cdot g$, $\mathrm{A}_2(\mathrm{E})$ et $\mathrm{S}_2^0(\mathrm{E})$.

Remarque : Les propositions précédentes montrent comment l'usage des contractions permet des décompositions du $\mathrm{O}(\mathrm{E}; g)$ module $\mathrm{T}_p(\mathrm{E})$ en sous-modules. Mais pour $p \geqslant 3$ ces sous-modules ne sont plus nécessairement irréductibles et des procédés plus fins doivent être utilisés pour poursuivre l'analyse. Le cas des sous-espace de tenseurs antisymétriques peut se traiter simplement comme nous allons le voir au paragraphe suivant.

V.5 – Dualité dans l'algèbre extérieure d'un espace métrique $(\mathrm{E}; g)$ de dimension finie. Irréductibilité des représentations de $\mathrm{O}(\mathrm{E}; g)$ sur les puissances extérieures $\wedge^p\mathrm{E}$. Commutant d'un module

Soient $e = (e_1, e_2, \dots, e_n)$ une base de E, e^* la base duale de E^*. Si ρ est l'isomorphisme de E sur E^* associé au produit scalaire g, l'isomorphisme de $\wedge^p\mathrm{E}$ sur $\wedge^p\mathrm{E}^*$ associé à l'extension de g à $\wedge^p\mathrm{E}$ est la puissance extérieurs p-ème de ρ, $\overline{\wedge}^p\rho$, tandis que l'isomorphisme associé à l'extension de g à $\wedge^p\mathrm{E}^*$ est $\overline{\wedge}^p\overset{-1}{\rho}$. Puisque, quel que soit $a \in \mathrm{O}(\mathrm{E}; g)$, $\rho a = \check{a}\rho$, on a $\overline{\wedge}^p\rho \cdot \overline{\wedge}^p a = \overline{\wedge}^p\check{a} \cdot \overline{\wedge}^p\rho$ ce qui montre que les représentations de $\mathrm{O}(\mathrm{E}; g)$ dans $\wedge^p\mathrm{E}$ et $\wedge^p\mathrm{E}^*$ sont isomorphes : $\overline{\wedge}^p\rho$ et $\overline{\wedge}^p\overset{-1}{\rho}$ sont des isomorphismes de $\mathrm{O}(\mathrm{E}; g)$-modules. $f = \overset{-1}{\rho}\, e^* = (f^1, f^2, \dots, f^n)$ est une base de E, supplémentaire de e : $(e_i \mid f^j) = \langle e_i, \rho f^j)\rangle = \langle e_i, e^{*j}\rangle = \delta_i^j$
$f^* = \rho e = (f_1^*, \dots, f_n^*)$ est une base de E^*, supplémentaire de e^* et duale de f : $\langle f^i, f_j^*\rangle = \langle \overset{-1}{\rho}\, e^{*i}, \rho e_j\rangle = (\overset{-1}{\rho}\, e^{*i} \mid e_j) = \langle e^{*i}, e_j\rangle = \delta_j^i$.
On a : $\overline{f}^* = f_1^* \wedge f_2^* \dots f_n^* = (\overline{\wedge}^p\rho) e_1 \wedge e_2 \wedge \dots e_n = (\overline{\wedge}^n\rho)\overline{e} = \overline{g} e^{*1} \wedge \dots e^{*n} = \overline{g}\cdot\overline{e}^*$. D'où : $\overline{f}^* = \overline{g e}^*$ et $\overline{f} = \overset{-1}{\overline{g}}\, \overline{e}$, avec $\overline{g} = \text{dét}|g_{ij}|$.

Pour une suite strictement croissante I de p indices pris parmi $(1, 2, \dots, n)$: $\mathrm{I} = \{1 \leqslant i_1 \leqslant i_2 < \dots i_p \leqslant n\}$, on note $e_\mathrm{I} = e_{i_1} \wedge \dots e_{i_p}$, $f^\mathrm{I} = f^{i_1} \wedge \dots f^{i_p}$. Dans $\wedge^p\mathrm{E}$, le produit scalaire $(e_\mathrm{I} \mid e_\mathrm{J}) = \text{dét}_{\substack{i\in\mathrm{I}\\ j\in\mathrm{J}}} |(e_i \mid e_j)| = g_{\mathrm{IJ}}$ est le déterminant de la sous-matrice de $\mathrm{G} = |g_{ij}|$ formée des lignes de I et des colonnes de J.
$(f^i \mid f^j) = (\overset{-1}{\rho}\, e^{*i} \mid \overset{-1}{\rho}\, e^{*j}) = (e^{*i} \mid e^{*j}) = g^{ij}$ et $(f^\mathrm{I} \mid f^\mathrm{J}) = g^{\mathrm{IJ}}$ tandis que $(f^*_\mathrm{I} \mid f^*_\mathrm{J}) = g_{\mathrm{IJ}}$. On a :
$(\overline{\wedge}^p\rho)(e_\mathrm{I}) = \rho e_{i_1} \wedge \dots \rho e_{i_p} = \sum g_{i_1k_1}e^{*k_1} \wedge \dots \sum g_{i_pk_p}e^{*k_p}$
$= \sum g_{\mathrm{IJ}}e^{*\mathrm{J}}$
$(\overline{\wedge}^p\, \overset{-1}{\rho})(e^{*1}) = f^\mathrm{I} = \sum g^{\mathrm{IJ}}e_\mathrm{J}$. Les bases e_I et f^I sont supplémentaires dans $\wedge^p\mathrm{E}$. Les bases $e^{*\mathrm{I}}$ et f^*_J sont supplémentaires dans $\wedge^p\mathrm{E}^*$.

Pour le produit scalaire étendu aux espaces vectoriels $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$, les applications linéaires associées, sommes des $\overline{\wedge}^p\rho$ ou $\overline{\wedge}^p\, \overset{-1}{\rho}$, seront simplement notées $\overline{\wedge}\rho$ et $\overline{\wedge}\, \overset{-1}{\rho}$, soit : $(t \mid u) = \langle t, (\overline{\wedge}\rho)u\rangle$ $= \langle(\overline{\wedge}\rho)t, u\rangle$ dans $\wedge\mathrm{E}$ et $(\varphi \mid \omega) = \langle\varphi, (\overline{\wedge}\, \overset{-1}{\rho})\omega\rangle = \langle(\overline{\wedge}\, \overset{-1}{\rho})\varphi, \omega\rangle$ dans $\wedge\mathrm{E}^*$.
$\overline{\wedge}\rho$ et $\overline{\wedge}\, \overset{-1}{\rho}$ sont des isomorphes *d'algèbres*.

Définition V.5.A :

Soit $(\mathrm{E}; g)$ un espace métrique dont on étend le produit scalaire à l'algèbre extérieure $\wedge\mathrm{E}$. L'opérateur adjoint de la multiplication extérieure à droite est le *produit intérieur à gauche* : $\lrcorner$, l'adjoint de la multiplication extérieure à gauche est le *produit intérieur à droite :* $\llcorner$:

$$(t \wedge u \mid v) = (t \mid u \lrcorner v) : (u \lrcorner) = (. \wedge u)^*$$
$$(v \mid u \wedge t) = (v \llcorner u \mid t) : (\llcorner u) = (u \wedge .)^*$$

Les éléments t, u, v étant supposés homogènes, on a $d^\circ t + d^\circ u = d^\circ v$, soit $d^\circ(u \lrcorner v) = d^\circ(v \llcorner u) = d^\circ v - d^\circ u$ tandis que :

$$(v \llcorner u \mid t) = (v \mid u \wedge t) = (-1)^{d^\circ t \cdot d^\circ u}(t \wedge u \mid v)$$
$$= (-1)^{d^\circ u(d^\circ v - d^\circ u)}(u \lrcorner v \mid t)$$

d'où il résulte que, sur des éléments homogènes, $\lrcorner$ et $\llcorner$ ne diffèrent que par le signe : $v \llcorner u = (-1)^{d^\circ u(d^\circ v - d^\circ u)} u \lrcorner v$.

Il faut vérifier que les produits intérieurs que nous venons de définir sont bien les mêmes que les produits intérieurs de la définition IV.10.A après identification de $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ par $\overline{\wedge}\rho$ et $\overline{\wedge}\,\overset{-1}{\rho}$. Or ces derniers sont les opérateurs transposés des multiplications extérieures. Si $d^\circ v = p$: $\langle t, u \lrcorner \alpha\rangle = \langle t \wedge u, \alpha\rangle = (\rho t \wedge \rho u \mid \alpha)$ $= (\rho t \mid \rho u \lrcorner \alpha) = \langle t, \rho u \lrcorner \alpha\rangle$ soit $u \lrcorner \alpha = \rho u \lrcorner \alpha$.

On peut donc sans craindre d'ambigüité, conserver la même notation.

L'associativité de la multiplication extérieure a pour conséquence les égalités suivantes :

$$(s \wedge t \wedge u \mid v) = (s \wedge t \mid u \lrcorner v) = (s \mid t \lrcorner (u \lrcorner v))$$
$$\text{d'où : } t \lrcorner (u \lrcorner v) = (t \wedge u) \lrcorner v$$

et de la même façon : $(v \llcorner t) \llcorner u = v \llcorner (t \wedge u)$.

Nous allons maintenant examiner comment la dualité entre $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ devient, grâce à l'identification par le produit scalaire, une dualité interne à $\wedge\mathrm{E}$.

Soient $\overline{e}$ une base de $\wedge^n\mathrm{E}$, $\overline{e}^*$ la base duale de $\wedge^n\mathrm{E}^*$: $\langle \overline{e}, \overline{e}^*\rangle = 1$, $\overline{f} = (\overline{\wedge}\,\overset{-1}{\rho}) \cdot \overline{e}^*$ et $\overline{f}^* = (\overline{\wedge}^n \rho) \cdot \overline{e}$.

Les isomorphismes linéaires H et H^* dits isomorphismes de dualité entre $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ sont définis par (Proposition IV.10.A)

$$\langle t \wedge u; \overline{e}^*\rangle = \langle t; \mathrm{H}(u)\rangle = \langle t; u \lrcorner \overline{e}^*\rangle$$
$$\langle \overline{e}; \varphi \wedge \omega\rangle = \langle \mathrm{H}^*(\varphi); \omega\rangle = \langle \overline{e} \llcorner \varphi; \omega\rangle \text{ et } \mathrm{H}^* = \overset{-1}{\mathrm{H}}$$

soient : $\boxed{\mathrm{H}(u) = u \lrcorner \overline{e}^* \quad \text{et} \quad \mathrm{H}^*(\varphi) = \overline{e} \llcorner \varphi}$

Or : $(t \wedge u \mid .\overline{e}) = \langle t \wedge u; \overline{\wedge}^p \rho \cdot \overline{e}\rangle = \overline{g}\langle t \wedge u; \overline{e}^*\rangle = \overline{g}\langle t; \mathrm{H}(u)\rangle$.

D'où : $(t \mid u \lrcorner \overline{e}) = (t \mid \overline{g}(\overline{\wedge}\,\overset{-1}{\rho})\mathrm{H}(u))$.

Pour définir une dualité interne associée à la métrique g sur l'algèbre extérieure $\wedge\mathrm{E}$, *on a le choix* entre plusieurs automorphismes naturels, dépendant toujours du choix d'une base $\overline{e}$ de $\wedge^n\mathrm{E}$:

1) l'application $\alpha = \overline{\wedge}\,\overset{-1}{\rho} \circ \mathrm{H}$ caractérisée par : $t \wedge u = (t \mid \alpha u)\overline{e}$;
2) l'application interne β : $\beta(u) = u \lrcorner \overline{e} = \overline{g}\alpha(u)$;

3) l'application interne étoile $*$ définie par l'égalité :

$$t \wedge *u = (t \mid u)\overline{e} \quad \text{(aussi égal à } u \wedge *t).$$

Exprimons cette dernière en fonction des autres :

$$t \wedge *u = (t \mid \alpha * u)\overline{e} = (t \mid u)\overline{e} \quad \text{soit : } \alpha * u = u.$$

Or, on verra ci-dessous que si $d^\circ u = p$ ou $(n-p)$ le carré de α est égal à : $(\alpha \circ \alpha)(u) = (-1)^{p(n-p)} \overset{-1}{g}\, u$. On a donc :

$$*u = (-1)^{p(n-p)}\overline{g}\alpha u = (-1)^{p(n-p)} u \,\lrcorner\, \overline{e}.$$

Cas particulier où (E, g) *est un espace pseudoeuclidien :*

On a alors (définition V.3.C) une base canonique de $\wedge^n\mathrm{E}$: *l'unité de volume orienté* ω, invariant défini pour toute base orientée positivement par : $\omega = |\overline{g}|^{-1/2}\overline{e}$. Les automorphismes de dualité de $\wedge\mathrm{E}$ précédents, dépendant de $\overline{e}$, deviennent canoniques lorsqu'on les rapporte à ω. En gardant pour ces automorphismes canoniques les notations $\alpha, \beta, *$, ils sont définis respectivement par :

$$\boxed{(t \wedge u = (t \mid \alpha u)\omega); \;\; \beta u = u \,\lrcorner\, \omega; \;\; t \wedge *u = (t \mid u)\omega}$$

La comparaison avec les automorphismes attachés à $\overline{e}$ donne $\alpha = |\overline{g}|^{1/2}\alpha_{\overline{e}}$; $\beta = |\overline{g}|^{-1/2}\beta_{\overline{e}}$; $* = |\overline{g}|^{1/2}*_{\overline{e}}$

Il en résulte que $\alpha = \beta$ et $* = (-1)^{p(n-p)}\alpha$.

Il n'y a donc plus en fait qu'un seul automorphisme naturel de dualité de $\wedge\mathrm{E}$ au signe près.

La géométrie différentielle choisit traditionnellement $*$ qui est donc déterminé sur $\wedge\mathrm{E}$ par $\boxed{t \wedge *u = (t \mid u)\omega}$ soit aussi

$$\boxed{*u = (-1)^{p(n-p)} u \,\lrcorner\, \omega}$$

On a $*1 = \omega$ et $*\omega = (-1)^s$.

On verra au chapitre IX (§IX.3) que l'opérateur d'adjonction $*$ est surtout utilisé dans l'algèbre extérieure des formes différentielles.

Si $\varphi \in \wedge^p\mathrm{E}^*$: $(*\varphi)_{j_1 j_2 \dots j_{n-p}} = (-1)^{p(n-p)} \sum \omega_{j_1 j_2 \dots j_{n-p} i_1 \dots i_p} \varphi^{i_1 \dots i_p}$.

soit : $$\boxed{(*\varphi)_{j_1 \dots j_{n-p}} = \sum \omega_{i_1 \dots i_p j_1 \dots j_{n-p}} \varphi^{i_1 \dots i_p}}.$$

expression qui, en raison de sa simplicité, est souvent prise comme définition de l'adjonction.

Revenons au cas d'un corps quelconque et supposons $\overline{e} = e_1 \wedge e_2 \wedge \dots e_n$.

Par un changement de base $e' = e\mathrm{S}$, d'où $\overline{e}' = (\text{dét S}) \cdot \overline{e}$: $\alpha_{e\mathrm{S}} = \overset{-1}{(\text{dét}}\ \mathrm{S})\alpha_e$.

Si a est un opérateur orthogonal de E, prolongé à l'espace vectoriel $\wedge$E, on a : $\alpha(au) = a \cdot \alpha(u)$dét a. Si α_p désigne la restriction de α à $\wedge^p$E, α_p est un isomorphisme de l'espace vectoriel $\wedge^p$E sur l'espace vectoriel $\wedge^{n-p}$E : $\alpha_p = \overline{\wedge}\ \overset{-1}{\rho} \cdot$H qui transforme la représentation de O(E; g) dans $\wedge^p$E en la représentation associée dans $\wedge^{n-p}$E. Géométriquement, si u est un p-vecteur décomposable non nul auquel correspond un sous-espace U, de dimension p, de E, $\alpha_p(u)$ est un $(n-p)$-vecteur décomposable auquel correspond le sous-espace $\mathrm{U}^\perp$ orthogonal de U.

Théorème V.5.A : Le carré α^2 de α est un automorphisme linéaire gradué de degré zéro de l'espace vectoriel $\wedge$E, égal sur $\wedge^p$E à une homothétie, soit :

$$\boxed{\alpha_{n-p} \circ \alpha_p = (-1)^{p(n-p)}\ \overset{-1}{\overline{g}}}\,.$$

Il en résulte que dans le cas d'un espace pseudoeuclidien d'indice (r, s) on a :

$$\boxed{*.* = (-1)^{p(n-p)+s}}\,.$$

Preuve : Comme p-vecteur décomposable non nul, il suffit de prendre le produit de p vecteurs de la base e, soit $e_{i_1} \wedge e_{i_2} \wedge \dots e_{i_p}$. On a (cf. §IV.10) avec toujours $f = \overset{-1}{\rho}\ e^*$:

$\mathrm{H}(e_\mathrm{I}) = \varepsilon(\mathrm{I}')e_{\mathrm{I}'}$ et $\boxed{\alpha(e_\mathrm{I}) = \varepsilon(\mathrm{I}')f_{\mathrm{I}'}}$ qui est bien décomposable. De plus chaque vecteur de base e_i de la base $(e_{i_1}, e_{i_2}, \dots e_{i_p})$ du sous-espace associé au p-vecteur e_I est orthogonal à chaque vecteur $f^{i'}$ de la base $(f^{i'_1}, \dots, f^{i'_p}$ du sous-espace associé au $(n-p)$-vecteur $f^{\mathrm{I}'}$. En attribuant un indice f aux automorphismes de dualité associés à la base f, remarquons que $\mathrm{H}_f(f^{\mathrm{I}'}) = \varepsilon(\mathrm{I}) f_{*,\mathrm{I}}$. Puisque, ainsi qu'on l'a vu ci-dessus :

$t \wedge u = \langle t, \mathrm{H}(u)\rangle \overline{e} = \langle t, \mathrm{H}_f(u)\rangle = \overline{f} = \langle t, \mathrm{H}_f(u)\rangle \overset{-1}{\overline{g}}\, \overline{e}$, on a $\mathrm{H}(u) = \overset{-1}{\overline{g}}\, \mathrm{H}_f(u)$ et $\mathrm{H}(f^{\mathrm{I}'}) = \varepsilon(\mathrm{I})\, \overset{-1}{\overline{g}}\, f^*_{\mathrm{I}}$. En appliquant maintenant $\overline{\wedge}^p \overset{-1}{\rho}$, puisque $\overset{-1}{\rho} f^*_j = \overset{-1}{\rho} \rho e_j = e_j$, on obtient : $\alpha_{n-p}(f'^{\mathrm{I}})$ $= \varepsilon(\mathrm{I})\, \overset{-1}{\overline{g}}\, e_{\mathrm{I}}$, et

$$\alpha_p \circ \alpha_{n-p} = \varepsilon(\mathrm{I})\varepsilon(\mathrm{I}')\, \overset{-1}{\overline{g}} .$$

On passe de la permutation de $(1, 2, \ldots, n)$ formée de $(\mathrm{I}, \mathrm{I}')$ à la permutation $(\mathrm{I}'\mathrm{I})$ en faisant passer successivement chaque élément de I par dessus ceux de I' ce qui fait en tout $p(n - p)$ transpositions, et $\varepsilon(\mathrm{I})\varepsilon(\mathrm{I}') = (-1)^{p(n-p)}$ cqfd. Puisque $* = (-1)^{p(n-p)}|\overline{g}|^{1/2}\alpha$, on a $** = (-1)^{p(n-p)}|\overline{g}| \cdot \overset{-1}{\overline{g}} = (-1)^{p(n-p)}$ signe de $\overline{g} = (-1)^{p(n-p)+s}$.

Corollaire V.5 : Sur l'algèbre extérieure $\wedge\mathrm{E}$ de l'espace pseudoeuclidien E de type (r, s), l'opérateur d'adjonction $*$ a pour adjoint $(-1)^s \overset{-1}{*} = (*u \mid v) = (u \mid (-1)^s \overset{-1}{*} v)$.

Il en résulte que l'opérateur $*$ est, au facteur $(-1)^s$ près, une *isométrie* : $(*v \mid *v) = (-1)^s(u \mid \overset{-1}{*} *v) = (-1)^s(u \mid v)$.

Preuve : Soient $u \in \wedge^p\mathrm{E}$ et $v \in \wedge^{n-p}\mathrm{E}$:

$$\begin{aligned}(u \mid \overset{-1}{*} v)\omega = u \wedge v = (-1)^{p(n-p)} v \wedge u = (v \mid (-1)^{p(n-p)})\, \overset{-1}{*} u)\omega \\ = (v \mid (-1)^s * u) \text{ cqfd.}\end{aligned}$$

Remarques :

1) Si n est pair, $n = 2r$, α_r est un automorphisme de l'espace vectoriel $\wedge^r\mathrm{E}$ de carré $(-1)^{r^2}(\overset{-1}{\overline{g}}) \cdot \mathrm{Id}$. Or $(-1)^{r^2} = (-1)^r$.

Si le scalaire $(-1)^r(\overline{g})$ est un carré dans le corps K, égal à λ^2 par exemple, $\lambda\alpha^r$ est un automorphisme involutif de $\lambda^r\mathrm{E}$ qui le décompose en la somme directe de ses sous-espaces propres, images et noyaux des projecteurs $\frac{1}{2}$ $(\mathrm{Id} + \lambda\alpha_r)$, $\frac{1}{2}$ $(Id - \lambda\alpha_r)$.

Un automorphisme orthogonal σ de déterminant (-1) applique chacun de ces sous-espaces sur l'autre. En effet : $\sigma(\mathrm{Id} + \varepsilon\lambda\alpha_r) = (\mathrm{Id} - \varepsilon\lambda\alpha_r)\sigma$ (où $\varepsilon = \pm 1$). Mais les automorphismes orthogonaux de $\mathrm{SO}(\mathrm{E}; g)$ (dét $= +1$), appliquent chacun de ces sous-espaces en lui-même : ce sont des $\mathrm{SO}(\mathrm{E}; g)$-modules. Insistons sur le fait que *ces sous-espaces ne contiennent aucun r-vecteur décomposable non nul.*

En effet, un élément u d'un de ces sous-espaces appartient au noyau de l'un des projecteurs, ce qui s'écrit $(\mathrm{Id} + \varepsilon\lambda\alpha_r)u = 0$ soit $u = \mu\alpha_r u$. Si u était décomposable, il résulterait de cette égalité que les sous-espaces de dimension r de E associés à u et à $\alpha_r u$ sont les mêmes. Or ils sont orthogonaux d'après le théorème V.5.A et c'est donc impossible.

2) la condition trouvée ci-dessus : « $(-1)^r(\overline{g})$ est un carré dans le corps K », est évidemment indépendante du choix de la base e. En effet, si $e' = e \cdot \mathrm{S}$ est une autre base, $\mathrm{G}' = {}^t\,\mathrm{SGS}$, $\overline{d}' = (\text{dét}\,\mathrm{S})^2 \cdot \overline{g}$.

Cas particulier : le cas le plus simple est celui du plan euclidien muni d'une base e orthonormale. Dans ce cas, $r = 1$ et $\wedge^1\mathrm{E} = \mathrm{E}$. L'automorphisme α_1 a pour carré moins l'identité et n'est autre que la rotation de $+\left(\frac{\pi}{2}\right)$ qui définit la structure complexe standard associé au plan euclidien orienté.

Théorème V.5.B : Soit $(\mathrm{E}; g)$ un espace vectoriel muni du produit scalaire g. Les représentations du groupe orthogonal $\mathrm{O}(\mathrm{E}; g)$ dans les puissances extérieures $\wedge^p\mathrm{E}$ sont irréductibles.

Preuve : Commençons par trois observations :

a) on vient de voir à la remarque précédente que si $\dim \mathrm{E} = n = 2r$, la représentation du sous-groupe $\mathrm{SO}(\mathrm{E}; g)$ de $\mathrm{O}(\mathrm{E}; g)$ dans $\wedge^r\mathrm{E}$ peut ne pas être irréductible. Les représentations de $\mathrm{SO}(\mathrm{E}; g)$ dans les puissances extérieures seront étudiées au numéro suivant.

b) le seuls hypothèses sur $(\mathrm{E}; g)$ sont que le corps K est de caractéristique zéro et que la forme bilinéaire symétrique g est non-dégénérée. Il faut bien se garder de vouloir adapter à cette situation très générale les raisonnements simplistes valables dans le cas d'un espace euclidien !

c) on peut écarter des raisonnements les cas $p = 0$ et n : $\wedge^0 \mathrm{E} = \mathrm{K}$ et $\wedge^n \mathrm{E}$ étant des droites la représentation ne peut être qu'irréductible.

Soient $e = (e_1, e_2, \ldots, e_n)$ une base *orthogonale* de E, $\wedge^p e = \{e_\mathrm{I}\}$ la base associée de $\wedge^p \mathrm{E}$ où I parcourt l'ensemble $\mathfrak{I}_p$ des suites strictements croissantes de p indices pris dans $(1, 2, \ldots, n)$ soient : $\mathrm{I} = \{1 \leqslant i_1 < i_2 < \ldots < i_p \geqslant n\}$.

L'ordre lexicographique des suites I est, sur $\mathfrak{I}_p$, un ordre total. La preuve est basée sur deux lemmes :

Lemme V.5.A : Soient V un sous-espace de $\wedge^p \mathrm{E}$ stable par les opérations de $\mathrm{O}(\mathrm{E}; g)$ et u un élément non nul de V. Quelle que soit la base *orthogonale* e de E, si $u = \lambda_1 e_{\mathrm{J}_1} + \lambda_2 e_{\mathrm{J}_2} + \ldots + \lambda_m e_{\mathrm{J}_m}$ avec $\mathrm{J}_1 < \mathrm{J}_2 < \ldots \mathrm{J}_m$ et $\lambda_1, \lambda_2, \ldots, \lambda_m$ non nuls, les éléments $e_{\mathrm{J}_1}, e_{\mathrm{J}_2}, \ldots, e_{\mathrm{J}_m}$ de la base $\wedge^p e$ appartiennent nécessairement à V.

Preuve du lemme : si $m = 1$, $e_{\mathrm{J}_1} \in \mathrm{V}$.

Si $m \geqslant 2$, soient j un indice appartenant à J_2 mais pas à J_1 ; et σ_j la symétrie orthogonale de E qui change e_j en $(-e_j)$ et laisse donc invariants tous les e_i, $i \neq j$. On a : $\sigma_j \cdot e_{\mathrm{J}_1} = e_{\mathrm{J}_1}, \sigma_j \cdot e_{\mathrm{J}_2} = -e_{\mathrm{J}_2}$ et d'une façon générale, $\sigma_j \cdot e_\mathrm{I} = \pm e_\mathrm{I}$.

L'élément $\dfrac{1}{2}(u + \sigma_j \cdot u) = \lambda_1 e_{\mathrm{J}_1} + \mu_3 e_{\mathrm{J}_3} + \ldots + \mu_m e_{\mathrm{J}_m}$ avec $\mu_k = \lambda_k$ ou 0 si $3 \leqslant k \leqslant m$, appartient à V et n'a, au plus, que $(m-1)$ composantes non nulles dans la base $\wedge^p e$. Une récurrence évidente aboutit à $e_{\mathrm{J}_1} \in \mathrm{V}$. Le même raisonnement avec des choix différents des σ_j montre que $e_{\mathrm{J}_2}, \ldots, e_{\mathrm{J}_m}$ appartiennent aussi à V.

Il résulte du lemme que pour toute base orthogonale e de E, certains des p-vecteurs e_J de la base $\wedge^p e$ forment une base de V, les autres formant une base de $\mathrm{V}^\perp$ qui est également stable par $\mathrm{O}(\mathrm{E}; g)$. Soient donc $e_{\mathrm{J}_1} \in \mathrm{V}$ avec $\mathrm{J}_1 = (j_1 < j_2 < \ldots < j_p)$, $j \in \mathrm{J}_1$ et $k \notin \mathrm{J}_1$; e_j et e_k forment une base orthogonale d'un plan P_{jk} de E. Soit J_2 la suite strictement croissante obtenue en remplaçant j dans J_1 par k. Nous allons montrer que $e_{\mathrm{J}_2} \in \mathrm{V}$. Pour cela, nous allons démontrer le

Lemme V.5.B : Dans le plan P_{jk} de base orthogonale (e_j, e_k), on peut construire une autre base orthogonale (e'_j, e'_k) : $e'_j = e_j + \lambda e_k$, $e'_k = e_j + \mu e_k$ où les coefficients λ et μ sont *non nuls*.

Preuve : $(e'_j \mid e'_j) = g_{jj} + \lambda^2 g_{kk}$ est non nul si $\lambda^2 \neq -\dfrac{g_{jj}}{g_{kk}}$; λ non nul étant ainsi choisi, déterminons μ par la condition : $(e'_j \mid e'_k) = 0 = g_{jj} + \lambda\mu g_{kk}$ soit $\mu = -\dfrac{g_{ii}}{\lambda_{g_{kk}}}$. Cette condition entraîne automatiquement $(e'_k \mid e'_k) \neq 0$.

On va maintenant remplacer la base orthogonale initiale e par une autre base orthogonale e' formée des mêmes vecteurs à l'exception de e_j et e_k qui sont remplacés par e'_j et e'_k. Or :

$$e_k = (\lambda - \mu)^{-1}(e'_j - e'_k) \quad \text{et} \quad e_j = (\lambda - \mu)^{-1}(\lambda e'_k - \mu e'_j).$$

Notons $e_{\mathrm{J}'_1}$ le produit extérieur e_{J_1} où l'on remplace e_j par e'_j et qui est aussi, au signe près le produit extérieur e_{J_2} où l'on remplace e_k par e'_j ! De même, notons $e_{\mathrm{J}''_1}$ le produit extérieur e_{J_1} où l'on remplace cette fois e_j par e'_k, et qui est aussi, au signe près, e_{J_2} où l'on remplace e_k par e'_k. Les égalités précédentes donnent $e_{\mathrm{J}_1} = \alpha e_{\mathrm{J}'_1} + \beta e_{\mathrm{J}'_2}$ et $e_{\mathrm{J}_2} = \gamma e_{J'_1} + \delta e_{\mathrm{J}'_2}$ avec $\alpha, \beta, \gamma, \delta$ non nuls. Le lemme précédent montre que, puisque e_{J_1} appartient à V, par hypothèse, $e_{\mathrm{J}'_1}$ et $e_{\mathrm{J}'_2}$ appartiennet à V donc aussi e_{J_2}. De proche en proche, tous les éléments e_{I} de $\wedge^p\mathrm{E}$ appartiennent à V, et le théorème est ainsi démontré.

Remarque : nous avons en particulier démontré que la représentation $\mathrm{O}(\mathrm{E}; g)$ dans E est irréductible !

Définition V.5 :

Commutant d'un module. Soit E un A-module, c'est-à-dire un espace vectoriel E dans lequel on a distingué une partie A de l'algèbre $\mathfrak{L}(\mathrm{E})$ (E est, par exemple, l'espace d'une représentation linéaire $\mathfrak{R}$ d'un groupe G et $\mathrm{A} = \mathfrak{R}(\mathrm{G})$). Le *commutant* de E est la sous-algèbre unitaire C de $\mathfrak{L}(\mathrm{E})$ que forment les opérateurs qui commutent avec tous ceux de A :

$$\gamma \in \mathrm{C} \Longleftrightarrow a\gamma = \gamma a, \quad \forall a \in \mathrm{A}.$$

C'est donc l'algèbre des endomorphismes du A-module E : $\forall x \in \mathrm{E}$, $\gamma(ax) = a\gamma(x)$. On retrouvera cette notion au §VI.4.

Exemples

1) si E est réductible en la somme directe de deux sous-modules F_1 et F_2, sous-espaces stables par les opérations de A : $E = F_1 \oplus E_2$, les projecteurs p_1 et p_2 sur ces sous-espaces appartiennent au commutant.

2) si G est un groupe, le commutant d'un G-module E est la sous-algèbre des invariants de la représentation associée de G dans $\mathfrak{L}(E)$ puisque $\gamma\mathfrak{R}(a) = \mathfrak{R}(a)\gamma$, s'écrit alors $\gamma = \mathfrak{R}(a) \cdot \gamma\mathfrak{R}(\overset{-1}{a})$.

Théorème V.5.C : L'image, le noyau, les sous-espaces propres d'un élément γ du commutant C d'un module E sont des sous-modules. Il en résulte que si un module est simple, son commutant est un corps (éventuellement non commutatif). Si de plus le corps K de E est algébriquement clos, le corps C ne peut alors être que $K \cdot \mathrm{Id}$.

Preuve :

a) si $\gamma x = 0$, $\gamma a x = a\gamma x = 0$, $\forall a \in A$, d'où $A \cdot \operatorname{Ker}\gamma \subset \operatorname{Ker}\gamma$.
b) si $x = \gamma y$, $ax = a\gamma y = \gamma a y$, $\forall a \in A$, d'où $A \cdot \operatorname{Im}\gamma \subset \operatorname{Im}\gamma$
c) quel que soit le scalaire λ, si $\gamma \in C$, $\gamma - \lambda \mid d \in C$ et d'après a) $\operatorname{Ker}(\gamma - \lambda \mid \mathrm{Id})$, sous-espace propre de γ pour sa valeur propre λ, est un sous-module.

Si E est simple, cela signifie que les seuls sous-modules sont 0 et E. Si γ est un élément non nul de C, $\operatorname{Im}\gamma \neq 0$ donc $\operatorname{Im}\gamma = E$, $\operatorname{Ker}\gamma \neq E$, donc $\operatorname{Ker}\gamma = 0$ d'où il résulte que γ est inversible : C est un corps.

De plus, les éléments de C qui ne sont pas des multiples scalaires de l'identité ne peuvent avoir aucune valeur propre dans le corps. Si le corps K de E est algébriquement clos tout opérateur linéaire de E possédant au moins une valeur propre, on a alors $C \equiv K \cdot \mathrm{Id}$.

Théorème V.5.D : Le commutant de la représentation du groupe orthogonal $O(E; g)$ dans $\wedge^p E$ est formé des multiples scalaires de l'opérateur identique : $C = K \cdot \mathrm{Id}$.

Preuve : Soit $e = (e_1, e_2, \ldots, e_n)$ une base orthogonale de E. Soit σ_j la symétrie orthogonale de E qui change e_j en $-e_j$ et laisse donc invariants les e_i, $i \neq j$. Les diverses σ_j commutent entre elles et leurs produits forment un groupe abélien fini T d'ordre 2^n contenu dans $\mathrm{O(E;}g)$. Chaque $e_\mathrm{I} = e_{i_1} \wedge \ldots e_{i_p}$ de la base $\wedge^p e$ de $\wedge^p \mathrm{E}$ est un vecteur propre pour les 2^n opérateurs de T, qui s'écrivent $\tau = \sigma_{j_1}\sigma_{j_2}\ldots\sigma_{j_k}$ avec $1 \leqslant j_1 < \ldots < j_k \leqslant n$, associé à une valeur propre collective ε_I, fonction sur T à valeurs $\pm 1 : \tau \to \varepsilon_\mathrm{I}(\tau)$, définie par $\tau \dot{e}_\mathrm{I} = \varepsilon_\mathrm{I}(\tau)e_\mathrm{I}$.

Il est clair que si $\mathrm{I}_1 \neq \mathrm{I}_2$, $\varepsilon_{\mathrm{I}_1} \neq \varepsilon_{\mathrm{I}_2}$ puisque si $j \in \mathrm{I}_2$ et $j \neq \mathrm{I}_1$, $\varepsilon_{\mathrm{I}_1}(\tau)$ et $\varepsilon_{\mathrm{I}_2}(\sigma_j) = -1$. Les vecteurs e_I de la base $\wedge^p e$ sont donc associés à des valeurs propres collectives distinctes des opérateurs de T. Soit maintenant γ un élément du commutant. Si u est un vecteur propre de tous les opérateurs de T associé à une valeur propre collective ε, il en est de même de γu puisque $\tau(\gamma u) = \gamma(\tau u) = \varepsilon(\tau)\gamma u$.

Il en résulte que tous les e_I sont des vecteurs propres de γ : $\gamma e_\mathrm{I} = \lambda_\mathrm{I} e_\mathrm{I}$ où $\lambda_\mathrm{I} \in \mathrm{K}$. Mais si V est le sous-espace propre de γ dans $\wedge^p\mathrm{E}$ pour la valeur propre λ, V est stable par $\mathrm{O(E;}g)$: $\gamma(av) = a(\gamma v)$, $\forall a \in \mathrm{O(E;}g)$. Puisque $\wedge^p\mathrm{E}$ est un $\mathrm{O(E;}g)$-module simple d'après le théorème V.5.B ci-dessus, V est ou nul ou égal à $\wedge^p\mathrm{E}$: tous les λ_I sont égaux et $\gamma = \lambda \cdot \mathrm{Id}$ avec $\lambda \in \mathrm{K}$, cqfd.

V.6 – Représentations du groupe $\mathrm{SO(E;}g)$ dans les puissances extérieures de E

On a vu au théorème V.5.B que l'automorphisme α_e « de dualité », de l'espace vectoriel $\wedge\mathrm{E}$, définit pour chaque entier p avec $0 \leqslant p \leqslant \dim \mathrm{E} = n$, un isomorphisme $\alpha_{e,p}$ du $\mathrm{SO(E;}g)$-module $\wedge^p\mathrm{E}$ sur le $\mathrm{SO(E;}g)$-module $\wedge^{n-p}\mathrm{E}$. Cet isomorphisme devient, lorsque n est pair et égal à $2r$, un automorphisme du $\mathrm{SO(E;}g)$-module $\wedge^r\mathrm{E}$, ce qui signifie, comme on vient de le voir, que $\alpha_{e,r}$ appartient au commutant du $\mathrm{SO(E;}g)$-module $\wedge^r\mathrm{E}$.

En écartant d'abord ce dernier cas, on a le :

Théorème V.6.A : Les représentations du groupe spécial orthogonal $\mathrm{SO(E;}g)$ dans les puissances extérieures $\wedge^p\mathrm{E}$:
a) sont toutes irréductibles si $n = \dim \mathrm{E}$ est impair.
b) sont irréductibles pour tout entier $p \neq \dfrac{n}{2}$ lorsque n est pair.

Preuve : elle est analogue à celle du théorème V.5.B. Soit e une base orthogonale de E. Si $u = \lambda_1 e_{\mathrm{J}_1} + \lambda_2 e_{\mathrm{J}_2} + \ldots + \lambda_m e_{\mathrm{J}_m}$ est un élément d'un sous-espace V dans $\wedge^p E$ appliqué en lui-même par $SO(E; g)$ avec $J_1 < J_2 < \ldots < J_n$ et $\lambda_1, \lambda_2, \ldots, \lambda_m$ non nuls, on peut, si $m > 2$, choisir un indice j appartenant à J_2 et n'appartenant pas à J_1, et un indice k qui, ou bien n'appartient ni à J_1 ni à J_2 si $p < \frac{n}{2}$, ou appartient aux deux si $p > \frac{n}{2}$.

Dès lors, le produit $\tau = \sigma_j \sigma_k$ des symétries orthogonales correspondantes est un élément de $SO(E; g)$. Si $\tau e_{J_1} = \varepsilon e_{J_1}$ alors $\tau e_{J_2} = -\varepsilon e_{J_2}$.

L'élément $\frac{1}{2}(u + \varepsilon \tau u)$ ne possède plus que $(m-1)$ composantes non nulles dans la base $\wedge^p e$.

Par récurrence, $e_{J_1}, e_{J_2}, \ldots, e_{J_m}$ appartiennent à V ; le lemme V.5.A est donc également valable pour $SO(E; g)$ et pour $p \neq \frac{n}{2}$. Le reste de la démonstration se poursuit exactement comme au théorème V.5.B à l'aide du lemme V.5.B.

Proposition V.6 :

Soient G un groupe, G_+ un sous-groupe de G tel que G/G_+ soit le groupe à deux éléments (exemple : $G = O(E; g)$, $G_+ = SO(E; g)$. Soit V un G-module simple. Par restriction à G_+, V est ou bien un G_+-module simple, ou bien une somme directe de deux sous-G_+-modules simples qu'échange tout élément σ de $G - G_+$.

Preuve : Soit L un sous G_+-module de dimension minimale non nul de V : L est un G_+-module simple. Si $\sigma \in G - G_+$, σL est également un sous G_+-module de V. En effet, $\sigma G_+ \overset{-1}{\sigma} = G_+$: quel que soit $a \in G_+$; il existe $a' \in G_+$ tel que $a\sigma = \sigma a'$.

Il en résulte que $a\sigma L = \sigma a' L$. Le sous-espace de V : $L + \sigma L$ est appliquée en lui-même par G. On l'a vu pour les éléments de G_+. Puisque $G = G_+ \cup \sigma G_+$, un élément $\sigma' \in G - G_+$ s'écrit $\sigma' = \sigma a$ et $\sigma a L = \sigma L$, $\sigma a \sigma L = \sigma^2 a' L = L$ car $\sigma^2 \in G^2 \in G_+$, ce qui prouve aussi que $\sigma(\sigma L) = L$.

Dès lors deux cas sont possibles. Puisque L est G_+-simple, ou bien $\sigma L = L$, qui est alors un G-module simple identique à V, ou bien $L \cap \sigma L = 0$, la somme $L \oplus \sigma L$ est directe et égale à V, c.q.f.d.

Lemme V.6 : Soient V un M-module et $V = L_1 \oplus L_2 \oplus \ldots L_k$ une décomposition de V en somme directe de sous-modules *simples*. Tout sous-module simple L de V est isomorphe à l'un des L_j. Si les L_j

sont des M-modules deux à deux *non-isomorphes*, la décomposition de V est unique à l'ordre près des termes.

Preuve : Tout élément x de L s'écrit de façon unique comme la somme de ses composantes suivant les L_j : $x = x_1 + x_2 + \ldots + x_k$ avec $x_j \in \mathrm{L}_j$. L'application $x \in \mathrm{L} \to x_j \in \mathrm{L}_j$ est un homomorphisme de M-modules puisque $m \cdot x = mx_1 + mx_2 + \ldots + mx_k$.

Puisque L et L_j sont simples, c'est donc ou bien 0, ou bien un isomorphisme. Comme les x_j ne peuvent être tous nuls, L est isomorphe à au moins l'un des L_j. Si les L_j sont deux à deux non-isomorphes, il en résulte que L ne peut être que l'un deux.

Rappelons que si un espace vectoriel V est décomposé en somme directe de sous-espaces $\mathrm{V} = \mathrm{L}_1 \oplus \mathrm{L}_2 \oplus \ldots \mathrm{L}_k$ l'étude d'un endomorphisme f vis-à-vis de cette composition conduit à l'écrire comme une matrice $|f_{\mathrm{L}_i\mathrm{L}_j}|$ (décomposition « en blocs ») où $f_{\mathrm{L}_i\mathrm{L}_j}$ désigne la restriction de f qui applique L_j dans L_i, soit avec les projecteurs associés : $f_{\mathrm{L}_i\mathrm{L}_j} = p_i f p_j$.

Si V est un M-module somme directe de sous-modules : $\mathrm{V} = \mathrm{L}_1 \oplus \ldots \mathrm{L}_k$, deux à deux non-isomorphes, un élément γ du commutant de V étant un endomorphisme de M-modules applique nécessairement chaque L_j en lui-même et le commutant de V, C(V) est donc la somme directe des commutants des L_j : $\mathrm{C(V)} = \mathrm{C}(\mathrm{L}_1) \oplus \mathrm{C}(\mathrm{L}_2) \oplus \ldots \mathrm{C}(\mathrm{L}_k)$.

Remarque : Soient V un $\mathrm{SO(E};g)$-module et R l'homomorphisme : $\mathrm{SO(E};g) \to \mathrm{Gl(V)}$ définissant la représentation.

Un élément $\sigma \in \mathrm{O(E};g) - \mathrm{SO(E};g)$ (autrement dit, $\mathrm{dét}_{\mathrm{E}}\sigma = -1$) détermine un automorphisme du groupe $\mathrm{SO(E};g)$ par : $a \to \widetilde{\sigma}(a) = \sigma a \overset{-1}{\sigma}$ et une nouvelle représentation R_σ de $\mathrm{SO(E};g)$ dans V par : $\mathrm{R}(a) = (\mathrm{R} \circ \widetilde{\sigma})(a) = \mathrm{R}(\sigma a \overset{-1}{\sigma})$.

Les automorphismes de $\mathrm{SO(E};g)$ définis par deux éléments σ et σ' de $\mathrm{O(E};g) - \mathrm{SO(E};g)$ ne diffèrent que par un automorphisme intérieur de $\mathrm{SO(E};g)$. En effet $\sigma' = u\sigma$ avec $u \in \mathrm{SO(E};g)$ et $\widetilde{\sigma}(a) = \sigma' a \overset{\prime -1}{\sigma} = u\sigma a \overset{-1}{\sigma} \overset{-1}{u}$.

Les représentations R_σ et $\mathrm{R}'_{\sigma'}$ obtenues à partir de R et de σ et σ' sont donc équivalentes.

$$\mathrm{R}_{\sigma'}(a) = \mathrm{R}(u)\mathrm{R}_\sigma(a)\mathrm{R}(\overset{-1}{u}).$$

Définition V.6 :

La classe des automorphismes $\widetilde{\sigma}(a) = \sigma a \overset{-1}{\sigma}$ de $\mathrm{SO}(\mathrm{E}; g)$ par les éléments $\sigma \in \mathrm{O}(\mathrm{E}; g) - \mathrm{SO}(\mathrm{E}; g)$ modulo les automorphismes intérieurs est appelé l'automorphisme miroir ("mirror automorphism"). La classe d'équivalence des représentations R_σ associées à une représentation R de $\mathrm{SO}(\mathrm{E}; g)$ est appelée la représentation miroir de R.

L'automorphisme miroir est l'identité si et seulement s'il se réduit à un automorphisme intérieur, c'est-à-dire s'il existe $u \in \mathrm{SO}(\mathrm{E}; g)$ tel que $\sigma a \overset{-1}{\sigma} = u a \overset{-1}{u}$, $\forall a \in \mathrm{SO}(\mathrm{E}; g)$. Cette condition est équivalente à l'existence d'un élément $\sigma_0 = \overset{-1}{u} \sigma$ de $\mathrm{O}(\mathrm{E}; g) - \mathrm{SO}(\mathrm{E}; g)$ qui commute avec *tous* les éléments de $\mathrm{SO}(\mathrm{E}; g)$. Un tel élément existe si $\dim \mathrm{E}$ est impaire : $\sigma_0 = -\mathrm{Id}_\mathrm{E}$, mais n'existe pas si la dimension de E est paire.

La représentation miroir de R est équivalente à R lorsque la représentation s'étend à (ou est la restriction d') une représentation de $\mathrm{O}(\mathrm{E}; g)$.

Exemple : Toutes les représentations irréductibles de $\mathrm{SO}(\mathrm{E}; g)$ dans les $\wedge^p \mathrm{E}$, $p \neq \dfrac{n}{2}$ sont équivalentes à leur représentation miroir.

Théorème V.6.B : Soit $(\mathrm{E}; g)$ un espace vectoriel de dimension paire $n = 2r$ muni d'un produit scalaire g. Si la puissance extérieure r-ème de E se décompose sous l'action du groupe spécial orthogonal en la somme directe de deux sous-modules simples, les représentations de $\mathrm{SO}(\mathrm{E}; g)$ dans ces sous-modules sont miroirs l'une de l'autre et inéquivalentes.

Preuve : Notons R la représentation de $\mathrm{O}(\mathrm{E}; g)$ dans $\wedge^r \mathrm{E}$, et soit $\wedge^r \mathrm{E} = \mathrm{L} \oplus \sigma \mathrm{L}$ où σ est un élément quelconque de $\mathrm{O}(\mathrm{E}; g) - \mathrm{SO}(\mathrm{E}; g)$ suivant la proposition V.6.A et $\sigma\mathrm{L}$ désigne $\mathrm{R}(\sigma)\mathrm{L}$. L'application σ de L sur $\sigma\mathrm{L}$ est un isomorphisme d'espace vectoriel et un isomorphisme de la représentation R de $\mathrm{SO}(\mathrm{E}; g)$, sur L, sur la représentation miroir R_σ de $\mathrm{SO}(\mathrm{E}; g)$, sur $\sigma\mathrm{L}$. En effet, si $x \in \mathrm{L}$, $\sigma a x = (\sigma a \overset{-1}{\sigma}) \cdot \sigma x$.

Si ces représentations étaient équivalentes, il existerait alors un élément γ du commutant de $\mathrm{O}(\mathrm{E}; g)$ dans $\wedge^r \mathrm{E}$, ce qui est impossible d'après le théorème V.5.C. c.q.f.d.

CHAPITRE VI

Algèbres de Clifford. Spineurs

Le chapitre précédent a commencé l'examen des propriétés particulières aux tenseurs d'un espace vectoriel E muni d'un produit scalaire symétrique g, le corps K de E étant un corps *de caractéristique zéro* non précisé. Dans cette même hypothèse générale sur K, nous allons voir qu'il s'introduit un quotient de l'algèbre tensorielle sur E : l'algèbre de Clifford de $(\mathrm{E}; g)$ qui traduit dans sa structure algébrique la géométrie de E. Elle conduit directement aux définitions des *spineurs* et du *groupe spinoriel* associés à $(\mathrm{E}; g)$. Ces notions, lorsque $(\mathrm{E}; g)$ est un espace euclidien ou pseudoeuclidien sont d'importance essentielle en physique. Les définitions générales qui sont données ici justifient les constructions et représentations matricielles, souvent arbitrairement imposées, des physiciens qui ne sont valables que dans des cas particuliers. Historiquement, l'algèbre de Clifford est ainsi intervenue comme sous-jacente à diverses constructions ou théories, comme celles des quaternions, du spin non relativiste de Pauli, ou du spin relativiste de Dirac.

Ce chapitre se contente de donner *succintement*, les bases algébriques et les théorèmes fondamentaux de la théorie, appuyés sur quelques exemples.

N.B. Tous les espaces vectoriels et algèbres considérés dans ce chapitre sont de *dimension finie*, à l'exception, bien entendu, de l'algèbre tensorielle intervenant dans la construction du théorème VI.1.A.

VI.1 – Algèbre de Clifford d'un espace quadratique (cf. [DR]). Exemples.

La donnée d'un produit scalaire symétrique g sur l'espace vectoriel E est équivalente à la donnée de la forme quadratique associée q. En effet, si $q(x) = g(x, x)$, on déduit g de q par :

$$g(x, y) = \frac{1}{2}\,\{q(x + y) - q(x) - q(y)\}.$$

La structure définie par la donnée de g sur E, qui fait de celui-ci un « espace géométrique », est donc un cas particulier de la structure d'*espace quadratique* : $(\mathrm{E}; q)$. En effet, cette dernière désigne un espace vectoriel muni d'une forme quadratique q, mais qui peut-être dégénérée, ainsi, par conséquent, que la forme bilinéaire symétrique associée que nous continuerons à désigner par g. Les espaces quadratiques sur un même corps K peuvent être considérés comme une catégorie (cf. §III.13) une fois définis les « morphismes ». Un morphisme de $(\mathrm{E}; q)$ dans $(\mathrm{E}'; q')$ est tout simplement une application linéaire de f des E dans E$'$ conservant la structure : $q'(fx) = q(x)$, $\forall x \in \mathrm{E}$.

La considération d'espaces quadratiques $(\mathrm{E}; q)$ où q est dégénérée n'est nullement étrangère à l'étude des espaces $(\mathrm{E}; g)$ munis d'un produit scalaire. En effet, un sous-espace F d'un tel espace, muni de la restriction g_{F} de g, peut être singulier comme par exemple, les droites isotropes de l'espace de Minkowski.

L'injection de $(\mathrm{F}; q_{\mathrm{F}})$ dans $(\mathrm{E}; q)$ est évidemment un morphisme.

Si L est un surcorps (commutatif) de K (par exemple $\mathrm{K} = \mathbb{R}$, $\mathrm{L} = \mathbb{C}$), l'application de $\mathrm{E} \times \mathrm{L} \times \mathrm{E} \times \mathrm{L}$ dans L : $(x, \lambda, y, \mu) \to \lambda\mu g(x, y)$ est quadrilinéaire (sur K) et détermine une forme L-bilinéaire $\widetilde{g}$, extension de g, sur $\mathrm{E} \otimes_{\mathrm{K}} \mathrm{L}$: $\widetilde{g}(x \otimes \lambda, y \otimes \mu) = \lambda\mu g(x, y)$. La forme quadratique $\widetilde{q}$ extension de q à $\mathrm{E} \otimes_{\mathrm{K}} \mathrm{L}$ est $\widetilde{q}(x \otimes \lambda) = \lambda^2 q(x)$.

Un problème, qui peut, à première vue, sembler n'être qu'un amusement, mais dont la solution a d'importantes conséquences, est celui-ci : peut-on exprimer une forme quadratique q sur E comme le carré d'une fonction linéaire φ : $q(u) = \varphi(u)^2$?

Exemple : L'exemple concret, résolu par P.A.M. Dirac, et à la base de sa théorie de l'électron, est celui où q est la métrique de

Lorentz sur l'espace vectoriel $\mathbb{R}^4$. Peut-on écrire $q(u) = t^2 - x^2 - y^2 - z^2 = (\alpha t + \beta x + \gamma y + \delta z)^2 = \varphi(u)^2$.

La fonction φ doit être linéaire et les coefficients $\alpha, \beta, \gamma, \delta$; qui représentent les vecteurs de la base standard, doivent pouvoir être ajoutés, multipliés par $\mathbb{R}$, et multipliés entre eux. φ doit donc prendre ses valeurs dans une algèbre sur $\mathbb{R}$.

L'identification des deux membres conduit aux équations :

$$\alpha^2 = -\beta^2 = -\gamma^2 = -\delta^2 = 1$$
$$\alpha\beta + \beta\alpha = \alpha\gamma + \gamma\alpha = \ldots = 0$$

et les coefficients $\alpha, \beta, \gamma, \delta$ ne peuvent les satisfaire que s'ils appartiennent à une algèbre non-commutative. Comme nous le verrons, on obtient une solution en prenant pour $\alpha, \beta, \gamma, \delta$ des matrices complexes d'ordre 4.

Définition VI.1.A :

Soient $(\mathrm{E}; q)$ un espace quadratique sur le corps K, A une algèbre associative sur K possédant une unité notée 1_{A}. Une *application de Clifford* de $(\mathrm{E}; q)$ dans A est une application linéaire φ de E dans A telle que $\varphi(x)^2 = q(x) \cdot 1_{\mathrm{A}}$, $\forall x \in \mathrm{E}$. *Une algèbre de Clifford* $\mathrm{C} = \mathrm{C}(\mathrm{E}; q)$ de l'espace quadratique $(\mathrm{E}; q)$ est une solution du problème universel (cf. III.13) posé par les applications de Clifford de la catégorie des espaces quadratiques sur K dans la catégorie des algèbres associatives avec unité sur K.

Rappelons la signification de cette dernière définition :
1) il existe une application de Clifford φ_C de E dans C ;
2) l'unité 1_{C} de C et l'image $\varphi_{\mathrm{C}}(\mathrm{E})$ engendrent l'algèbre C ;
3) quelle que soit l'application de Clifford φ de E dans A, il existe un homomophisme d'algèbres avec unité Φ de C dans A tel que :

$$\varphi = \Phi \circ \varphi_c : \quad \begin{tikzcd} \mathrm{E} \arrow[rr, "\varphi"] \arrow[dr, "\varphi_c"'] & & \mathrm{A} \\ & \mathrm{C} \arrow[ur, "\Phi"'] & \end{tikzcd}$$

Cet homomorphisme Φ est unique d'après 2).

Si elle existe, l'algèbre de Clifford $\mathrm{C}(\mathrm{E}; q)$ est unique à un isomorphisme près en vertu des considérations générales du chapitre III, §13.

Réciproquement, tout homomorphisme Φ d'algèbres unitaires de C dans A détermine une application de Clifford de E dans A : $x \in \mathrm{E} \to \Phi(\varphi_c(x)) \in \mathrm{A}$ puisque : $\Phi(\varphi_c(x))^2 = \Phi(\varphi_c(x)^2) = \Phi(q(x) \cdot 1_c) = q(x) \cdot 1_{\mathrm{A}}$.

Si φ est une application de Clifford de E dans A, et α un automorphisme de l'algèbre unitaire A, $\alpha \circ \varphi$ est encore une application de Clifford de E dans A auquel correspond l'homomorphisme $\alpha \circ \Phi$ de C dans A.

Remarquons que si A est une algèbre de matrice $m \times m$ et si $\mathrm{X} = \varphi(x)$ est la matrice image de $x \in \mathrm{E}$, la propriété de Clifford $\mathrm{X}^2 = q(x)$. I entraîne $(\det \mathrm{X})^2 = q(x)^m$.

La forme bilinéaire symétrique g associée à q peut, elle aussi, être exprimée à l'aide de la multiplication dans une algèbre A dans laquelle E est appliqué par une application de Clifford.

En effet, en remplaçant x par $x + y$ dans $\varphi(x)^2 = q(x) \cdot 1_{\mathrm{A}}$, on obtient : $\varphi(x)\varphi(y) + \varphi(y)\varphi(x) = 2g(x, y) \cdot 1_{\mathrm{A}}$.

L'orthogonalité de x et y : $g(x, y) = 0$ est ainsi équivalente à l'anticommunauté de leurs images $\varphi(x)$ et $\varphi(y)$. Cela interdit à A d'être commutative dès que $\dim \mathrm{E} \geqslant 2$.

D'autre part, $\varphi(x) = 0$ entraîne $g(x, y) = 0$ quel que soit y.

Si g est non-dégénérée, φ est donc nécessairement injective.

Exemples d'applications de Clifford :

1) Le plongement φ de la droite réelle $\mathbb{R}$ dans le corps complexe $\mathbb{C}$, considéré comme algèbre réelle, défini par $\varphi(x) = ix$ est une application de Clifford de $\mathbb{R}$, muni de la forme quadratique $q(x) = -x^2$.

2) Plus généralement, la forme quadratique q sur la droite K étant définie par $q(\alpha \cdot 1) = \alpha\lambda^2$, on obtient une application de Clifford φ de $(\mathrm{K}; q)$ dans K(2) par : $\varphi(\lambda \cdot 1) = \begin{vmatrix} 0 & \lambda\alpha \\ \lambda & 0 \end{vmatrix}$.

3) l'application φ de l'espace euclidien $\mathbb{E}_3$ dans l'algèbre $\mathbb{C}(2)$ des matrices complexes d'ordre deux, considérée comme algèbre réelle :

$$\varphi(u) = \varphi(xe_1 + ye_2 + ze_3) = \begin{vmatrix} z & x - iy \\ x + iy & -z \end{vmatrix}$$
$$\text{et} \quad \varphi(u)^2 = (x^2 + y^2 + z^2) \cdot \mathrm{I}$$

4) Les exemples ci-dessus ne font que montrer l'existence d'applications de Clifford dans quelques cas. Nous allons construire une application de Clifford par une méthode standard valable pour tout espace quadratique $(\mathrm{E}; q)$. Rappelons (théorème IV.13.B) qu'une forme linéaire $\alpha \in \mathrm{E}^*$ détermine une antidérivation d_α de degré (-1) de l'algèbre $\Lambda\mathrm{E}$, de carré nul.

A chaque vecteur x de E correspond la forme linéaire ρx qui lui est associée par g : $g(x, y) = \langle \rho x, y\rangle$ et par conséquent une dérivation, que nous noterons d_x, de $\Lambda\mathrm{E}$, qui s'écrit sur un p-vecteur décomposable :

$$\begin{aligned} d_x(y_1 \wedge y_2 \wedge \dots y_p) &= \textstyle\sum(-1)^{i-1}\langle \rho x, y_i\rangle y_1 \wedge \dots \widehat{y_i} \wedge \dots y_p, \\ &= \textstyle\sum(-1)^{i-1} g(x, y_i) y_1 \wedge \dots \widehat{y_i} \wedge \dots y_p. \end{aligned}$$

En particulier, si x est orthogonal à chacun des vecteurs y_i, $d_x(y_1 \wedge \dots y_p) = 0$. Soit L_x l'opérateur linéaire de $\wedge\mathrm{E}$ de multiplication à gauche par x : $\mathrm{L}_x(t) = x \wedge t$; L_x est aussi de carré nul. Nous allons montrer que l'application φ :

$$x \in \mathrm{E} \xrightarrow{\varphi} \varphi(x) = d_x + \mathrm{L}_x \in \mathfrak{L}(\wedge\mathrm{E}),$$

de E dans l'algèbre des opérateurs linéaires de $\wedge\mathrm{E}$ est une application de Clifford. En effet, si $u \in \wedge\mathrm{E}$, $\varphi(x)u = d_x u + x \wedge u$ et $\varphi(x)^2 u = d_x^2 u + g(x, x)u - x \wedge d_x u + x \wedge d_x u + x \wedge x \wedge u = q(x)u$ soit $\varphi(x)^2 = q(x) \cdot \mathrm{Id}$

Cette application de Clifford est injective puisque, quel que soit $x \in \mathrm{E}$, l'image de l'unité 1 de l'algèbre $\lambda\mathrm{E}$ par $\varphi(x)$ est $\varphi(x) \cdot 1 = x$.

Proposition VI.1 :

Soit φ une application de Clifford de $(\mathrm{E}; q)$ dans l'algèbre associative unitaire A. La sous-algèbre engendrée par 1_A et $\varphi(\mathrm{E})$ est de dimension finie, inférieure ou égal à 2^n ou $n = \dim \mathrm{E}$.

Preuve : Soient $e = (e_1, e_2, \dots, e_n)$ une base orthogonale de E et u un élément de A de la forme : $u = \varphi(e_{i_1})\varphi(e_{i_2}) \dots \varphi(e_{i_m})$

a) Si $i_k > i_{k+1}$, permutons $\varphi(e_{i_k})$ et $\varphi(e_{i_{k+1}})$ pour rétablir l'ordre naturel des indices et multiplions par (-1) : u ne change pas puisque $\varphi(e_{i_k})$ et $\varphi(e_{i_{k+1}})$ anticommutent.

b) Si $i_k = i_{k+1}$, remplaçons $\varphi(e_{i_k})\varphi(e_{i_{k+1}}) = \varphi(e_{i_k})^2$ par $q(e_{i_k})\cdot 1_A$, u ne change pas. A chaque opération a) le nombre d'inversions de la suite des indices diminue d'une unité. A chaque opération b) le nombre des facteurs diminue de deux. Au bout d'un nombre fini d'opérations u s'écrit : $u = \lambda\varphi(e_{j_1})\ldots\varphi(e_{j_p})$ soit symboliquement $\lambda\varphi(e_J)$ ou $\lambda \in K$ et $J = \{1 \leqslant j_1 < j_2 \ldots j_p \leqslant n\}$ est une suite strictement croissante, qui peut être vide, avec $p \equiv m \mod 2$.

Ainsi, tout élément de la sous-algèbre de A engendrée par l'unité 1_A et l'image $\varphi(E)$ est une combinaison linéaire des 2^n éléments $\{1_A, \varphi(e_J)\}$, cqfd.

Nous allons maintenant aborder le problème de l'existence de l'algèbre de Clifford $C(E; q)$.

Théorème VI.1.A : Tout espace quadratique $(E; q)$ possède une algèbre de Clifford $C(E; q)$ qui peut être définie comme le quotient de l'algèbre tensorielle sur E : $\otimes E$ par l'idéal bilatère Q engendré par les tenseurs symétriques de la forme : $\{x \otimes x - q(x)\cdot 1 ; \forall x \in E\}$.

L'application φ_c de E dans $(C(E; q)$ est alors la composée de l'injection de E dans $\otimes E$ et de la projection de $\otimes E$ sur son quotient. Si E est de dimension n, $C(E; q)$ est de dimension 2^n. Si $e = (e_1, e_2, \ldots, e_n)$ est une base orthogonale de E, les 2^n éléments 1_C, $e_I = e_{i_1}e_{i_2}\ldots e_{i_p}$ avec $(1 \leqslant i_1 < i_2 < \ldots i_p \leqslant n)$ forment une base de $C(E; q)$. En particulier, φ_c est injective, ce qui permet de considérer E comme plongé dans $C(E; q)$ et de l'identifier ainsi à son image $\varphi_c(E)$. (voir en complément le théorème VI.2.A)

Si L est un surcorps commutatif de K et $\widetilde{q}$ l'extension de q à $E\otimes L$ (voir ci-dessus), l'algèbre de Clifford $C(E \otimes_K L; \widetilde{q})$ est évidemment $C(E; q) \otimes_K L$.

Preuve : Puisque $\otimes E$ est engendré par 1 et E, $(\otimes E)/Q$ est engendré par son unité, image $\varphi_c(1)$ de l'unité de $\otimes E$, et par $\varphi_c(E)$ image de E. Si $(\otimes E)/Q$ n'est pas nulle, φ_c est une application de Clifford de E dans $(\otimes E/Q)$ puisque $\varphi_c(x \otimes x - q(x) \cdot 1) = \varphi_c(x) \cdot \varphi_c(x) - q(x)\varphi_c(1) = 0$. Si maintenant φ est une application de Clifford de E dans A, l'application linéaire φ se prolonge en un homomorphisme d'algèbres unitaires ψ de $\otimes E$ dans A (cf. III;12) qui s'annulant sur l'idéal Q en vertu de la condition de Clifford : $\psi(x \otimes x - q(x) \cdot 1) = \varphi(x) \cdot \varphi(x) - q(x) \cdot 1_A = 0$, passe au quotient et détermine un homomorphisme d'algèbres unitaires Φ de $(\otimes E)/Q$

dans A tel que $\Phi \circ \varphi_c = \varphi$; $C(E;q) = (\otimes E)/Q$ possède donc la propriété universelle vis-à-vis des applications de Clifford de $(E;q)$. Cette propriété va nous permettre de démontrer que non seulement $(\otimes E)/Q$ n'est pas nulle mais qu'elle est de dimension 2^n, ce que nous ferons à l'aide de l'application de Clifford standard de E dans $\mathfrak{L}(\wedge E)$ de l'exemple 4 ci-dessus.

Puisque $C(E;q) = (\otimes E/Q)$ possède la propriété universelle, il existe un homomorphisme d'algèbre unitaire Φ de $C(E;q)$ dans $\mathfrak{L}(\wedge E)$ tel que $\Phi(x) = \varphi(x) = d_x + L_x$. On a alors : $\Phi(x_1 x_2 \ldots x_p) = \varphi(x_1)\varphi(x_2)\ldots\varphi(x_p)$. Or, si 1 est l'unité de $\wedge E$ et si x et y sont orthogonaux, $\Phi(xy)\cdot 1 = \varphi(x)\varphi(y)\cdot 1 = (d_x + L_x)y = g(x,y) + x\wedge y = x \wedge y$. On va raisonner par récurrence. Soient $e_{i_1} e_{i_2}, \ldots e_{i_p}$ des vecteurs de la base orthogonale $e = (e_1, e_2, \ldots, e_n)$ de E avec $i_1 < i_2 < \ldots i_p$. Supposons démontrée l'égalité : $\Phi(e_{i_2} e_{i_3}, \ldots e_{i_p})\cdot 1 = e_{i_2} \wedge e_{i_3} \wedge \ldots e_{i_p}$. On a alors :

$$\begin{aligned}
\Phi(e_{i_1} e_{i_2}, \ldots e_{i_p}) \cdot 1 &= \varphi(e_{i_1}) \cdot \Phi(e_{i_2} \ldots e_{i_p}), \\
&= (d_{e_{i_1}} + L_{e_{i_1}}) e_{i_2} \wedge \ldots e_{i_p}, \\
&= e_{i_1} \wedge e_{i_2} \wedge \ldots e_{i_p}.
\end{aligned}$$

L'égalite est ainsi démontré pour toutes les suites strictement croissantes de $(1, 2, \ldots n)$. Les produits de vecteurs e_i correspondants dans $C(E;q)$ pour lesquels nous adoptons la même notation e_I que pour l'algèbre $\wedge E$, en forment donc une base puisque leurs images par l'application linéaire : $u \longrightarrow \Phi(u)\cdot 1$ sont linéairement indépendantes, c.q.f.d.

Corollaire VI.1 : Soient $(E;q)$ un espace quadratique, $e = (e_1, e_2, \ldots, e_n)$ une base orthogonale de E, φ une application de Clifford de E dans l'algèbre associative unitaire A.

Si la sous-algèbre C de A engendrée par 1_A et $\varphi(E)$ est de dimension 2^n, c'est-à-dire si et seulement si les éléments $\{1_A, \varphi(e_{i_1})\varphi(e_{i_2})\ldots\varphi(e_{i_p}); 1 \leqslant i_1 < i_2 < \ldots < i_p < n\}$ de A sont linéairement indépendants, C est une algèbre de Clifford de $(E;q)$ isomorphe à $C(E;q)$.

Ce corollaire permet la détermination rapide de nombreuses algèbres de Clifford.

Exemples :

1) Lorsque la forme quadratique q est nulle, une application de Clifford n'est autre qu'une application alternée (cf. §IV.9) et l'algèbre de Clifford $\mathrm{C}(\mathrm{E}; \mathrm{O})$ est l'algèbre extérieure $\wedge \mathrm{E}$.

2) La droite $\mathrm{K} \cdot e$ munie de la forme quadratique $q(\mu e) = \mu^2 a$ a pour algèbre de Clifford la sous-algèbre (commutative) de dimension 2 de $\mathrm{K}(2)$ formée par les matrices :

$$\begin{vmatrix} \lambda & \mu a \\ \mu & \mu \end{vmatrix}.$$

C'est un corps commutatif si et seulement si $a \notin \mathrm{K}^2$. En particulier $\mathrm{C}(\mathbb{R}; q(x) = -x^2) = \mathbb{C}$.

3) *Définition VI.1.B :* On appelle *algèbre de quaternions* toute algèbre de Clifford d'un plan quadratique régulier. Le plan $\mathbb{R}^2$ muni de la forme quadratique $q(e_1 x + e_2 y) = -x^2 - y^2$ a pour algèbre de Clifford le corps non-commutatif $\mathbb{H}$ des quaternions d'Hamilton généralisant le corps complexe. La représentation matricielle classique de $\mathbb{H}$ dans $\mathbb{C}(2)$ est la suivante. En désignant par E_1 et E_2 les matrices images des vecteurs e_1, e_2 de la base canonique de $\mathbb{R}^2$ et par E_3 le produit $\mathrm{E}_1\mathrm{E}_2$, on définit la représentation matricielle d'un élément de $\mathbb{H}$ par :

$$h = \alpha \mathrm{I} + \beta \mathrm{E}_1 + \gamma \mathrm{E}_2 + \delta \mathrm{E}_3 = \begin{vmatrix} \alpha + i\delta & \beta + i\gamma \\ -\beta + i\gamma & \alpha - i\delta \end{vmatrix},$$

$\mathbb{H}$ peut donc *s'identifier* à la sous-algèbre réelle, de dimension 4, de $\mathbb{C}(2)$ formée par les matrices :

$$h = \begin{vmatrix} u & v \\ -\overline{v} & \overline{u} \end{vmatrix},$$

et l'application de Clifford de $\mathbb{R}^2$ dans $\mathbb{H} \subset \mathbb{C}(2)$ s'écrit :

$$\beta e_1 + \gamma e_2 \longrightarrow \begin{vmatrix} 0 & \beta + i\gamma \\ -\beta + i\gamma & 0 \end{vmatrix}.$$

Cette application définit aussi une application de Clifford dans $\mathbb{C}(2)$ du plan quadratique complexe $(\mathbb{C}^2; q(e_1 x + e_2 y) = -x^2 - y^2)$, complexifié du plan antieuclidien précédent. On en déduit sans peine que l'algèbre des quaternions complexes est l'isomorphe à l'algèbre de matrices $\mathbb{C}(2)$.

4) Pour l'application de Clifford déjà vue de l'espace euclidien E_3 dans $\mathbb{C}(2)$:

$$\varphi(xe_1 + ye_2 + ze_3) = \begin{vmatrix} z & x - iy \\ x + iy & -z \end{vmatrix}$$

l'image $\varphi(\mathrm{E}_3)$ est l'espace des matrices hermitiennes de trace nulle de $\mathbb{C}(2)$.

Définition VI.1.C :

On appelle *matrices de Pauli* les matrices complexes d'ordre deux images des vecteurs de la base canonique de $\mathbb{E}_3$ par l'application de Clifford ci-dessus :

$$\sigma_1 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \quad \sigma_2 = \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix} \quad \sigma_3 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}.$$

On note parfois σ_0 la matrice unité I. Plus généralement, on peut appeler matrices de Pauli tout ensemble de trois matrices $\sigma_1, \sigma_2, \sigma_3$ complexes 2×2 vérifiant les relations : $\sigma_j^2 = \mathrm{I}$, $\sigma_i\sigma_j + \sigma_j\sigma_i = 0 \ \forall i \neq j \in (1,2,3)$ et $\sigma_1\sigma_2\sigma_3 = i\mathbb{1}$. Il en résulte : $\sigma_1\sigma_2 = i\sigma_3$, $\sigma_2\sigma_3 = i\sigma_1$, $\sigma_3\sigma_2 = i\sigma_2$.

Les huit matrices $\sigma_0, \sigma_1, \ldots, \sigma_1\sigma_2\sigma_3$ sont linéairement indépendantes sur $\mathbb{R}$. En vertu du corollaire VI.1. ci-dessus, l'algèbre réelle $\mathbb{C}(2)$ est donc une algèbre de Clifford de l'espace euclidien de dimension trois.

5) La restriction au plan sous-tendu par (e_1, e_3) dans l'application de Clifford ci-dessus montre comment on peut obtenir une application de Clifford, dans l'algèbre des matrices $\mathrm{K}(2)$, du plan quadratique sur K de type euclidien : $(\mathrm{K}^2; q(e_1x+e_2y) = x^2+y^2)$: $\varphi(e_1x + e_2y) = \begin{vmatrix} y & x \\ x & -y \end{vmatrix}$. Si $\mathrm{E}_1 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$, $\mathrm{E}_2 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}$, $\mathrm{E}_3 = \mathrm{E}_1\mathrm{E}_2 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$ elles forment avec la matrice unité une base de $\mathrm{K}(2)$ qui est donc l'algèbre de Clifford de ce plan quadratique.

6) L'espace antieuclidien de dimension trois, $E_{0,3}$ défini comme l'espace réel $\mathbb{R}^3$ muni de la forme quadratique : $q(e_1x^1 + e_2x^2 + e_3x^3) = -(x^1)^2 - (x^2)^2 - (x^3)^2$ possède une application de Clifford naturelle φ dans le corps des quaternions d'Hamilton H, qui est de dimension 4 sur $\mathbb{R}$, alors que $C(E_{0,3})$ est de dimension 8 :

$$\varphi(x) = \varphi(e_1x^1 + e_2x^2 + e_3x^3) = E_1x^1 + E_2x^2 + E_3x^3$$

avec les notations du n° 3) ci-dessus.

Si Ψ est l'application de $E_{0,3}$ dans l'algèbre $\mathbb{H}(2)$ des matrices quaternioniennes d'ordre deux définies par :

$$x \in E_{0,3} \longrightarrow \Psi(x) = \begin{vmatrix} 0 & \varphi(x) \\ \varphi(x) & 0 \end{vmatrix}$$

la matrice unité I et $\Psi(E_{0,3})$ engendrent alors une *sous-algèbre réelle* de dimension 8 de $\mathbb{H}(2)$ qui est une algèbre de Clifford de $E_{0,3}$. Elle est isomorphe à $\mathbb{H} \oplus \mathbb{H}$.

7) L'espace de Minkowski $E_{1,3}$ défini comme l'espace réel $\mathbb{R}^4$ muni de la forme quadratique : $q(u) = q(e_0x^0 + e_1x^1 + e_2x^2 + e_3x^3) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$ a pour algèbre de Clifford $C(E_{1,3})$ une algèbre réelle de dimension $2^4 = 16$. En écrivant un vecteur u de $E_{1,3}$ comme somme de e_0x^0 et d'un vecteur x de $E_{0,3}$ on obtient une application de Clifford Ψ de $E_{1,3}$ dans $\mathbb{H}(2)$, extension de l'application du n° précédent :

$$\begin{aligned}\Psi(u) = \Psi(e_0x^0 + x) &= \begin{vmatrix} x^0 1_{\mathbb{H}} & 0 \\ 0 & -x^0 1_{\mathbb{H}} \end{vmatrix} + \Psi(x) \\ &= \begin{vmatrix} x^0 1_{\mathbb{H}} & \varphi(x) \\ \varphi(x) & -x^0 1_{\mathbb{H}} \end{vmatrix}\end{aligned}$$

et l'on prouve ainsi aisèment que $C(E_{1,3})$ est isomorphe à l'algèbre réelle $\mathbb{H}(2)$.

Or, l'algèbre réelle $\mathbb{H}(2)$ est une forme réelle de l'algèbre complexe $\mathbb{C}(4)$: $\mathbb{C}(4) = \mathbb{H}(2) \oplus i\mathbb{H}(2)$, et on peut donc l'y plonger, ce qui permet de réaliser $C(E_{1,3})$ comme sous-algèbre réelle de l'algèbre $\mathbb{C}(4)$. On est ainsi conduit à rechercher les applications de Clifford de $E_{1,3}$ dans $\mathbb{C}(4)$.

Une telle application est entièrement déterminée par les images $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ des vecteurs de la base canonique de $E_{1,3}$, ce qui conduit à la définition :

Définition VI.1.D :

Un ensemble de *matrices de Dirac* est un ensemble de quatre matrices complexes d'ordre quatre : $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ satisfaisant aux relations :

a) $(\gamma_0)^2 = 1$, $(\gamma_1)^2 = (\gamma_2)^2 = (\gamma_3)^2 = -1$

b) d'anticommutativité deux à deux : $\gamma_i\gamma_j + \gamma_j\gamma_i = 0$, $\forall i \neq j$ qui expriment qu'elles définissent une application de Clifford de $\mathbb{E}_{1,3}$ dans $\mathbb{C}(4)$.

On peut, par exemple, prendre pour γ_0 l'une ou l'autre des matrices : $\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}$ ou $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ et pour les autres $\gamma_j = \begin{vmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{vmatrix}$ où les σ_j, $j = 1, 2, 3$ sont les matrices de Pauli.

Les exemples précédents montrent l'intérêt pratique de représenter les éléments d'une algèbre de Clifford C par des matrices, c'est-à-dire de construire une « *représentation matricielle* », de C, isomorphisme de C sur une sous-algèbre, ou quand cela est possible, une algèbre, de matrices.

Ce point de vue sera systématisé au §VI.4 ci-dessous.

Dans les sept exemples développés ci-dessus, on a pu déterminer les algèbres de Clifford d'espaces quadratiques de petites dimensions grâce à des applications de Clifford astucieusement choisies dans chaque cas. Nous allons, dans les deux exemples ci-dessous, déterminer cette fois toute une famille d'algèbres de Clifford.

8) *Théorème VI.1.B :* L'algèbre de Clifford C(H) de l'espace *hyperbolique* H de dimension $2r$ sur le corps K est isomorphe à l'algèbre de matrices $K(2^r)$. Si l'on représente H comme somme directe de deux sous-espaces totalement isotropes maximaux : $H = F \oplus F'$, C(H) s'identifie naturellement à l'algèbre $\mathfrak{L}(\wedge F)$ des endomorphismes de l'espace vectoriel $\wedge F (= C(F))$.

Preuve : Tout vecteur x de H s'écrit de façon unique :

$$x = \zeta + \alpha, \text{ avec } \zeta \in F, \ \alpha \in F', \text{ et } q(x) = 2g(\alpha, \zeta).$$

Appelons $\rho(\zeta)$ la multiplication extérieure par ζ dans $\wedge$F et $\rho(\alpha)$ l'antidérivation (Thm IV.13.B) de l'algèbre $\wedge$F déterminée par la forme linéaire sur F : $\zeta \to 2g(\alpha;\zeta)$. L'application de H dans $\mathfrak{L}(\wedge \mathrm{F})$:

$$x = \zeta + \alpha \to \rho(x) = \rho(\zeta) + \rho(\alpha)$$

est alors une application de Clifford. En effet, si $f_1 \wedge f_2 \wedge \dots f_p \in \wedge^p \mathrm{F}$:

$$\rho(\zeta) \cdot f_1 \wedge \dots f_p = \zeta \wedge f_1 \wedge \dots f_p \\ + \sum (-1)^{i-1} 2g(\alpha; f_i) f_1 \wedge \dots \widehat{f_i} \wedge \dots f_p$$

Puisque $\rho(\zeta^2) = 0$ et $\rho(\alpha)^2 = 0$, $\rho(x)^2 = \rho(\zeta)\rho(\alpha) + \rho(\zeta)\rho(\zeta)$, et

$$\rho(\zeta)\rho(\alpha) f_1 \wedge f_2 \wedge \dots f_p = \sum (-1)^{i-1} 2g(\alpha; f_i) \zeta \wedge f_1 \dots \widehat{f_i} \wedge \dots f_p \\ \rho(\alpha)\rho(\zeta) f_1 \wedge f_2 \wedge \dots f_p = 2g(\alpha;\zeta) f_1 \wedge f_2 \wedge \dots f_p \\ + \sum (-1)^{i} 2g(\alpha; f_i) \zeta \wedge f_1 \wedge \dots \widehat{f_i} \wedge \dots f_p.$$

En ajoutant, il vient $\rho(x)^2 = q(x) \cdot \mathrm{Id}$.

ρ se prolonge donc en un homomorphisme $\widetilde{\rho}$ d'algèbres unitaires de C(H) dans $\mathfrak{L}(\wedge \mathrm{F})$. Un examen des bases et l'utilisation du théorème VI.2.A ci-dessous, ou le fait que la dimension de H étant paire, l'algèbre C(H) est simple (théorème VI.4.A) avec $\dim \mathrm{C(H)} = \dim (\mathfrak{L}(\wedge \mathrm{F})$ montrent que $\widetilde{\rho}$ est un isomorphisme c.q.f.d.

9) *La construction de Brauer-Weyl.*

Cette construction généralise celle de l'algèbre de Clifford $\mathrm{C}(\mathrm{E}_3)$ à l'aide des matrices de Pauli $\sigma_1\sigma_2\sigma_3$ (Définition VI.1.C). Soit E l'espace quadratique complexe standard de dimension n, de base orthonormale $(e_1, e_2, \dots, e_n)$ avec $q(x) = \sum (x^j)^2$.

Si $n = 2r$ ou $2r+1$, soit A le produit tensoriel d'algèbres complexes de r copies de l'algèbre de matrices $\mathbb{C}(2)$. A s'identifie à $\mathbb{C}(2^r)$: $\mathrm{A} = \mathbb{C}(2) \otimes \mathbb{C}(2) \otimes \dots \otimes \mathbb{C}(2) \equiv \mathbb{C}(2^r)$.

Définissons une application p de E dans A. Pour $1 \leqslant j \leqslant r$:

$$e_j \longrightarrow p_j = \underbrace{\sigma_3 \otimes \dots \sigma_3}_{(j-1)} \otimes \sigma_1 \otimes \underbrace{\dots \ I \otimes I}_{(r-1)}$$

$$e_{r+j} \longrightarrow p_{r+j} = \overbrace{\sigma_3 \otimes \dots \sigma_3}^{(j-1)} \otimes \sigma_2 \quad \overbrace{\otimes I \otimes \dots I}^{(r-1)}$$

et si n est impair $(n = 2r + 1)$:

$$e_n = e_{2r+1} \rightarrow p_{2r+1} = \sigma_3 \otimes \sigma_3 \otimes \sigma_3 \quad (r \text{ fois}).$$

On vérifie immédiatement que c'est une application de Clifford : $p_k^2 = l$, et si $k \neq l$: $p_k p_l = -p_l p_k$.

Lorsque n est impair, l'élément $u = p_1 p_2 \dots p_n$ de A commute avec tous les autres et est égal à $i\mathbb{1}$ si $n = 4s+3$, à $\mathbb{1}$ si $n = 4s+1$.

On voit alors directement que les p_j engendrent toute l'algèbre A. Si n est pair, on a alors : $\mathrm{C}(\mathrm{E}_{2r}) \equiv \mathbb{C}(2^r)$.

Mais si n est impair, l'application de $\mathrm{C}(\mathrm{E}_{2r+1})$ dans $\mathrm{C}(2^r)$ n'est pas injective puisque les deux éléments de la base de $\mathrm{C}(\mathrm{E}_{2r+1})$: 1_{C} et $e_1 e_2 \dots e_n$ ont des images liées.

En composant la symétrie de E_n par rapport à l'origine avec p, on obtient une autre application de Clifford p' de E' dans A : $e_j \rightarrow p'_j = -p_j$. Mais puisque l'image de $e_1 e_2 \dots e_n$ est $p'_1 p'_2 \dots p'_n = -p_1 p_2 \dots p_n$, cette représentation de $\mathrm{C}(\mathrm{E}_{2r+1})$ dans A est inéquivalente à la première (elle ne peut résulter d'un automorphisme intérieur de $\mathbb{C}(2^r)$, voir la Remarque, §VI.1.2.). Une vérification élémentaire montre cette fois que $\mathbb{C}(\mathrm{E}_{2r+1}) \equiv \mathbb{C}(2^r) \oplus \mathbb{C}(2^r)$ est la somme directe de deux algèbres de matrices complexes.

VI.2 – Structures sur les algèbres de Clifford

On peut jouer de la propriété universelle pour obtenir deux propriétés importantes des algèbres de Clifford : l'existence d'un automorphisme Π et d'un antiautomorphisme τ, involutifs, liés à la structure, et appelés pour cela *principaux*. Soit $\mathrm{C} = \mathrm{C}(\mathrm{E}; q)$. Si u est un automorphisme orthogonal de E : $q(ux) = q(x)$, $\forall x \in \mathrm{E}$, l'application composée $\Psi : x \in \mathrm{E} \rightarrow \Psi(x) = ux$ dans C est encore une application de Clifford : $\Psi(x)^2 = (ux)^2 = q(ux) \cdot 1 = q(x) \cdot 1$. Il existe donc un automorphisme U de C qui prolonge

u par $\mathrm{U}(x_1 x_2 \ldots x_p) = (ux_1)(ux_2)\ldots(ux_p)$. On obtient ainsi une représentation du groupe orthogonal $\mathrm{O}(\mathrm{E}; q)$ comme groupe d'automorphismes de l'algèbre unitaire C. En particulier, en prenant pour $u : (-\mathrm{Id}_\mathrm{E})$, soit $\Psi(x) = -x$, on obtient *l'automorphisme involutif de* C. Π peut donc être défini comme l'unique automorphisme de C qui change les signes des vecteurs de E. Le sous-espace propre de Π pour la valeur propre $+1$ est une sous-algèbre, dite *sous-algèbre paire* de C et notée C_+. Elle contient tous les produits en nombre pair de vecteurs de E. Le sous-espace propre C_- contient tous les produits en nombre impair de vecteurs de E et en particulier E lui-même. Si $e = (e_1, e_2, \ldots, e_n)$ est une base orthogonale de E, les e_I, $|\mathrm{I}|$ pair, forment une base de C_+, les e_I, $|\mathrm{I}|$ impair de C_-. On a donc $\dim \mathrm{C}_+ = \sum\limits_k \binom{n}{2k} = 2^{n-1} = \dim \mathrm{C}_- = \sum\limits_k \binom{n}{2k+1}$. La décomposition de C en somme directe : $\mathrm{C}_+ \oplus \mathrm{C}_-$ peut être considérée comme une *graduation* de l'algèbre C dans $\mathbb{Z}_2$.

Remarque : il est utile dans les applications de savoir si l'automorphisme Π est, ou non, un automorphisme intérieur de $\mathrm{C}(\mathrm{E}; q)$, c'est-à-dire s'il existe un élément u tel que $\Pi a = ua\overset{-1}{u}$, $\forall a \in \mathrm{C}(\mathrm{E}; q)$.

Or ceci entraîne pour tout vecteur x de E : $\Pi x = -x = ux\overset{-1}{u}$ soit $ux = -xu$, $uxy = xyu$, etc... : l'élément u, anticommute avec les éléments impairs, commute avec les éléments pairs. De tels éléments forment l'« *anticentre* » de $\mathrm{C}(\mathrm{E}; q)$. Lorsque $(\mathrm{E}; q)$ est *régulier* de dimension *impaire*, l'anticentre est nul : Π *n'est pas* un automorphisme intérieur. On a vu une application de ce fait dans la construction de Brauer-Weyl (Exemple 9, fin du §VI.1)

Considérons maintenant une nouvelle algèbre C^0 obtenue en définissant sur l'espace vectoriel C une autre multiplication, notée par un astérisque. Le nouveau produit $a * b$ de deux éléments a et b est l'ancien produit de ces éléments dans l'ordre inverse : $a * b = ba$. La nouvelle algèbre C^0, appelée l'opposée de C, est, comme elle, associative et unitaire, et contient E. On a, si $x \in \mathrm{E}$, $x * x = xx = x^2 = q(x) \cdot 1$, et C^0 est donc aussi une algèbre de Clifford pour $(\mathrm{E}; q)$. La propriété universelle implique donc l'existence d'un isomorphisme τ d'algèbres unitaires de C sur C^0 que l'on peut plus pratiquement interpréter comme une *antiautomorphisme* (involutif) de C. C'est l'unique antiautomorphisme laissant fixes les éléments de E : $\tau(ab) = \tau b \cdot \tau a$. Si $x_1, x_2, \ldots x_p \in$ E, $\tau(x_1, x_2, \ldots, x_p) = x_p x_{p-1} \ldots x_1$.

Puisque $\Pi \circ \tau$ et $\tau \circ \Pi$ sont deux antiautomorphismes de C coïncidant sur les générateurs : 1 et E, ils sont égaux.

$\nu = \Pi\tau = \tau\Pi$ est l'unique antiautomorphisme de C qui change le signe des vecteurs de E. Il est involutif et appelé *conjugaison* de C. On note simpement $\breve{u} = \nu u$ le conjugué de u.

Remarque : Une représentation matricielle de C est d'autant plus maniable que Π, τ et ν ont une expressions simple. (On verra au théorème VI.6.B que toute antiinvolution de $\mathrm{K}(m)$ est composée de la transposition et d'un antomorphisme intérieur).

Définition VI.2.A :

L'application quadratique de $\mathrm{C} = \mathrm{C}(\mathrm{E}; q)$ en elle-même : $u \to \mathrm{N}(u) = u\breve{u}$ est appelée la *norme* de C. L'application : $u \to \mathrm{S}(u) = u(\tau u)$ est la *norme spinorielle*. Elles coïncident sur C_+.

Si u est décomposable, c'est-à-dire le produit de p vecteurs de E : $u = x_1, x_2, \ldots, x_p$, $\mathrm{S}(u) = x_1, x_2, \ldots, x_p x_p, \ldots x_2 x_1 = q(x_1)q(x_2)\ldots q(x_p)$ et $\mathrm{N}(u) = (-1)^p \mathrm{S}(u)$ sont à valeurs scalaires.

Exemples :

1) Soient E un plan quadratique régulier, e_1, e_2 une base orthogonale de E, $e_3 = e_1 e_2$ le produit de ces vecteurs dans l'algèbre de Clifford C de E. C est donc une algèbre de quaternions, et si $h = \alpha \cdot 1 + \beta e_1 + \gamma e_2 + \delta e_3 \in$ C. $\Pi h = \alpha \cdot 1 - \beta e_1 - \gamma e_2 + \delta e_3$, $\tau h = \alpha \cdot 1 + \beta e_1 + \gamma e_2 - \delta e_3$ tandis que le conjugué $\nu h = \breve{h}$ est égal à $\breve{h} = \alpha \cdot 1 - \beta \varepsilon_1 - \gamma e_2 - \delta e_3$. Si $q(e_1) = a$ et $q(e_2) = b$, la norme de h est le scalaire : $\mathrm{N}(h) = h\breve{h} = \alpha^2 - a\lambda^2 - b\mu^2 + ab\delta^2$. C'est une forme quadratique multiplicative sur C. $\mathrm{N}(hh') = hh'(\breve{hh'}) = hh'\breve{h}'\breve{h} = \mathrm{N}(h)\mathrm{N}(h')$. (Tous les quaternions non nuls sont décomposables.) Un quaternion est inversible si et seulement si sa norme est non nulle. En effet, de $h \overset{-1}{h} = 1$ on déduit $\mathrm{N}(h)\mathrm{N}(h)^{-1} = 1$. Réciproquement, si $\mathrm{N}(h) \neq 0$, $\overset{-1}{h} = \breve{h}/\mathrm{N}(h)$.

Sur l'algèbre $\mathrm{K}(2)$ des matrices 2×2 à coefficients dans K, on a une antiinvolution canonique : $\mathrm{A} \to \breve{\mathrm{A}}$, que nous appellerons conjugaison, définie par :

si $A = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$, $\breve{A} = \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}$, soit, si $C = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$, par $\breve{A} = C^t A \overset{-1}{C}$.

On a : $A\breve{A} = \breve{A}A = (\text{dét } A)I = N(A)$.

Dans la représentation matricielle classique des quaternions réels et complexes vues à l'exemple 3 du paragraphe précédent la conjugaison de $\mathbb{C}(2)$ est celle des quaternions. Dans le cas des quaternions réels, et seulement dans ce cas, $\breve{A} =^t \overline{A}$

La norme des quaternions d'Hamilton est $N(h) = N(\alpha 1 + \beta e_1 + \gamma e_2 + \delta e_3) = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$. Sous forme matricielle dans $\mathbb{C}(2)$:

$$\text{si} \quad h = \begin{vmatrix} u & v \\ -\overline{v} & \overline{u} \end{vmatrix}, \ N(h) = \text{dét } h = u\overline{u} + v\overline{v} = |u|^2 + |v|^2$$

$$\text{et } h\breve{h} = h^t\overline{h} = N(h)I$$

Tout élément non nul de $\mathbb{H}$ est donc inversible : $\mathbb{H}$ est un corps non commutatif.

La propriété $N(hh') = N(h)N(h')$ exprime le produit de deux sommes de quatre carrés comme somme de quatre carrés (formule d'Euler).

2) Si (e_1, e_2, e_3) est une base orthogonale ordonnée de l'espace euclidien E_3, elle y définit une orientation.

L'élément $J = e_1e_2e_3$ de l'algèbre de Clifford C de E_3 a pour carré $J^2 = -1$ et définit donc une structure complexe sur C, qui devient la structure opposée $(-J)$ si l'on change l'orientation. Sur la réalisation matricielle $\mathbb{C}(2)$ de $C(E_3)$ construite au paragraphe précédent : $J = \sigma_1\sigma_2\sigma_3 = iI$, ce qui est bien la structure complexe naturelle de $\mathbb{C}(2)$. Π et τ changent J en $(-J)$ et sont donc antilinéaires. La conjugaison ν laisse J invariant et change le signe des matrices de Pauli :

$$\nu J = \nu(e_1e_2e_3) = (-e_3)(-e_2)(-e_1) = -e_3e_2e_1 = e_1e_2e_3 = J$$

ν est donc $\mathbb{C}$-linéaire. On a : $\nu A = \breve{A}$ (voir 1); ci-dessus), $\tau A =^t \overline{A}$, $\Pi A =^t \breve{\overline{A}}$.

Une autre structure de l'algèbre de Clifford $C(E; q) = C$ est celle de « filtration ». Appelons $C^{(k)}$ le sous-espace vectoriel de C formé des combinaisons linéaires de produits d'au plus k vecteurs de E. $C^{(k)}$ est l'image de $\left(\sum_0^k \otimes^k E\right)$ dans l'expression de $C(E;q)$ comme quotient $(\otimes E)/Q$. Les $C^{(k)}$, $k = 0, 1, \ldots$ forment une suite croissante de sous-espaces emboîtés : $(C^{(0)} = K \cdot 1) \subset (C^{(1)} = K \cdot 1 + E) \subset \ldots$ ou filtration de C.

Théorème VI.2.A :

1) Si $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ est une base *quelconque* de l'espace quadratique $(E;q)$, les 2^n éléments de l'algèbre de Clifford $C = C(E;q)$: $\{1, \varepsilon_J = \varepsilon_{j_1}\varepsilon_{j_2}\ldots\varepsilon_{j_p};\ 1 \leqslant j_1 < \ldots j_p \leqslant n\}$ en forment une base.

2) Si F est un sous-espace vectoriel quelconque de E, q_F la restriction de q à F, la sous-algèbre C_F de C engendrée par 1_C et F est une algèbre de Clifford pour $(F;q_F)$.

Preuve :

1) Si $e = (e_1, e_2, \ldots, e_n)$ est une base orthogonale de E, on peut exprimer les e_j en fonction des ε_k. En utilisant les relations de Clifford entre les ε_k : $\varepsilon_k^2 = q(\varepsilon_k) \cdot 1$ et $\varepsilon_j\varepsilon_k + \varepsilon_k\varepsilon_j = 2g(\varepsilon_j\varepsilon_k) \cdot 1$ on prouve immédiatement que les e_I, $|I| = p$ sont des combinaisons linéaires des ε_I, $|I| = p$ modulo le sous-espace $C^{(p-1)}$ défini ci-dessus.

2) est prouvé immédiatement en prenant une base ε de E dont les p premiers vecteurs forment une base de F et en utilisant 1).

Théorème VI.2.B : Si la forme quadratique q n'est pas nulle, la sous-algèbre paire C_+ de $C(E;q)$ peut être considérée comme l'algèbre de Clifford de tout sous-espace $E_1 = e_1^\perp$ de E, orthogonal d'un vecteur régulier e_1 pour la forme quadratique $q_1 = -q(e_1)q$. Pour une telle structure d'algèbre de Clifford sur C_+, la conjugaison ν_1 est toujours la restriction de la conjugaison ν de C, et l'automorphisme principal Π_1 est l'automorphisme intérieur de C_+ par e_1.

Preuve : Pour tout $y \in e_1^\perp$, y et e_1 anticommutent dans C. Si $q(e_1) \neq 0$, $e_1^\perp$ est un sous-espace de codimension un dans E supplémentaire de e_1. L'application $y \in e_1^\perp \to \varphi_1(y) = e_1 y \in C_+$ est une application de Clifford pour la forme quadratique $q_1 = -q(e_1)q$ sur $e_1^\perp$ puisque $\varphi_1(y)^2 = e_1 y e_1 y = -e_1^2 y^2 = -q(e_1)q(y)$.

On vérifie immédiatement que l'unité et $\varphi_1(e_1^\perp)$ engendrent C_+, d'où : $C_+ = C(e_1^\perp; q_1)$. La restriction de ν à C_+ est un antiautomorphisme dont la restriction à $\varphi_1(e_1^\perp)$ est : $\nu(\varphi_1(y)) = \nu(e_1 y) = (-y)(-e_j) = ye_1 = -e_1 y = -\varphi_1(y) = \nu_1\varphi_1(y)$ d'où $\nu = \nu_1$ sur C_+. De la même façon, $e_1\varphi_1(y)\overline{e}_1^1 = e_1(e_1 y)\ \dfrac{e_1}{q(e_1)}$ $= -e_1 y = -\varphi_1(y)$. L'automorphisme principal Π_1 coïncide donc avec l'automorphisme intérieur par e_1, et dépend ainsi du choix de e_1, contrairement à ν_1.

Exemple : Soient E_3 l'espace euclidien de dimension trois et (e_1, e_2, e_3) une base orthonormale de E_3. La sous-algèbre paire C_+ de $C(E_3)$ est l'algèbre de Clifford du plan $e_1^\perp = (e_2, e_3)$ muni de la forme quadratique $q_1 = -q$, c'est-à-dire d'un plan antieuclidien. $C_+(E_3)$ est donc isomorphe au corps non-commutatif $\mathbb{H}$ des quaternions d'Hamilton, et puisque $J = e_1e_2e_3$ définit la structure complexe de $C(E_3) = \mathbb{C}(2)$, on a $\mathbb{C}(2) = \mathbb{H} \oplus i\mathbb{H}$, ce qui montre que $\mathbb{H}$ est une forme réelle de l'algèbre complexe $\mathbb{C}(2)$ ($\mathbb{R}(2)$ en est une autre).

On a déjà observé au paragraphe précédent VI.1 qu'un premier problème dans l'étude des algèbres de Clifford est un problème concret : étant donné un espace quadratique particulier, peut-on calculer son algèbre de Clifford? Existe-t'il, par exemple, des procédés généraux de récurrence sur la dimension permettant de déduire la structure de $C(E; q)$ de la connaissance des algèbres de Clifford des sous-espaces de E? Le théorème VI.2.B et le théorème suivant permettent une approche théorique du problème de récurrence :

Théorème VI.2.C : Si $(E; q) = (E_1, q_1)\oplus(E_2; q_2)$ est la somme directe orthogonale de deux sous-espaces E_1, E_2, son algèbre de Clifford $C(E; q)$ est naturellement isomorphe au produit tensoriel gauche (Définition III.11.C) de leurs algèbres de Clifford :

$$C(E; q) = C(E_1, q_1)\widehat{\otimes}C(E_2, q_2).$$

En particulier, si $q \equiv 0$: $\wedge E = (\wedge E_1)\widehat{\otimes}(\wedge E_2)$.

Preuve : Tout vecteur x de E se décompose en $x = x_1 + x_2$ suivant la décomposition de E en somme directe $E = E_1 \oplus E_2$.

L'application de E dans $C(E_1; q_1)\widehat{\otimes}C(E_2; q_2)$:

$$x \xrightarrow{\varphi} x_1 \otimes 1 + 1 \otimes x_2$$

est une application de Clifford puisque (Déf. III.11.C) :

$$\begin{aligned}\varphi(x)^2 &= (x_1 \otimes 1 + 1 \otimes x_2)(x_1 \otimes 1 + 1 \otimes x_2)\\ &= x_1^2 \otimes 1 + x_1 \otimes x_2 - x_1 \otimes x_2 + 1 \otimes x_2^2\\ &= (q(x_1) + q(x_2))1 \otimes 1 = q(x) \cdot 1.\end{aligned}$$

Elle se prolonge en un homomorphisme d'algèbre Φ. La réunion de bases orthogonales de E_1 et de E_2 est une base orthogonale de E et les éléments des bases associées de $C(E; q)$ et $C(E_1; q_1)\widehat{\otimes}C(E_2; q_2)$ se correspondent par Φ, c.q.f.d.

Remarque : la manipulation des produits tensoriels gauches est malcommode et l'on recherche plutôt des théorèmes faisant intervenir le produit tensoriel ordinaire ! (cf. Exercices VI.3).

VI.3 – Groupes de Clifford et groupes spinoriels

Définition VI.3.A : Le *groupe de Clifford* $G(E; q)$ d'un espace quadratique *régulier* est le groupe que forment pour la multiplication les éléments de l'algèbres de Clifford $C(E; q)$ qui sont produits de vecteurs réguliers de E. Le *groupe de Clifford pair* $G_+(E; q)$ est le sous-groupe de $G(E; q)$ que forment les produits d'un nombre pair de vecteurs réguliers de E, soit : $G_+(E; q) = G(E; q) \cap C_+$.

Puisque les éléments de $G(E; q)$ sont décomposables par définition, les restrictions des normes N et S à $G(E; q)$ sont à valeurs scalaires et par conséquent multiplicatives $N(uv) = (uv)(\check{uv}) = u(v\check{v})\check{u} = N(u)N(v)$ et de même $S(uv) = S(u)S(v)$. N et S sont donc des homomorphismes du groupe $G(E; q)$ dans le groupe multiplicatif K^* des éléments non nuls du corps des scalaires.

Définition VI.3.B : On appelle *groupe spinoriel* $\mathrm{Spin}(E; q)$ *de l'espace quadratique régulier* $(E; q)$ le sous-groupe du groupe de Clifford pair $G_+(E; q)$ noyau de l'homomorphisme norme, c'est-à-dire le sous-groupe de $G_+(E; q)$ formé des éléments de norme un.

Exemples :

1) Déterminons le groupe spinoriel de l'espace euclidien usuel de dimension trois : $\mathrm{Spin}(\mathrm{E}_3) = \mathrm{Spin}(3)$. Notons $\mathrm{C} = \mathrm{C}(\mathrm{E}_3)$. La sous-algèbre paire C_+ est le corps $\mathbb{H}$ des quaternions d'Hamilton. Nous allons vérifier que tous les éléments non nuls de $\mathrm{C}_+ = \mathbb{H}$ sont décomposables et, plus précisément, peuvent s'écrire comme produit de deux vecteurs de E_3. Le groupe de Clifford $\mathrm{G}_+(\mathrm{E}_3)$ est donc le groupe multiplicatif $\mathbb{H}^+$ des quaternions non nuls, et le groupe spinoriel Spin(3) est le groupe des quaternions de norme un. Désignons par $\sigma_1, \sigma_2, \sigma_3$ les éléments de C correspondant à une base orthonormale orientée de E_3 et simplement par i l'élément $\sigma_1\sigma_2\sigma_3$ qui détermine une structure d'algèbre complexe sur C. Dans la réalisation de C comme $\mathbb{C}(2)$ vue au §VI.1, les σ_j sont des matrices de Pauli. $\{1, \sigma_2\sigma_3 = i\sigma_1,\ \sigma_3\sigma_1 = i\sigma_2,\ \sigma_1\sigma_2 = i\sigma_3\}$ forment une base de C_+.

Si $\mathrm{X} = x^1\sigma_1 + x^2\sigma_2 + x^3\sigma_3$ et $\mathrm{Y} = y^1\sigma_1 + y^2\sigma_2 + y^3\sigma_3$ sont les éléments de C correspondant à des vecteurs x et y de E, leur produit, élément de C_+, s'écrit :

$$\mathrm{XY} = (x \mid y) \cdot 1 + i\{(x^2y^3 - x^3y^2)\sigma_1 + (x^3y^1 - x^1y^3)\sigma_2 \\ +(x^1y^2 - x^2y^1)\sigma_3\};$$

si $z = x \times y$ est le produit vectoriel de x et de y, on a donc :

$$\mathrm{XY} = (x \mid y) \cdot 1 + i\mathrm{Z}.$$

Réciproquement, tout élément non nul h de C_+ s'écrit $\rho \cdot 1 + i\mathrm{U}$, ρ et U non tous deux nuls, et les équations en x, y : $(x \mid y) = \rho$, $x \times y = u$ possèdent une infinité de solutions. h peut donc toujours s'écrire $h = \mathrm{XY}$. Calculons la norme quaternionienne de $h = \mathrm{XY}$ $= \rho \cdot 1 + i\mathrm{Z}$:

$$\mathrm{N}(h) \cdot \mathrm{I} = \mathrm{YXXY} = \|x\|^2 \|y\|^2 \cdot \mathrm{I},$$

soit $\mathrm{N}(h) = \|x\|^2\|y\|^2 = \rho^2 + \|z\|^2 = (x \mid y)^2 + \|x \times y\|^2$.

Dans la représentation matricielle standard des quaternions d'Hamilton : $\mathrm{N}(h) = \det \mathrm{X} \det \mathrm{Y} = \det h$ et $\mathrm{N}(h)\mathrm{I} =^t \overline{h}h$. Les quaternions de norme un s'identifient donc aux matrices spéciales unitaires de $\mathbb{C}(2)$: ${}^t\overline{h}h = \mathrm{I}$ et $\det h = +1$.

Le groupe spinoriel Spin(3) *s'identifie ainsi au groupe spécial unitaire* SU(2) *:*

$$h \in \mathrm{Spin}(3) \Longleftrightarrow h = \begin{vmatrix} u & v \\ -\overline{v} & \overline{u} \end{vmatrix} \text{ avec } u\overline{u} + v\overline{v} = 1.$$

En identifiant E_3 à son image dans $\mathbb{C}(2)$, les vecteurs de $\mathrm{E} : x = e_1 x^1 + e_2 x^2 + e_3 x^3$ sont représentés par les matrices hermitiennes de trace nulle :

$$\mathrm{X} = \begin{vmatrix} x^3 & x^1 - ix^2 \\ x^1 + ix^2 & -x^3 \end{vmatrix} \quad \text{et} \quad \mathrm{X}^2 = \{(x^1)^2 + (x^2)^2 + (x^3)^2\} \cdot \mathrm{I}$$
$$= -\text{dét}\,\mathrm{X} \cdot \mathrm{I}.$$

On obtient élémentairement l'épimorphisme canonique de $\mathrm{Spin}(3) = \mathrm{SU}(2)$ sur le groupe des rotations $\mathrm{SO}(3)$ en faisant correspondre à tout élément h de $\mathrm{SU}(2)$ une rotation $\widetilde{h}$ de E_3 par : si $h \in \mathrm{SU}(2)$, $\mathrm{Y} = \widetilde{h}(\mathrm{X}) = h\mathrm{X}\overset{-1}{h}$. En effet les $h\mathrm{X}\overset{-1}{h}$ parcourent les matrices hermitiennes de trace nulle et $\mathrm{Y}^2 = \mathrm{X}^2$ (ou $\text{dét}\,\mathrm{Y} = \text{dét}\,\mathrm{X}$).

2) L'extension à l'espace de Minkowski $\mathrm{E}_{1,3}$ des considérations élémentaire précédentes conduit à identifier $\mathrm{E}_{1,3}$ à l'espace vectoriel des matrices hermitiennes 2×2 : à $x = e_0 x^0 + e_1 x^1 + e_2 x^2 + e_3 x^3$ correspond

$$\mathrm{X} = \begin{vmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{vmatrix}.$$

L'application $x \to \mathrm{X}$ *n'est pas* une application de Clifford mais on récupère la métrique de $\mathrm{E}_{1,3}$ avec le déterminant de la matrice : $\text{dét}\,\mathrm{X} = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = q(x)$.

Une analogie *partielle* avec l'exemple précédent conduit à considérer le groupe $\mathrm{Sl}(2;\mathbb{C})$ des matrices complexes 2×2 de déterminant $+1$, et de le faire opérer sur $\mathrm{E}_{1,3}$ par : si $g \in \mathrm{Sl}(2;\mathbb{C})$, $\mathrm{Y} = g\mathrm{X}{}^t\overline{g}$. En effet, les $g\mathrm{X}{}^t\overline{g}$ parcourent les matrices hermitiennes, et puisque $\text{dét}\,g = 1$, $\text{dét}\,\mathrm{Y} = q(y) = \text{dét}\,\mathrm{X} = q(x)$. On vérifie que l'orientation de l'axe des temps est conservée et on obtient un épimorphisme du groupe connexe et simplement connexe. $\mathrm{Sl}(2;\mathbb{C})$ sur $\mathrm{SO}(1,3)$, de noyau $(-1,+1)$.

Nous allons montrer que $\mathrm{Sl}(2;\mathbb{C})$ s'identifie au groupe $\mathrm{Spin}(1,3)$ (définition VI.3.B). Pour cela formons une application de Clifford de $\mathrm{E}_{1,3}$ dans $\mathbb{C}(4)$ à l'aide de l'application ci-dessus :

$$x \in \mathrm{E}_{1,3} \longrightarrow \mathfrak{X} = \begin{vmatrix} 0 & \mathrm{X} \\ \check{\mathrm{X}} & 0 \end{vmatrix} \quad \text{où} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix}^{\vee} = \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}$$

(cf. §VI.2). On a bien $\mathfrak{X}^2 = q(x) \cdot \mathrm{I}$ puisque $\mathrm{X}\check{\mathrm{X}} = \check{\mathrm{X}}\mathrm{X} = \mathrm{dét}\,\mathrm{X} \cdot \mathrm{I}$ $= q(x) \cdot \mathrm{I}$. Le produit $\mathfrak{X}\mathfrak{Y}$ des matrices associées à deux vecteurs s'écrit : $\mathfrak{X}\mathfrak{Y} = \begin{vmatrix} \mathrm{X}\check{\mathrm{Y}} & O \\ O & \check{\mathrm{X}}\mathrm{Y} \end{vmatrix}$ et sa norme dans $\mathrm{C}(\mathrm{E}_{1,3})$ est $\mathrm{N}(xy) =$ $\mathrm{dét}\,\mathrm{X} \cdot \mathrm{dét}\,\mathrm{Y} = \mathrm{dét}\,|\mathrm{X}\check{\mathrm{Y}}|$.

On a vu au §VI.2 que l'automorphisme principal Π de $\mathrm{C}(\mathrm{E}_3)$ s'écrivait, dans la représentation de $\mathrm{C}(\mathrm{E}_3)$ par $\mathbb{C}(2)$, $\mathrm{A}^{\Pi} = {}^t\check{\bar{\mathrm{A}}}$. Puisque X et Y sont hermitiennes $\check{\mathrm{X}}\mathrm{Y} = (\mathrm{XY})^{\Pi}$. On voit alors immédiatement (cf. DR) que la sous-algèbre paire $\mathrm{C}_+(\mathrm{E}_{1,3})$, qui s'identifie à l'algèbre de Clifford $\mathrm{C}(\mathrm{E}_3)$ (Théorème VI.2.B : $\mathrm{E}_3 \equiv \{e_0^{\perp}$ muni de la forme $-q\}$) est ici réalisée comme la sous-algèbre de $\mathbb{C}(4)$ que forment les matrices $\begin{vmatrix} a & 0 \\ 0 & a^{\Pi} \end{vmatrix}$, et qui est bien isomorphe à $\mathbb{C}(2)$.

$\mathrm{C}(\mathrm{E}_{1,3})$ est réalisée comme la sous-algèbre réelle de $\mathbb{C}(4)$ formée des matrices $g = \begin{vmatrix} a & b \\ b^{\Pi} & a^{\Pi} \end{vmatrix}$. Le groupe de Clifford pair $\mathrm{G}_+(\mathrm{E}_{1,3})$ s'identifie alors au sous-groupe de $\mathrm{Gl}(2;\mathbb{C})$ que forment les matrices g de déterminant *réel* et $\mathrm{Spin}(1,3)$ à $\mathrm{Sl}(2;\mathbb{C})$.

La condition $g \in \mathrm{Sl}(2;\mathbb{C})$ est équivalente à $\check{g} = \overset{-1}{g}$, et l'action de $\mathrm{Spin}(1,3) = \mathrm{Sl}(2;\mathbb{C})$ dans $\mathrm{E}_{1,3}$ s'écrit :

$$\begin{vmatrix} g & 0 \\ 0 & g^{\Pi} \end{vmatrix} \begin{vmatrix} 0 & \mathrm{X} \\ \check{\mathrm{X}} & 0 \end{vmatrix} \begin{vmatrix} \overset{-1}{g} & 0 \\ 0 & \overset{-1}{g}{}^{\Pi} \end{vmatrix} = \begin{vmatrix} 0 & g\mathrm{X}{}^t\overline{g} \\ (g\mathrm{X}{}^t\overline{g})^{v} & 0 \end{vmatrix}.$$

Elle se réduit à l'action trouvée ci-dessus : $\mathrm{X} \to g\mathrm{X}{}^t\overline{g}$.

Nous allons étudier les relations qui existent entre les groupes de Clifford, spinoriel, et orthogonal d'un espace quadratique régulier. Observons d'abord que l'automorphisme principal Π de $\mathrm{C}(\mathrm{E};q)$ détermine un automorphisme involutif de $\mathrm{G}(\mathrm{E};q)$ dont $\mathrm{G}_+(\mathrm{E};q)$ est le sous-groupe des éléments invariants.

La symétrie orthogonale σ_a de $(\mathrm{E}; q)$ qui change le vecteur régulier a en $-a$ s'écrit, en reprenant la notation $(x \mid y)$ pour le produit scalaire $g(x, y)$ associé à q :

$$\forall x \in \mathrm{E} : \sigma_a(x) = x - 2\, \frac{(x \mid a)}{q(a)}\, a.$$

Or, on peut écrire cette équation à l'aide de la multiplication de $\mathrm{C}(\mathrm{E}; q)$, puisque $2(x \mid a) = ax + xa$ et que $\dfrac{a}{q(a)} = a^{-1}$,ce qui donne :

$$\sigma_a(x) = x - (ax + xa)\, \overset{-1}{a} = -ax\, \overset{-1}{a}.$$

Ceci encourage à essayer de représenter plus généralement un automorphisme quadratique de E à l'aide de la multiplication dans $\mathrm{C}(\mathrm{E}; q)$. Un tel automorphisme u peut toujours être écrit comme produit de symétries orthogonales : $u = \sigma_{a_1}\sigma_{a_2}\ldots\sigma_{a_k}$ et peut donc être représenté dans $\mathrm{C}(\mathrm{E}; q)$ par l'application :

$$x \longrightarrow (-1)^k (a_1 a_2 \ldots a_k) \times (a_1 a_2 \ldots a_k)^{-1}.$$

On voit apparaître les éléments du groupe de Clifford $\mathrm{G}(\mathrm{E}; q)$, produits de vecteurs réguliers de $\mathrm{C}(\mathrm{E}; q)$, l'automorphisme Π de ce groupe, et un homomorphisme naturel Ψ de $\mathrm{G}(\mathrm{E}; q)$ *sur* le groupe orthogonal $\mathrm{O}(\mathrm{E}; q)$: $g \in \mathrm{G}(\mathrm{E}; q) \to \Psi(g) \in \mathrm{O}(\mathrm{E}; q)$ avec $\Psi(g)x = (\Pi g)x\, \overset{-1}{g}$, $\forall x \in \mathrm{E}$.

On peut vérifier *a priori* que $\Psi(g) \in \mathrm{O}(\mathrm{E}; q)$ en calculant :

$$\begin{aligned} q(\Psi(g)x)1 &= (\Psi(g)x)^2 = -(\Psi(g)x)(\Pi\Psi(g)x) \\ &= -(\Pi g)x\, \overset{-1}{g}\, g(-x)\, \overset{-1}{\Pi g} = x^2 = q(x)\cdot 1 \end{aligned}$$

De plus, Ψ applique $\mathrm{G}_+(\mathrm{E}; q)$ sur $\mathrm{SO}(\mathrm{E}; q)$, puisque les éléments de ce dernier groupe sont des produits d'un nombre pair de symétries orthogonales, et $\mathrm{G}_-(\mathrm{E}; q)$ sur $\mathrm{O}_-(\mathrm{E}; q)$.

On vient de faire opérer le groupe de Clifford $\mathrm{G}(\mathrm{E}; q)$ comme groupe d'automorphismes quadratiques de E. Mais on a vu que tout automorphisme quadratique de E s'étend naturellement en un automorphisme de l'algèbre $\mathrm{C}(\mathrm{E}; q)$. $\mathrm{G}(\mathrm{E}; q)$ opère donc comme groupe d'automorphismes de l'algèbre $\mathrm{C}(\mathrm{E}; q)$ par prolongement de son action sur E. Si $g \in \mathrm{G}_+(\mathrm{E}; q)$, ce prolongement est évidemment l'automorphisme intérieur : $\alpha(g) \cdot t = gt \cdot \overset{-1}{g}$.

Si $g \in \mathrm{G}_-(\mathrm{E}; q)$ et $x \in \mathrm{E}$, $(\Pi g)x\overset{-1}{g} = -gx\overset{-1}{g} = \Pi(gx\overset{-1}{g}) = g(\Pi x)\overset{-1}{g}$; C'est donc l'automorphisme composé $\Pi \circ \alpha(g) = \alpha(g) \circ \Pi$ de $\mathrm{C}(\mathrm{E}; q)$ qui prolonge l'action de g sur E.

On démontre sans peine la théorème suivant dans le cas des corps réel ou complexe (cf. DR) :

Théorème VI.3.A :

a) Soit $(\mathrm{E}; q)$ un espace quadratique régulier *complexe*, ou *réel euclidien* de dimension $n \geqslant 2$.
Le groupe spinoriel de E est alors le groupe que forment dans $\mathrm{C}(\mathrm{E}; q)$ les produits d'un nombre pair de vecteur unitaires de E. Il est connexe et simplement connexe par arcs et c'est un revêtement à deux feuillets du groupe des rotations $\mathrm{SO}(\mathrm{E}; q)$.

b) Soit $(\mathrm{E}; q) = \mathrm{E}_{p,m}$ un espace *pseudoeuclidien* d'indice (p, m), p et m non nuls. Le groupe spinoriel $\mathrm{Spin}(p, m)$ de E est le groupe que forment dans $\mathrm{C}(\mathrm{E}_{p,m})$ les produits d'un nombre pair de vecteurs a_j tels que $q(a_j) = +1$ et d'un nombre pair de vecteur b_k tels que $q(b_k) = -1$. Il est connexe et simplement connexe par arcs si $(p, m) \neq (1, 1)$. C'est un revêtement à deux feuillets du groupe des *rotations propres* $\mathrm{SO}_+(p, m)$ (conservant l'orientation complète de $\mathrm{E}_{p,m}$).

Par contre, le groupe spinoriel $\mathrm{Spin}(1, 1)$ du plan hyperbolique a, lui, deux composantes connexes. Si (e_1, e_2) est une base hyperbolique : $q(e_1) = +1$, $q(e_2) = -1$ les éléments de $\mathrm{Spin}(1, 1)$ dans $\mathrm{C}(\mathrm{E}_{1,1})$ peuvent s'écrire : $\pm(\operatorname{ch}\theta \cdot 1 + \operatorname{sh}\theta \cdot e_1e_2)$. Ils décrivent dans le plan $(1, e_1e_2)$ les deux branches d'une hyperbole.

Sous l'action du groupe orthogonal $\mathrm{O}(\mathrm{E}; q)$, l'espace vectoriel $\mathrm{C}(\mathrm{E}; q)$ se décompose naturellement en une somme directe de sous-modules que nous allons déterminer.

Théorème VI.3.B : Soit $(\mathrm{E}; q)$ un espace quadratique régulier de dimension n. Pour chaque $p = 0, 1, 2, \ldots, n$ on a un isomorphisme naturel Φ_p de $\mathrm{O}(\mathrm{E}; q)$-modules de $\wedge^p\mathrm{E}$ sur un sous-espace C_p de l'algèbre de Clifford $\mathrm{C} = \mathrm{C}(\mathrm{E}; q)$.

C est somme directe des sous-espaces C_p et $\oplus\Phi_p$ est un isomorphisme de $\mathrm{O}(\mathrm{E}; q)$-modules de $\wedge\mathrm{E}$ sur C. Pour toute base orthogonale $e = (e_1, e_2, \ldots, e_n)$, $\Phi_p(e_{i_1} \wedge e_{i_2} \wedge \ldots e_{i_p}) = e_{i_1} e_{i_2} \ldots e_{i_p} = e_{\mathrm{I}}$ pour $i_1 < i_2 < \ldots i_p$, et les e_{I} ; $|\mathrm{I}| = p$ forment une base de C_p. En particulier quelle que soit la base orthogonale $e' = (e'_1, e'_2, \ldots, e'_n)$, le produit $e'_1 e'_2 \ldots e'_n$ est un multiple scalaire de $e_1 e_2 \ldots e_n$. La décomposition naturelle de C en $\oplus \mathrm{C}_p$ est une graduation de l'espace vectoriel de C dans $\mathbb{Z}$.

Preuve : On a d'abord $\wedge^\circ \mathrm{E} = \mathrm{K} \cdot 1_{\wedge \mathrm{E}}$, $\mathrm{C}_0 = \mathrm{K} \cdot (1_c)$ et $\Phi_0(1_{\wedge \mathrm{E}}) = 1_c$. Si $p \geqslant 1$, considérons l'application de $\mathrm{E} \times \mathrm{E} \times \mathrm{E} \ldots \mathrm{E}$ (p fois) dans C définie par :

$$\boxed{(x_1, x_2, \ldots, x_p) \to \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) x_{\sigma 1} x_{\sigma 2} \ldots x_{\sigma p}}.$$

Elle est linéaire alternée et détermine une application linéaire Φ_p de $\wedge^p \mathrm{E}$ dans C. Si $e_{i_1}, e_{i_2}, \ldots, e_{i_p}$ sont p vecteurs d'une base orthogonale de E dont les indices forment une suite strictement croissante, ils anticommutent dans C, et $e_{i_{\sigma 1}} \cdot e_{i_{\sigma 2}} \ldots e_{i_{\sigma p}} = \varepsilon(\sigma) e_{i_1} \cdot e_{i_2} \ldots e_{i_p}$. On a donc : $\Phi_p(e_{i_1} \wedge e_{i_2} \wedge \ldots e_{i_p}) = e_{i_1} \cdot e_{i_2} \ldots e_{i_p}$, d'où il résulte que Φ_p est injectif et que $\oplus \Phi_p$ est un isomorphisme de $\wedge \mathrm{E}$ sur C. Si $u \in \mathrm{O}(\mathrm{E}; q)$, $(\wedge^p u) \cdot x_1 \wedge x_2 \wedge \ldots x_p = u x_1 \wedge u x_2 \wedge \ldots u x_p$ tandis que, dans C, $\mathrm{U}(x_1 x_2 \ldots x_p) = (u x_1)(u x_2) \ldots (u x_p)$. Φ_p est donc un isomorphisme de $\mathrm{O}(\mathrm{E}; q)$-modules.

Si e' est une autre base orthogonale : $e' = e\mathrm{S}$, on a $e'_1 e'_2 \ldots e'_n = (\text{dét S}) \cdot e_1 e_2 \ldots e_n$ en vertu de la multilinéarité.

Notation : Par un abus d'écriture commode, on peut noter $x_1 \wedge x_2 \wedge \ldots x_p$ l'élément de $\mathrm{C}(\mathrm{E}; q)$ image par Φ_p de l'élément correspondant de $\wedge^p \mathrm{E}$. On matérialise ainsi l'isomorphisme canonique de $\mathrm{O}(\mathrm{E}; q)$-modules entre $\wedge \mathrm{E}$ et $\mathrm{C}(\mathrm{E}; q)$ lorsque $(\mathrm{E}; q)$ est régulier. On peut ainsi écrire *dans* $\mathrm{C}(\mathrm{E}; q)$ avec cette notation : $xy = \frac{1}{2}(xy + yx) + \frac{1}{2}(xy - yx)$ soit : $xy = (x \mid y) \cdot 1 + x \wedge y$.

Proposition VI.3.A :

Le noyau de l'homomorphisme Ψ du groupe de Clifford $\mathrm{G}(\mathrm{E}; q)$ d'un espace quadratique régulier sur $\mathrm{O}(\mathrm{E}; q)$ est égal à $\mathrm{K}^* \cdot 1_c$, sous-groupe que forment les scalaires non nuls, qui est contenu dans $G_+(E;q)$. C'est donc aussi le noyau de l'homomorphisme de $G_+(E;q)$ sur $SO(E;q)$. Le noyau de l'homomorphisme du groupe spinoriel $\mathrm{Spin}(E;q)$ dans $SO(E;q)$ est le sous-groupe à deux éléments : $(-1,+1)$ de K^*, qui sont les scalaires de norme 1.

Preuve : Si $g \in G_+(E;q)$, $\Psi(g)$ est l'identité de E si et seulement si $gx\overset{-1}{g} = x$, soit $gx = xg$, pour tout x de E. Il en résulte que g commute avec tous les éléments de $C(E;q)$: g est donc un élément pair *du centre* Z de l'algèbre. Si $g \in G_-(E;q)$ c'est $gx = -xg$ que l'on doit avoir pour tout x de E : g doit anticommuter avec tous les éléments impairs : g est donc un élément impair de ce qui est appelé *l'anticentre* A de $C(E;q)$, donc nul et la proposition est une conséquence de la suivante :

Proposition VI.3.B :

Soient $(E;q)$ un espace quadratique régulier de dimension n, $e = (e_1, e_2, \ldots, e_n)$ une base orthogonale de E, $J = e_1 \cdot e_2 \ldots e_n$ le produit des vecteurs de e dans $C(E;q)$, Z le centre, A l'anticentre de $C(E;q)$.

1) Si n est pair, Z et A sont pairs et sont de dimension un : $Z = K \cdot 1_c$, $A = K \cdot J$

2) Si n est impair, $A = 0$, $Z = \{\lambda 1_c + \mu J; \lambda, \mu \in K\}$ et $\Pi(\lambda 1_c + \mu J) = \lambda 1_c - \mu J$

Preuve : Soit z un élément du centre. Pour chaque i, il doit commuter avec e_i dont être invariant par l'automorphisme intérieur $t \to e_i t \overset{-1}{e}_i = e_i t \dfrac{e_i}{q(e_i)}$. Ecrivons z dans la base e comme somme de quatre termes, où P, P′ sont pairs, Q, Q′ impair :

$$z = \sum_{P, i \in P} \lambda_P e_P + \sum_{P', i \notin P'} \lambda_{P'} e_{P'} + \sum_{Q, i \in Q} \lambda_Q e_Q + \sum_{Q', i \notin Q'} \lambda_{Q'} e_{Q'}.$$

Puisque $e_i e_P \overset{-1}{e}_i = -e_P$, $e_i e_Q \overset{-1}{e}_i = -e_{Q'}$, et que l'on doit avoir $e_i z \overset{-1}{\underset{i}{e}} = z$ pour chaque i, on en déduit que :

a) si n est pair, $z = \lambda \cdot 1_c$.

b) si n est impair, $z = \lambda \cdot 1_c + \mu J$.

Un raisonnement analogue permet de calculer l'anticentre.

Nous allons maintenant nous intéresser, *lorsque le corps est réel* à l'algèbre de Lie commune au groupe orthogonal $\mathrm{O(E};q)$ et au groupe spinoriel $\mathrm{Spin(E};q)$ d'un espace régulier $(\mathrm{E};q)$. Un arc différentiable issu de l'élément neutre 1 dans le groupe de Clifford pair $\mathrm{G}_+(\mathrm{E};q) : t \to g(t)$ avec $t \in [0,\alpha]$, $g(0) = 1$ vérifie quel que soit $x \in \mathrm{E}$:

$$g(t)xg(\overset{-1}{t}) = y(t) \in \mathrm{E}.$$

En dérivant par rapport à t, pour $t = 0$, cette égalité ainsi que l'égalité $g(t)g(\overset{-1}{t}) = 1$, on obtient :

$$g'(0) + \overset{-1}{g}{}'(0) = 0 \text{ et } g'(0)x - xg'(0) = y'(0) \in \mathrm{E}.$$

Un vecteur tangent $g'(0)$ en l'élément neutre de G_+ est donc un élément a de la sous-algèbre C_+ qui satisfait à la condition : $\forall x \in \mathrm{E}$, $[a,x] = ax - xa \in \mathrm{E}$. Cette condition s'analyse aisément en prenant une base orthogonale $(e_1, e_2, \dots, e_n)$ de E et en écrivant que $[a, e_j] \in \mathrm{E}$, $\forall j$, par un raisonnement analogue à la démonstration de la Proposition VI.3.B. On voit alors que a est nécessairement de la forme $a = \lambda \cdot 1 + \sum_{i<j} \lambda_{ij} e_i e_j$. Si l'arc différentiable $g(t)$ est contenu dans le groupe $\mathrm{Spin(E};q)$, la norme de $g(t)$ est constamment égale à un : $\breve{g}(t)g(t) = 1$, d'où : $\breve{g}'(0) + g(0) = 0$. Un vecteur tangent en l'élément neutre de $\mathrm{Spin(E};q)$ satisfait donc en outre à l'équation $\breve{a} + a = 0$, équivalente à $\lambda = 0$.

Nous avons ainsi démontré que l'algèbre de Lie de $\mathrm{Spin(E};q)$ est *contenue* dans le sous-espace $\mathrm{C}_2(\mathrm{E};q)$ que forment les combinaisons linéaires $\sum_{i<j} \lambda_{ij} e_i e_j$ pour une base orthogonale e arbitraire et prouvé la moitié de la

Proposition VI.3.C :

L'algèbre de Lie du groupe spinoriel $\mathrm{Spin(E};q)$ s'identifie naturellement au sous-espace $\mathrm{C}_2(\mathrm{E};q)$ de l'algèbre de Clifford $\mathrm{C(E};q)$.

Preuve : Vérifions d'abord que $\mathrm{C}_2(\mathrm{E};q)$ est une sous-algèbre de Lie de l'algèbre de Lie associée à $\mathrm{C(E};q)$.

En effet, le crochet $[e_ie_j, e_ke_l]$ est nul si les quatre indices sont différents ou si $i = k$ et $j = l$. Et lorqu'il y a que trois indices distincts :

$$[e_ie_j, e_ie_k] = -2q(e_i)e_je_k \text{ et } [e_ie_j, e_je_k] = 2q(e_j)e_ie_k.$$

Soit maintenant P un plan régulier de $(\mathrm{E}; q)$. Son orthogonal $\mathrm{P}^\perp$ en est un supplémentaire.

Les automorphismes quadratiques du plan P prolongés par l'identité sur $\mathrm{P}^\perp$ identifient le groupe orthogonal $\mathrm{O(P)}$ à un sous-groupe de $\mathrm{O}(\mathrm{E}; q)$. Il en est ainsi en particulier des rotations de P qui se prolongent en des rotations de $(\mathrm{E}; q)$ appelées *rotations élémentaires « d'axe* $\mathrm{P}^\perp$ ». Une telle rotation est le produit de deux symétries orthogonales $\sigma_b\sigma_a$ relatives à deux vecteurs réguliers a, b de P.

Son image réciproque dans le groupe de Clifford pair $\mathrm{G}_+(\mathrm{E}; q)$ est l'ensemble des éléments $\{\lambda ba\}$, λ scalaire non nul, qui définissent le même automorphisme intérieur de $\mathrm{C}(\mathrm{E}; q)$: $x \longrightarrow \lambda bax(\lambda \overset{-1}{b} a)$.

La rotation $\sigma_b\sigma_a$ se relève dans $\mathrm{Spin}(\mathrm{E}; q)$ si a et b sont « de même signe » lorsque $\mathrm{K} = \mathbb{R}$. On peut alors normaliser a et b en prenant $q(a) = q(b) = +1$ ou $q(a) = q(b) = -1$ et la rotation $\sigma_b\sigma_a$ a pour relèvements dans $\mathrm{Spin}\, E$ les deux éléments ba et $(-ba)$.

Rapportons le plan P à une base orthonormale (e_1, e_2) et considérons une rotation de P d'« angle » θ. Lorsque P est euclidien ou antieuclidien on peut la décrire en prenant $a = e_1$, $b = e_1 \cos\frac{\theta}{2} + \mathrm{e}_2 \sin\frac{\theta}{2}$. Ses relèvements dans Spin(E) sont : $\pm ba = \pm\left(\cos\frac{\theta}{2}\, 1 + e_2e_1 \sin\frac{\theta}{2}\right)$, si P est euclidien, et $\pm\left(\cos\frac{\theta}{2}\, 1 - e_2e_1 \sin\frac{\theta}{2}\right)$ si P est antieuclidien. Si P est hyperbolique, avec $q(e_1) = 1$, $q(e_2) = -1$, on prend $a = e_1$, $b = e_1 \operatorname{ch}\frac{\theta}{2} + e_2 \operatorname{sh}\frac{\theta}{2}$. Les relèvements de $\sigma_b\sigma_a$ dans Spin(E) sont $\pm\left(\operatorname{ch}\frac{\theta}{2} \cdot 1 + e_2e_1 \operatorname{sh}\frac{\theta}{2}\right)$.

En faisant varier θ et en prenant les signes $(+)$ on définit dans chaque cas un chemin différentiable $g(\theta)$ issu de l'élément neutre $(g(0) = 1)$ dans Spin E dont le vecteur tangent à l'origine $g'(0)$ est $\frac{1}{2}\, e_2e_1$ dans les cas euclidien et hyperbolique, $\frac{1}{2}\, e_1e_2$ dans le cas antieuclidien. On voit ainsi que *chaque produit* $e_i e_j$, $i \neq j$ *représente une rotation infinitésimale attachée au plan* (e_i, e_j) *et un élément de l'algèbre de Lie de* Spin(E) c.q.f.d.

Exemple : Dans la représentation de l'algèbre de Clifford $\mathrm{C}(\mathrm{E}_3)$ par les matrices de $\mathbb{C}(2)$, vue au §VI.1., les rotations infinitésimales autour des axes $0x, 0y, 0z$ s'écrivent à l'aide des matrices de Pauli :

$$\begin{aligned} 0x :\ \rho_1 &= -\frac{1}{2}\,\sigma_2\sigma_3 = -\frac{i}{2}\,\sigma_1; \\ 0y :\ \rho_2 &= -\frac{1}{2}\,\sigma_3\sigma_1 = -\frac{i}{2}\,\sigma_2; \\ 0z :\ \rho_3 &= -\frac{1}{2}\,\sigma_1\sigma_2 = -\frac{i}{2}\,\sigma_3; \end{aligned}$$

ρ_1, ρ_2, ρ_3 forment une base de l'algèbre de Lie $\mathfrak{O}(3)$, commune aux groupes SO(3) et Spin(3) = SU(2). Leurs crochets ou relations de commutation sont :

$$[\rho_1, \rho_2] = \rho_3 \quad [\rho_2, \rho_3] = \rho_1 \quad [\rho_3, \rho_1] = \rho_2.$$

Plus généralement, si $\overrightarrow{l}$ est un vecteur unitaire de E_3 représenté dans $\mathbb{C}(2)$ par la matrice L, la rotation infinitésimale d'axe $\overrightarrow{l}$ est représentée par $-\frac{i}{2}$ L (cf. DR).

Ces opérateurs de rotation infinitésimale sont antihermitiens et leurs valeurs propres sont par conséquent imaginaires pures. On préfère, en mécanique quantique, considérer les *opérateurs cinétiques* correspondants, qui sont les opérateurs de rotation infinitésimale multipliés par i, et qui sont hermitiens. Cela revient à complexifier l'algèbre de Lie SO(3). Si $\mathrm{J}_k = i\rho_k$, $k = 1, 2, 3$, les relations de commutation deviennent alors :

$$[\mathrm{J}_1, \mathrm{J}_2] = i\mathrm{J}_3 \quad [\mathrm{J}_2, \mathrm{J}_3] = i\mathrm{J}_1 \quad [\mathrm{J}_3, \mathrm{J}_1] = i\mathrm{J}_2.$$

En privilégiant l'axe $0z$, on pose $\mathrm{J}_+ = \mathrm{J}_1 + i\mathrm{J}_2$ et $\mathrm{J}_- = \mathrm{J}_1 - i\mathrm{J}_2$ ce qui donne les relations de commutation.

$$[\mathrm{J}_3, \mathrm{J}_+] = \mathrm{J}_+ \quad [\mathrm{J}_3, \mathrm{J}_-] = -\mathrm{J}_- \quad [\mathrm{J}_+, \mathrm{J}_-] = 2\mathrm{J}_3.$$

fort commodes pour la détermination des représentations complexes irréductibles de l'algèbre de Lie O(3) qui sera vue au §VI.5.

VI.4 – Nature des algèbres de Clifford. Spineurs

L'exposé qui vient d'être fait des propriétés des algèbres de Clifford n'est pas vraiment de nature à convaincre de leur utilité. Du point de vue linéaire, par exemple, et des représentations du groupe orthogonal, elles sont isomorphes aux algèbres extérieures. Par contre, lorsque l'espace est régulier, nous avons vu que leur structure multiplicative fait apparaître des extensions du groupe orthogonal. C'est l'approfondissement de ce phénomène qui est essentiel pour les applications.

Une façon d'étudier un ensemble A muni d'une certaine structure consiste à tenter de le représenter comme un ensemble d'opérateurs linéaires sur un espace vectoriel M, qui prend alors le nom de A-*module*, en utilisant toutes les ressources de $\mathcal{L}(\mathrm{M})$. Ces dernières permettent ainsi de représenter : les groupes, les algèbres associatives de Lie et de Jordan (cf. III.11).

Définition VI.4.A :

Une *représentation linéaire* d'une algèbre A associative unitaire (resp. de Lie, resp. de Jordan) dans l'espace vectoriel de dimension finie M est un homomorphisme ρ de A dans l'algèbre associative unitaire $\mathcal{L}(\mathrm{M})$ (resp. de Lie $\mathcal{L}(\mathrm{M})_{-}$, resp. de Jordan $\mathcal{L}(\mathrm{M})_{+}$ avec les notations du §III.11). Muni d'une telle représentation, M prend le nom de A-module. Les sous-espaces de M stables, par A sont appelés des *sous-modules*. On fait usuellement opérer une algèbre associative. A sur M *à gauche*, soit : $\rho(ab) \cdot m = \rho(a)(\rho(b) \cdot m)$. Si une partie G des éléments inversibles de A forme un groupe pour la multiplication, la restriction de ρ à G est évidemment une représentation linéaire de G dans M. D'autre ρ est également une représentation linéaire d'algèbres de Lie, de A_{-} dans $\mathcal{L}(\mathrm{M})_{-}$ ou de Jordan, de A_{+} dans $\mathcal{L}(\mathrm{M})_{+}$ pour les structures multiplicatives affaiblies de A et $\mathcal{L}(\mathrm{M})$ (cf. §III.11).

Remarque : L'existence de valeurs propres pour tout opérateur d'un espace vectoriel M sur un corps algébriquement clos impose, pour des raisons de clarté, et de simplicité, l'usage systématique de représentations linéaires dans ces espaces. Même si A est une algèbre réelle, on considère de façon privilégiée ses représentations linéaires complexes, ce qui revient à étudier $\mathrm{A} \otimes \mathbb{C}$ au lieu de A ; A est *une* forme réelle de $\mathrm{A} \otimes \mathbb{C}$ mais il y en a généralement d'autres.

Cela prouve que, l'on perd une certaine information par la seule considération de $\mathrm{A} \otimes \mathbb{C}$!

Nous allons étudier les représentations linéaires des algèbres de Clifford des espaces quadratiques *réguliers* en commençant par le

Théorème VI.4.A : Soient $(\mathrm{E}; q)$ un espace quadratique régulier, $\mathrm{C} = \mathrm{C}(\mathrm{E}; q)$ son algèbre de Clifford, le corps des scalaires K étant de caractéristique nulle.

1) Si la dimension n de E est paire, C est *simple*, c'est-à-dire qu'elle n'admet aucun idéal bilatère propre, et puisque dans ce cas, son centre est formé des multiples scalaires de l'unité, elle est dite *centrale simple*.

2) Si la dimension n de E est impaire, deux cas sont possibles suivant la nature algébrique du centre Z de C qui est une algèbre commutative de dimension deux sur K (prop. VI.3.B) :

a) si Z est un corps, C est une Z-algèbre centrale simple. Z est un corps si et seulement si $(1)^r \overline{g}$ *n'est pas* un carré dans le corps K, où $n = 2r + 1$ et $\overline{g} = \text{dét}\,|g_{ij}|$.

b) si Z n'est pas un corps, $\mathrm{Z} = \mathrm{K}\varepsilon_1 \oplus \mathrm{K}\varepsilon_2$ où les ε_j sont des idempotents, et C est somme directe de deux sous-algèbres $\mathrm{C}_+\varepsilon_1 \oplus \mathrm{C}_+\varepsilon_2$ toutes deux isomorphes à la sous-algèbre paire C_+, donc simples, et échangées par Π.

Remarque : Si le corps K de E est algébriquement clos, seul *b)* intervient. Si K n'est pas algébriquement clos ($\mathrm{K} = \mathbb{R}$ par exemple) et si l'on est dans le cas *a)* la propriété pour C d'être simple est « instable ». Si l'on tensorise par la clôture algébrique L de K, elle disparaît et on se retrouve dans le cas *b)*.

Preuve : Soit e une base orthogonale de E. Choisissons un ordre total sur les vecteurs e_{I} de la base associée de C, en utilisant par exemple l'ordre naturel des cardinaux $|\mathrm{I}| = p$ et pour chaque p l'ordre lexicographique. Soit $\mathfrak{I}$ un idéal bilatère de C. Un élément non nul b de $\mathfrak{I}$ s'écrit $b = \mu_1 e_{\mathrm{J}_1} + \ldots + \mu_k e_{\mathrm{J}_k}$. L'élément e_{J_1} est inversible et $\mathfrak{I}$ contient $\overset{-1}{e}_{\mathrm{J}_1} b$. Il en résulte que si $\mathfrak{I}$ n'est pas nul, il contient nécessairement un élément de la forme $a = \lambda_0 \cdot 1 + \lambda_1 e_{\mathrm{I}_1} + \ldots + \lambda_m e_{\mathrm{I}_m}$ avec $\lambda_0, \ldots, \lambda_m$ non nuls et $\mathrm{I}_1 < \mathrm{I}_2 < \ldots \mathrm{I}_m$. Si I_1 est pair et $i \in \mathrm{I}_1$, ou si I_1 est impair et $i \notin \mathrm{I}_1$, on a $e_i e_{\mathrm{I}_1} \overset{-1}{e}_i = -e_{\mathrm{I}_1}$ tandis que $e_i e_J \overset{-1}{e}_i = \pm e_{\mathrm{J}}$ d'où il résulte que l'élément $\dfrac{1}{2}\,(a + e_i a \overset{-1}{e}{}_i)$ est non nul, appartient à $\mathfrak{I}$, et contient au plus $(m-1)$coefficients non nuls. On en déduit immédiatement que si $\mathfrak{I}$ est non nul :

1) si la dimension de E est paire, $\mathfrak{I}$ contient l'unité et coïncide donc avec C,

2) si la dimension de E est impaire, $\mathfrak{I}$ contient nécessairement un élément de la forme $\lambda 1 + \mu \mathrm{J}$ avec $\mathrm{J} = e_1 e_2 \dots e_n$ et $\lambda \neq 0$, c'est-à-dire un élément non nul du centre Z de C (proposition VI.2.B). Deux cas sont alors possibles :

a) tout élément non nul de Z est inversible : Z est un corps, extension quadratique de K. $\mathfrak{I}$ contenant un élément inversible contient l'unité et est égal à C, qui est donc centrale simple sur Z.

b) l'élément $\lambda \cdot 1 + \mu \mathrm{J}$ n'est pas inversible. Or, le produit $(\lambda \cdot 1 + \mu\mathrm{J})(\lambda \cdot 1 - \mu \mathrm{J}) = \lambda^2 \cdot 1 - \mu^2 \mathrm{J}^2$ est un scalaire, puisque $\mathrm{J}^2 = (e_1 e_2 \dots e_n)^2 = (-1)^{\frac{n(n-1)}{2}}\; q(e_1) q(e_2) \dots q(e_n) \cdot 1$; n étant impair, on peut l'écrire $n = 2r+1$ d'où $\dfrac{n(n-1)}{2} = r(2r+1) \equiv r \bmod 2$. D'autre part, la base étant orthogonale, le produit $q(e_1) \dots q(e_n)$ n'est autre que le déterminant de la matrice $|g_{ij}|$, que nous avons appelé $\overline{g}$. On a donc : $\mathrm{J}^2 = (-1)^r \overline{g} \cdot 1$ et $(\lambda 1 + \mu \mathrm{J})(\lambda 1 - \mu \mathrm{J}) = (\lambda^2 - (-1)^r \overline{g} \cdot \mu^2) \cdot 1$

Lorsque $\beta = \lambda^2 - (-1)^r \overline{g} \mu^2$ est non nul, $\lambda \cdot 1 + \mu \mathrm{J}$ est inversible d'inverse $\beta^{-1} \cdot (\lambda 1 - \mu \mathrm{J})$. Ainsi, Z est un corps si et seulement si $(-1)^r \overline{g}$ *n'est pas* un carré dans le corps K. Dans le cas où le centre Z n'est pas un corps, on a donc nécessairement $\mathrm{J}^2 = \alpha^2 \in \mathrm{K}$. Soient :

$$\varepsilon_1 = \frac{1}{2}\left(1 + \frac{1}{\alpha}\,\mathrm{J}\right) \text{ et } \varepsilon_2 = \frac{1}{2}\left(1 - \frac{1}{\alpha}\,\mathrm{J}\right).$$

On a $\varepsilon_1^2 = \varepsilon_1$, $\varepsilon_2^2 = \varepsilon_2$, $\varepsilon_1 \varepsilon_2 = \varepsilon_2 \varepsilon_1 = 0$, $\varepsilon_1 + \varepsilon_2 = 1$, $\mathrm{J} = \alpha(\varepsilon_1 - \varepsilon_2)$.

Autrement dit ε_1 et ε_2 sont des idempotents algébriquement orthogonaux formant une décomposition de l'unité. Puisque J est impair, ce sont des éléments *inhomogènes*, échangés par Π, car $\Pi \mathrm{J} = -\mathrm{J}$. De plus, $\mathrm{C}_- = \mathrm{J}\mathrm{C}_+$ soit $\mathrm{C} = \mathrm{C}_+ \oplus \mathrm{J}\mathrm{C}_+$. Tout élément u de C s'écrit donc d'une façon et d'une seule sous la forme : $u = a + \mathrm{J}b$ avec a et $b \in \mathrm{C}_+$, soit :

$$u = a(\varepsilon_1 + \varepsilon_2) + \alpha(\varepsilon_1 - \varepsilon_2)b = (a + b\alpha)\varepsilon_1 + (a - b\alpha)\varepsilon_2$$

donc d'une façon et d'une seule comme somme $p\varepsilon_1 + q\varepsilon_2$ avec p et $q \in C_+$. On a donc un isomorphisme : $C = C_+\varepsilon_1 \oplus C_+\varepsilon_2$, et C est isomorphe à la somme directe de deux sous-algèbres isomorphes à la sous-algèbres C_+, donc simples, puisque C_+ est, d'après le théorème VI.2.B, l'algèbre de Clifford d'un sous-espace régulier de E de codimension un, et par conséquent de dimension paire.

Exemples : Les exemples suivants illustrent les divers cas du théorème :

1) $n = 2$. Les algèbres de Clifford (exemples 3) et 5) du §VI.1) :

$$\mathrm{C}(\mathrm{K}^2; q(e_1x + e_2y) = x^2 + y^2) = \mathrm{K}(2)$$
$$\mathrm{C}(\mathbb{R}^2; q(e_1x + e_2y) = -x^2 - y^2) = \mathbb{H}$$

sont centrales simples.

2) $n = 1$ ou 3. Les algèbres de Clifford (exemples 2) 4) 6) du §VI.1) :

$$\mathrm{C}(\mathrm{K}; q(x \cdot 1) = x^2) = \mathrm{K} \oplus \mathrm{K};\;\; \mathrm{C}(\mathbb{R}; q(x \cdot 1) = -x^2) = \mathbb{C}$$
$$\mathrm{C}(\mathbb{R}^3; q(e_1x + e_2y + e_3z) = -x^2 - y^2 - z^2)) = \mathbb{H} \oplus \mathbb{H}$$
$$\mathrm{C}(\mathbb{R}^3; q(e_1x + e_2y + e_3z) = x^2 + y^2 + z^2) = \mathbb{C}(2)$$

sont centrales simples ou sommes directes de deux algèbres centrales simples.

3) Espaces quadratiques réels réguliers de dimension 4. On a vu (Exemple 7 du §VI.1 et Théorème VI.1.B) :

$$\mathrm{C}(\mathrm{E}_{1,3}) = \mathbb{H}(2); \quad \mathrm{C}(\mathrm{E}_{2,2}) = \mathbb{R}(4).$$

On prouve par un choix convenable d'applications de Clifford (cf. DR)

$$\mathrm{C}(\mathrm{E}_{4,0}) = \mathrm{C}(\mathrm{E}_{0,4}) = \mathbb{H}(2).$$

4) Soit $(\mathrm{E}; q)$ un espace quadratique régulier *réel* (c'est-à-dire euclidien ou pseudoeuclidien) de dimension n *impaire*. En choisissant une base orthonormale e de $\mathrm{E}(q(e_j) = \pm 1)$, l'élément $\mathrm{J} = e_1e_2 \ldots e_n$ a pour carré J^2 égal à plus un ou moins un. Lorsque $\mathrm{J}^2 = -1$, puisque J est impair, on a $C = C_+ \oplus \mathrm{J}C_+$ et C est isomorphe à la complexifiée de sa sous-algèbre paire C_+. Si l'on

complexifie alors C, les idempotents $\varepsilon_1 = \dfrac{1+i\mathrm{J}}{2}$ et $\varepsilon_2 = \dfrac{1-i\mathrm{J}}{2}$ décomposent $\mathrm{C} \otimes \mathbb{C}$ en la somme directe de deux sous-algèbres :

$$\mathrm{C} \otimes \mathbb{C} = (\mathrm{C}_+ \otimes \mathbb{C})\varepsilon_1 \oplus (\mathrm{C}_+ \otimes \mathbb{C})\varepsilon_2 \equiv \mathrm{C} \oplus \overline{\mathrm{C}}$$

isomorphes à la complexifiée de C_+ c'est-à-dire à C elle-même considérée avec J comme algèbre complexe.

Dans une représentation linéaire complexe de $\mathrm{C} \otimes \mathbb{C}$ dans M, homomorphisme φ d'algèbres unitaires : $\mathrm{C}\otimes\mathbb{C} \to \mathfrak{L}(\mathrm{M})$, $\varphi(\varepsilon_1) = 0$ est équivalent à $\varphi(\mathrm{J}) = i\mathbb{1}_{\mathfrak{L}(\mathrm{M})}$ et $\varphi(\varepsilon_2) = 0$, équivalent à $\varphi(\mathrm{J}) = -i\mathbb{1}_{\mathfrak{L}(\mathrm{M})}$. Les représentations linéaires complexes de $\mathrm{C} \otimes \mathbb{C}$ telles que $\varphi(\varepsilon_1) = 0$ sont donc les représentations linéaires complexes de $(\mathrm{C};\mathrm{J})$, celles pour lesquelles $\varphi(\varepsilon_2) = 0$ les représentations antilinéaires complexes de $(\mathrm{C};\mathrm{J})$.

5) Les algèbres de Clifford des espaces pseudoeuclidiens $\mathrm{E}_n = \mathrm{E}_{p,q}$ se déterminent aisément par récurrence. Elles se classent naturellement suivant les huit classes de $(p-q)$ modulo 8

1	$\mathbb{R}(n) \oplus \mathbb{R}(n)$
2 ou 0	$\mathbb{R}(n)$
3 ou 7	$\mathbb{C}(n)$
4 ou 6	$\mathbb{H}(n)$
5	$\mathbb{H}(n) \oplus \mathbb{H}(n)$

tandis que les algèbres de Clifford des espaces quadratiques complexes sont tout simplement :

$$\mathrm{C}(\mathrm{E}_{2p} \otimes \mathbb{C}) = \mathbb{C}(2^{2p}) \quad \text{et} \quad \mathbb{C}(\mathrm{E}_{2p+1} \otimes \mathbb{C}) = \mathbb{C}(2^{2p}) \oplus \mathbb{C}(2^{2p}).$$

Remarque : Il est clair que la manipulation d'algèbres de matrices à coefficients quaternioniens n'a rien de particulièrement attirant. C'est pourquoi, en géométrie et en physique, on complexifie souvent les espaces quadratiques réels pour s'épargner des désagréments. Ce procédé brutal à l'inconvénient de faire diparaître une certaine structure fine associée à $\mathbb{R}$. On peut obtenir un plongement naturel dans des algèbres de matrices complexes par le théorème suivant :

Théorème VI.4.B : Soit $E_{r,s}$ un espace euclidien ou pseudo-euclidien de dimension paire : $r + s = 2p$. Son algèbre de Clifford $C(E_{r,s})$ (qui est une algèbre complète de matrices sur $\mathbb{R}$ ou sur $\mathbb{H}$), en s'identifiant à une sous-algèbre de l'algèbre de Clifford de l'espace pseudoeuclidien de dimension $2p + 1$: $F = E_{r,s} \oplus \mathbb{R}\varepsilon$, avec $q(\varepsilon) = (-1)^{\frac{r-s}{2}} + 1$,. se réalise comme une algèbre de matrices complexes, sous-algèbre et forme réelle de l'algèbre $C(F) = \mathbb{C}(2^p)$.

Preuve : En effet si $e_1, e_2, \ldots e_{2p}$ est une base orthogonale de $E_{r,s}$ l'élément $J - e_1, e_2, \ldots, e_{2p}\varepsilon$ de $C(F)$ a pour carré :

$$J^2 = -1^{\frac{2p(2p+1)}{2}}(-1)^s q(\varepsilon).$$

Puisque $\dfrac{2p(2p+1)}{2} + s \equiv p + s = \dfrac{r+s}{2} + s \equiv \dfrac{r-s}{2} \mod 2$, on a bien $J^2 = -1$. D'après le théorème VI.4.A, $C(F)$ est une algèbre complexe simple et $C(F) = C(E_{r,s}) \oplus JC(E_{r,s})$.

Définition VI.4.B :

Un A-module non nul M est dit *simple* s'il ne contient aucun autre sous-module que lui-même et 0.

On dit aussi que la représentation de A dans M est *irréductible*. M est dit *semi-simple* s'il est somme directe de sous-modules simples. La représentation de A est dite dans ce cas *complètement réductible*. Une algèbre A est dite *simple* si elle ne contient aucun idéal bilatère propre et si sa multiplication est non nulle.

Un A-module M est dit *fidèle* si l'application de l'ensemble A qui définit sa structure dans $\mathfrak{L}(M)$ est injective.

Remarque : Si $a \in A$, $m \in M$, on note simplement $a \cdot m$ l'action de a sur m.

Théorème VI.4.C : Si un module M est engendré par un ensemble $(M_i,\ i \in I)$ de sous-modules simples, il est semi-simple et tout sous-module N est facteur direct : $M = N \oplus N'$ où N' est un sous-module supplémentaire.

Preuve : Soit $J \subset I$ une partie maximale pour la propriété : « la somme $S = N + \sum M_j$, $j \in J$, est directe », où l'on peut prendre $N = 0$.

Soit M_i un sous-module quelconque de l'ensemble. Puisque M_i est simple, l'intersection de M_i et de S est un sous-module de M_i, donc 0 ou M_i. Cela ne peut être 0 puisqu'alors la somme $M_i + S$ serait directe. C'est donc que $M_i \subset S$ qui est égal à M.

Si A est une algèbre associative avec unité, elle opère sur elle-même par translations à gauche, et les sous-modules de cette représentation sont les idéaux à gauche.

Théorème VI.4.D : Soient A une algèbre associative avec unité, J un idéal à gauche non nul minimal de A, M un A-module.

a) si $m \in M$, $J \cdot m$ est ou nul ou un sous-module simple isomorphe à J.

b) *si* M *est simple*, ou bien $JM = 0$, ou bien il existe $m \in M$ tel que $J \cdot m \neq 0$ et J et M sont des A-modules isomorphes.

c) *si* A *est simple*, tous les A-modules simples sont isomorphes : *il n'y a qu'une seule représentation linéaire irrédutible de* A à un isomorphisme près. Toute représentation non nulle de A est fidèle.

d) si A est simple, tous les A-modules sont semi-simples.

Preuve :

a) L'application φ de J dans M : $\varphi(a) = am$ est un homomorphisme de A-modules puisque quel que soit $b \in A$, $\varphi(ba) = b\varphi(a)$. Son noyau est un sous-module de J, soit un idéal à gauche de A contenu dans J : $\forall x \in \operatorname{Ker}\varphi$, $\forall b \in A$, $bx \in \operatorname{Ker}\varphi$.
Comme J est non nul minimal, c'est ou 0 ou J. Si Jm n'est pas nul, c'est donc un sous-module de M isomorphe à J et simple.

b) $AJ = J$ d'où $JM = AJM$ est un sous-module de M. Si M est simple, ou bien $JM = 0$, ou il existe m tel que $J \cdot m \neq 0$.
Dès lors, Jm, sous-module de M est égal à M, et isomorphe à J d'après a).

c) Si maintenant A est simple, JA, qui est un idéal bilatère non nul, est égal à A. Dès lors, $M = AM = JAM$. Si M est simple M est isomorphe à J d'après b). Le noyau d'une représentation linéaire ρ de A est un idéal bilatère, qui ne peut être, puisque A est simple, que A ou 0.

d) si M est un A-module, puisque $A = J \cdot A = \sum_{a \in A} J \cdot a$, M : $AM = \sum_{a \in A} J \cdot a \cdot m$.

M est donc engendré par les sous-modules $J \cdot a \cdot m$ qui sont ou nuls, ou simples.

Définition VI.4.C :

Une algèbre A (d'un type quelconque) est dite *semi-simple* si elle est somme directe d'idéaux bilatères minimaux $A = \bigoplus_{j=1}^{r} A_j$, la multiplication dans chacun d'eux étant non nulle.

Le produit de deux éléments appartenant à des idéaux différents de la décomposition devant appartenir à chacun d'eux, donc à leur intersection, est nul.

Si $a = a_1 + a_2 + \ldots + a_r$, $b = b_1 + b_2 + \ldots + b_r$ et $x \in A_j$, on a donc $axb = a_j x b_j$. Il en résulte qu'un idéal bilatère de l'algèbre A_j est un idéal de A contenu dans A_j.

Puisque A_j est minimal, c'est 0 ou A_j et A_j est donc une algèbre simple. Pour ne pas allonger inutilement cette étude, nous nous limiterons désormais aux seules algèbres associatives avec unité.

Proposition VI.4.A :

Dans une algèbre associative avec unité semi-simple : $A = \bigoplus_{j=1}^{r} A_j$ il n'y a d'autre idéal bilatère simple non nul que les A_j et la décomposition de A est donc unique à l'ordre près des facteurs. Si $1_A = e_1 + e_2 + \ldots + e_r$, e_j est l'unité de A_j.

Preuve : Soit I un idéal bilatère simple non nul de A : $A_j I A_k \subset A_j A_k = 0$ si $j \neq k$, et $I = AIA = \sum_j A_j I A_j$. L'un au moins des $A_j I A_j$ n'est pas nul et dans ce cas est un idéal bilatère contenu dans I, donc égal à A_j. Si $a_j \in A_j$, on a :

$$a_j = a_j 1_A = 1_A \cdot a_j = \sum_k e_k a_j = \sum_k a_j e_k \text{ d'où : } a_j = a_j e_j = e_j a_j.$$

Théorème VI.4.E : Soit A une algèbre associative avec unité semi-simple. *a)* Tout A-module est semi-simple. *b)* Tout idéal à gauche de A est engendré par un idempotent.

Preuve :

a) Soient $A = \overset{r}{\underset{i=1}{\oplus}} A_i$, A_i simples et M un A-module $M = AM = \overset{r}{\underset{i=1}{\oplus}} A_i M = \sum_{i=1}^{r} J_i A_i M$, où, pour chaque i, J_i est un idéal à gauche simple de A_i ; D'où : $M = \sum_{i, a_i, m}, J_i \cdot a_i \cdot m$. Or $J_i \cdot m$ est ou bien nul ou bien un sous-module simple de M d'après le théorème précédent. M engendré par des sous-modules simples est semi-simple.

b) Pour les translations à gauche, A est elle-même un A-module, donc semi-simple. Les sous-modules sont les idéaux à gauche. Si J est un idéal à gauche, il est facteur direct d'après le théorème VI.3.B et $A = J \oplus J'$, d'où $1_A = e + e'$. Si $a \in J$, $a = a \cdot 1_A = ae + ae'$.

L'élément $ae' = a - ae \in J \cap J'$ est nul. En prenant $a = e$, on obtient $e = e^2$. Puisque $J = Je$, on a donc trouvé dans J un idempotent qui l'engendre, et puisque $A = Ae \oplus Ae'$, $J = Ae$. Remarquons aussi que les idempotents e et e' sont algébriquement orthogonaux : $ee' = e'e = 0$.

Théorème VI.4.F : Soient A une algèbre associative avec unité semi-simple, somme directe de r algèbres simples : $A = A_1 \oplus A_2 \oplus \dots A_r$, et $1_A = e_1 + e_2 + \dots + e_r$ avec $e_j \in A_j$. Alors :

1) Tout A-module M est somme directe des sous-modules $e_j M$.

2) Un A-module fidèle M contient au moins un A_j-module simple M_j pour chaque $j = 1, 2, \dots, r$: $M_j \subset e_j M$.

3) A un isomorphisme près il n'y a qu'un seul A-module fidèle minimal S, qui est somme directe de r sous-modules simples $\overset{r}{\underset{i=1}{\oplus}} M_j$, chaque M_j étant un module simple de la sous-algèbre A_j.

4) Il n'y a d'autre sous-module simple de S que les M_j.

Preuve :

1) l'intersection de $e_j M$ avec la somme des $e_k M$, $k \neq j$ est nulle : il suffit de multiplier par e_j une égalité : $e_j m = \sum_{k \neq j} e_k m_k$.

2) si l'homomorphisme ρ de A dans $\mathfrak{L}(M)$ est injectif, $\rho(A_j)$ est non nul. Si les opérateurs de $\rho(A_j)$ ne sont pas tous nuls, il existe

$m \in \mathrm{M}$ tel que $\mathrm{A}_j\dot{m} \neq 0$. Si J_j est un idéal à gauche minimal de A_j, on a : $\mathrm{A}_j = \mathrm{J}_j\mathrm{A}_j$ et $\mathrm{J}_i\mathrm{A}_j m \neq 0$. Il existe donc un élément $\overline{m}$ tel que $\mathrm{J}_j \cdot \overline{m} \neq 0$ et $\mathrm{J}_j \cdot \overline{m}$ est un A-module simple contenu dans $e_j \cdot \mathrm{M}$.

3) Soit $\mathrm{S} = \bigoplus_{i=1}^{r} \mathrm{J}_j$. C'est un A-module à gauche et si $a = a_1 + a_2 + \ldots + a_r \in \mathrm{A}$, $a \cdot \mathrm{S} = \sum a_j \cdot \mathrm{J}_j$. L'opérateur $\rho(a)$ est nul si et seulement si tous les $\rho(a_j)$ sont nuls. Mais la représentation de A_j dans J_j est fidèle et cela impose $a_j = 0$, $\forall j$.

S est donc un module fidèle, minimal d'après 1). Réciproquement, si M est un module fidèle minimal $e_j\mathrm{M}$ est nécessairement simple et isomorphe à J_j.

4) Si P est un sous-module simple non nul de M, on a $\mathrm{P} = 1_{\mathrm{A}} \cdot \mathrm{P} = \sum e_j\mathrm{P}$.

L'un au moins des $e_j\mathrm{P}$ est non nul. On a $\mathrm{A}e_j\mathrm{P} = \mathrm{A}_j e_j\mathrm{P}$. C'est un A-module non nul contenu dans les modules simples $e_j\,\mathrm{M} = \mathrm{M}_j$ et P et qui leur est donc égal.

Le théorème précédent entraîne que *l'algèbre de Clifford d'un espace quadratique régulier* $\mathrm{C}(\mathrm{E}; q)$ *possède un module fidèle minimal unique* $\mathrm{S}(\mathrm{E}; q)$ qui est simple ou somme directe de deux modules simples suivant que $\mathrm{C}(\mathrm{E}; q)$ est simple ou non. Le groupe spinoriel $\mathrm{Spin}(\mathrm{E}; q) \subset \mathrm{C}(\mathrm{E}; q)$ opère donc dans l'espace vectoriel $\mathrm{S}(\mathrm{E}; q)$ par une représentation linéaire.

Cependant, $\mathrm{Spin}(\mathrm{E}; q)$ est contenu dans la sous-algèbre paire $\mathrm{C}_+(\mathrm{E}; q)$. Si l'on s'intéresse essentiellement aux représentations linéaires de ce groupe, il est naturel de considérer plutôt le module fidèle minimal $\mathrm{S}_+(\mathrm{E}; q)$ de l'algèbre $\mathrm{C}_+(\mathrm{E}; q)$. Il n'en est naturellement plus de même si l'on s'intéresse au groupe de revêtement $\mathrm{RO}(\mathrm{E}; q)$ du groupe orthogonal tout entier !

En ce qui concerne $\mathrm{C}_+(\mathrm{E}; q)$, on peut l'investir d'une structure d'algèbre de Clifford (cf. théorème VI.2.B), mais d'un espace quadratique de dimension inférieure d'une unité à celle de E, donc de parité différente. Il en résulte que si $\dim \mathrm{E}$ est impaire, $\mathrm{C}_+(\mathrm{E}; q)$ est une algèbre simple et $\mathrm{S}_+(\mathrm{E}; q)$ est simple. Si $\dim \mathrm{E}$ est paire, $\mathrm{S}_+(\mathrm{E}; q)$ est simple ou somme directe de deux modules simples. Les points de vue différents qui viennent d'être exposés montrent qu'il existe un certain flottement dans ce que l'on appelle spineurs ou semi-spineurs ! Nous adopterons la définition générale suivante :

Définition VI.4.D :

On appelle espace de *spineurs* tout module fidèle minimal de l'algèbre de Clifford d'un espace quadratique régulier. Lorsque ce module est somme directe de deux modules simples, ces derniers sont appelés espaces de semi-spineurs.

Exemples : 6) $C(E_3)$ est l'algèbre *réelle* $\mathbb{C}(2)$ qui est simple. Elle a donc un seul espace de spineurs : $\mathbb{C}^2$. Mais sa compléxifiée $C(E_3)\otimes \mathbb{C}$ (Exemple 4) du §VI.4.) est somme directe de deux sous-algèbres complexes isomorphes à $\mathbb{C}(2)$, échangées par l'automorphisme Π, et possède donc un espace de spineurs $\mathbb{C}^4$ décomposé en somme directe de deux sous-espaces de semi-spineurs. L'algèbres $C(E_3)$ considérée comme complexe (avec J) opère dans l'un par une représentation linéaire complexe, dans l'autre par la représentation antilinéaire complexe conjuguée.

7) On a vu à l'Exemple du §VI.3 que l'algèbre de Clifford de l'espace de Minkowski $C(E_{1,3})$ est l'algèbre de matrices 2×2 sur $\mathbb{H}$ dont la complexifiée est l'algèbre des matrices complexes 4×4 : $C(E_{1,3}) \otimes \mathbb{C} = \mathbb{H}(2) \otimes \mathbb{C} = \mathbb{C}(4)$. Ces algèbres sont simples et l'espace des spineurs complexes $S(E_{1,3})$ est l'espace vectoriel complexe $\mathbb{C}^4$. La sous-algèbre $C_+(E_{1,3}) \equiv C(E_3) \equiv \mathbb{C}(2)$ est, dans la représentation matricielle classique de $C(E_{1,3})$ étudiée à l'exemple 2 du §VI.3. formée des matrices :

$$\begin{vmatrix} a & 0 \\ 0 & \tilde{a} \end{vmatrix} \in \mathbb{C}(4) \quad \text{avec} \quad a \in \mathbb{C}(2)$$

où $a \rightarrow \tilde{a} = ({}^t\overline{a})^\nu$ est l'automorphisme Π de $C(E_{1,3})$ déterminé par la symétrie d'espace de $E_{1,3}$.

(Rappelons que $\begin{vmatrix} a & b \\ c & d \end{vmatrix}^\nu = \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}$)

$S(E_{1,3})$ se décompose donc *relativement à* $C_+(E_{1,3})$ e, la somme directe de deux sous-espaces de semi-spineurs $S = \mathfrak{S} \oplus \overline{\mathfrak{S}}$. Le groupe spinoriel $\mathrm{Spin}(1,3) \equiv \mathrm{Sl}(2;\mathbb{C}) \subset C_+(E_{1,3})$ opère dans l'un par la représentation linéaire complexe standard de $\mathrm{Sl}(2;\mathbb{C})$ dans $\mathbb{C}^2$, dans l'autre par la représentation antilinéaire complexe conjuguée (qui est une représentation linéaire *réelle* inéquivalente à la première).

VI.5 – Un exemple de l'emploi des spineurs : le moment angulaire en mécanique quantique

Le postulat de la quantification du moment angulaire, énoncé par Niels Bohr en 1913, et de sa direction, suggéré par Sommefeld en 1916, ainsi que les règles empiriques du couplage des moments angulaires inspirées par la spectroscopie (Landé, 1923) ont trouvé une explication très naturelle dans le cadre de la mécanique quantique et des représentations de l'algèbre de Lie du groupe des rotations.

Cette explication permet de rendre compte des spins demi-entiers des particules, qui n'ont aucun droit à l'existence en mécanique classique.

Rappelons (Chap. 1) que la mécanique quantique représente les systèmes physiques par des fonctions et les grandeurs physiques par des opérateurs agissant sur elles. Le résultat d'une mesure est une valeur propre de l'opérateur. Elle représente un « état » du système. En particulier l'état dynamique relativement à un point est déterminé par les opérateurs de moment angulaire en ce point. Ces derniers forment l'algèbre de Lie $\mathfrak{O}(3)$ du groupe des rotations. L'espace vectoriel des états du système est ainsi l'espace d'une représentation linéaire de $\mathfrak{O}(3)$.

La quantification du moment angulaire est alors tout simplement conséquence du fait que le groupe SO(3) est compact, d'où il résulte que ses représentations irréductibles sont de dimension finie : un opérateur de moment angulaire a des valeurs propres discrètes !

On peut passer des moments classiques aux moments quantiques par un raisonnement élémentaire. La position relativement à la coordonnée x^j étant représentée par l'opérateur de multiplication par x^j, le moment linéaire p^j relativement à x^j, par l'opérateur $-i\hbar\,\dfrac{\partial}{\partial x^j}$ ($\hbar = \dfrac{h}{2\Pi}$, h constante de Planck), leurs relations de commutation s'écrivent :

$$[x^j, p^k] \cdot f = x^j p^k f - p^k x^j f = i\hbar\delta^{jk} \cdot f, \text{ soit } [x^j, p^k] = i\hbar\delta^{jk}.$$

Elles déterminent les relations de commutation des opérateurs de moment angulaire, rassemblés en le produit vectoriel de l'opérateur de position r et de l'opérateur de moment linéaire p soit :

$\mathrm{L} = r \times p$. Dans la représentation précédente :

$$\mathrm{L}_x = -i\hbar\Big(y\,\frac{\partial}{\partial z} - z\,\frac{\partial}{\partial y}\Big)$$
$$\mathrm{L}_y = -i\hbar\Big(z\,\frac{\partial}{\partial x} - \frac{\partial}{\partial z}\Big)$$
$$\mathrm{L}_z = -i\hbar\Big(x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}\Big)$$

expriment les moments angulaires quantiques comme opérateurs de rotations infinitésimales autour des axes (§VI.3, VII.11 et IX.3) multipliés par $i\hbar$. Les relations de commutation :

$$[\mathrm{L}_x, \mathrm{L}_y] = i\hbar \mathrm{L}_z;\ [\mathrm{L}_y, \mathrm{L}_z] = i\hbar \mathrm{L}_x;\ [\mathrm{L}_z, \mathrm{L}_x] = i\hbar \mathrm{L}_y$$

peuvent être réunies en : $\mathrm{L} \times \mathrm{L} = i\hbar \mathrm{L}$.

Cette relation n'étant évidemment pas invariante par un éventuel changement d'unité : $\mathrm{L} \to \lambda \mathrm{L}$, démontre l'existence d'une unité absolue pour la mesure du moment angulaire quantique, que l'on choisit égale à $\hbar$. L'opérateur $\mathrm{J} = \dfrac{\mathrm{L}}{\hbar}$ est l'opérateur cinétique et $\mathrm{J} \times \mathrm{J} = i\mathrm{J}$ (tandis que si l'on pose $\mathrm{M} = -i\mathrm{J}$ on a $\mathrm{M} \times \mathrm{M} = \mathrm{M}$).

On a vu à la fin du §VI.3 que l'algèbre de Lie $\mathfrak{O}(3)$ des rotations infinitésimales de E_3 s'identifie à l'algèbre de Lie $\mathfrak{S}\mathrm{U}(2)$ du groupe $\mathrm{SU}(2)$, formée des matrices 2×2 antihermitiennes de trace nulle. La complexifiée de $\mathfrak{S}\mathrm{U}(2)$ est l'algèbre de Lie complexe $\mathfrak{S}l(2;\mathbb{C})$, formée des matrices complexes 2×2 de trace nulle. Il est cependant préférable d'éviter l'utilisation d'une réalisation matricielle particulière pour la détermination des représentations irréductibles de $\mathfrak{O}(3)$.

La démarche suivie devient plus claire si l'on part, sans autre précision, d'une algèbre de Lie $\mathfrak{g}$, simple : $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, de dimension 3 sur un corps K,et qu'on analyse sa structure. Puisque tout élément H de $\mathfrak{g}$ peut s'écrire $\mathrm{H} = \sum [\mathrm{A}_i, \mathrm{B}_j]$, il en résulte que l'opérateur ad $\mathrm{H}(\mathrm{ad\ H} \cdot \mathrm{X} = [\mathrm{H}, \mathrm{X}])$ a une trace nulle. Ayant aussi un déterminant nul (puisque $[\mathrm{H}, \mathrm{H}] = 0$), son polynôme caractéristique s'écrit : $\text{dét}\,(tl - \mathrm{ad\ H}) = t^3 - \dfrac{1}{2}\,\mathrm{Tr}(\mathrm{ad\ H})^2 \cdot t$.

Par exemple, un calcul simple pour l'algèbre de Lie $\mathfrak{O}(3)$ montre que :

$$\text{si H} = \begin{vmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{vmatrix}, \text{ dét}\,(t\mathrm{I} - \mathrm{ad\ H}) = t(t^2 + a^2 + b + c^2).$$

S'il existe dans $\mathfrak{g}$ un élément H tel que l'opérateur ad H ait une valeur propre α non nulle dans le corps K, $(0, \alpha, -\alpha)$ sont alors les valeurs propres distinctes de ad H qui est diagonalisable, et $\mathfrak{g}$ *est dite diagonalisable*. C'est toujours le cas si $\mathfrak{g}$ est complexe. Par contre $\mathfrak{O}(3)$ ne l'est pas. Si X et Y sont des vecteurs propres de ad H pour α et $-\alpha$, ils sont définis à l'addition de λH près, et on peut les choisir de telle sorte que pour le *triplet* (H, X, Y) la structure de $\mathfrak{g}$ se réduise à :

$$[\mathrm{H}, \mathrm{X}] = \alpha \mathrm{X}; \; [\mathrm{H}, \mathrm{Y}] = -\alpha \mathrm{Y}; \; [\mathrm{X}, \mathrm{Y}] = \rho \mathrm{H}.$$

Notons bien qu'on peut *choisir arbitrairement* les coefficients α et ρ (non nuls) du triplet à l'aide de multiplication de H et Y par des constantes.

Les mathématiciens aiment prendre : $\alpha = 2$, $\rho = 1$. Les physiciens préfèrent le contraire $\alpha = 1$, $\rho = 2$.

Ces derniers, depuis 1925, et à la suite d'Elie Cartan (1894), diagonalisent l'algèbre de Lie complexe des moments angulaires $\mathfrak{S}l(2; \mathbb{C})$, dont on a vu les relations de commutation : $\mathrm{J} \times \mathrm{J} = i\mathrm{J}$, soient $[\mathrm{J}_2, \mathrm{J}_3] = i\mathrm{J}_1$; $[\mathrm{J}_3, \mathrm{J}_1] = i\mathrm{J}_2$; $[\mathrm{J}_1, \mathrm{J}_2] = i\mathrm{J}_3$ à l'aide de l'élément J_3, représentant le moment angulaire relativement à l'axe Oz. En définissant :

$$\mathrm{J}_+ = \mathrm{J}_1 + i\mathrm{J}_2 \text{ et } \mathrm{J}_- = \mathrm{J}_1 - i\mathrm{J}_2$$

on a : $[\mathrm{J}_3, \mathrm{J}_+] = \mathrm{J}_+$; $[\mathrm{J}_3, \mathrm{J}_-] = -\mathrm{J}_-$ et $[\mathrm{J}_+\mathrm{J}_-] = 2\mathrm{J}_3$.

L'élément $\mathrm{J}^2 = \mathrm{J}_1^2 + \mathrm{J}_2^2 + \mathrm{J}_3^2 = \dfrac{1}{2}\,(\mathrm{J}_+\mathrm{J}_- + \mathrm{J}_-\mathrm{J}_+) + \mathrm{J}_3^2$, de l'algèbre enveloppante, commute avec J_1, J_2, J_3 et a donc pour image, dans toute représentation irréductible complexe, un scalaire (les phycisiens déclarent, sans humour, qu'ils « diagonalisent simultanément » J_3 et J^2 ! !)

Soit θ une représentation linéaire, dans un espace vectoriel E, complexe, de dimension finie, de l'algèbre de Lie $\mathfrak{g}$ simple, complexe, de dimension trois, rapportée au triplet (H, X, Y).

On a : $[\theta(\mathrm{H}), \theta(\mathrm{X})] = \theta(\mathrm{H})\theta(\mathrm{X}) - \theta(\mathrm{X})\theta(\mathrm{H}) = \alpha\theta(\mathrm{X})$, etc.

Les valeurs propres de l'opérateur $\theta(\mathrm{H})$ sont appelées les *poids* de la représentation. Nous allons voir que ce sont toujours des multiples demi-entiers de α.

On remarque que si y est un vecteur propre de $\theta(\mathrm{H})$ pour un poids λ, $\theta(\mathrm{X})y$ et $\theta(\mathrm{Y})y$, s'ils sont non nuls, sont également des vecteurs propres pour des poids $(\lambda+\alpha)$ et $(\lambda-\alpha)$. Par exemple : $\theta(\mathrm{H})\theta(\mathrm{X})y = \{\theta([\mathrm{H},\mathrm{X}]) + \theta(\mathrm{X})\theta(\mathrm{H})\}y = (\lambda+\alpha)\theta(\mathrm{X})y$.

Les opérateurs $\theta(\mathrm{X})$ et $\theta(\mathrm{Y})$ sont donc des opérateurs de montée et de descente des valeurs propres. E étant de dimension finie et les valeurs propres associées aux $\theta(\mathrm{X})^p y$ distinctes, il existe un entier m maximal tel que $\theta(\mathrm{X})^m y$ soit non nul. En effet, puisque $\theta(\mathrm{X})^{q+1} = \dfrac{1}{\alpha}\,[\theta(\mathrm{X})^q\theta(\mathrm{H}),\,\theta(\mathrm{X})]$ pour tout q, il en résulte que les traces de tous les $\theta(\mathrm{X})^p$ sont nulles : *$\theta(\mathrm{X})$ est donc nilpotent, ainsi que $\theta(\mathrm{Y})$.* On vérifie immédiatement les égalités : $[\theta(\mathrm{H}),\theta(\mathrm{X})^p] = p\alpha\theta(\mathrm{X})^p$; $[\theta(\mathrm{X}),\theta(\mathrm{Y})^{p+1}] = (p+1)\rho\{\theta(\mathrm{H}) + \dfrac{p\alpha}{2}\}\,\theta(\mathrm{Y})^p$ et $[\theta(\mathrm{Y}),\theta(\mathrm{X})^{p+1}] = -(p+1)\rho\{\theta(\mathrm{H}) - \dfrac{p\alpha}{2}\}\,\theta(\mathrm{X})^p$.

Dès lors, soit m le plus grand entier tel que $\theta(\mathrm{X})^m \neq 0$. On a donc $\theta(\mathrm{X})^{m+1} = 0$. Pour tout vecteur x non nul appartenant au sous-espace $\theta(\mathrm{X})^m\mathrm{E}$, $x = \theta(\mathrm{X})^m y$, on a : $\theta(\mathrm{X})x = 0$ et d'après la dernière égalité : $\theta(\mathrm{H})x = \dfrac{m\alpha}{2}\,x$.

Un vecteur propre de $\theta(\mathrm{H})$ annulé par $\theta(\mathrm{X})$ est dit *primitif*, ou de *poids maximal*.

Supposons maintenant la représentation θ *irréductible*. D'après ce qui précède, il existe un vecteur v dont le poids maximal, $\dfrac{m\alpha}{2}$ est un multiple demi-entier de α. Les vecteurs propres successifs $\theta(\mathrm{Y})^p\cdot v$, $p = 1,3,\ldots$ de $\theta(\mathrm{H})$ ont des poids $\left(\dfrac{m}{2}-p\right)\alpha$ qui sont tous des multiples demi-entiers de α. Ils sont en nombre fini. Soit p le plus grand entier tel que $\theta(\mathrm{Y})^p\cdot v \neq 0$. Une égalité précédente donne alors :

$$0 = [\theta(\mathrm{X}),\theta(\mathrm{Y})^{p+1}]v = (p+1)\rho\{\theta(\mathrm{H}) + \frac{\pi}{2}\,\alpha\}\theta(\mathrm{Y})^p v,$$

soit : $\dfrac{m}{2} - p + \dfrac{p}{2} = 0$ et $p = m$.

Un raisonnement par récurrence montre alors que la suite des $(m+1)$ vecteurs : $\theta(\mathrm{Y})^m v, \theta(\mathrm{Y})^{m-1}v, \ldots \theta(\mathrm{Y})v, v$ ayant pour poids $-\dfrac{m}{2}\,\alpha, \left(-\dfrac{m}{2}+1\right)\,\alpha, \ldots, \dfrac{m}{2}\,\alpha$ est stable également par l'opérateur $\theta(\mathrm{X})$: ils forment donc la base d'un $\mathfrak{g}$-module qui, en raison de l'irréductibilité supposée, coïncide avec E.

Réciproquement, pour tout entier $m \geqslant 1$ on construit de cette façon une représentation irréductible de $\mathfrak{g}$ de dimension $m+1$, de poids maximal $\frac{m}{2}\,\alpha$, que l'on désigne par D_m. Si D_m est la représentation attachée au *spin d'une particule*, ce *spin* $\frac{m}{2}$ est demi-entier si m est impair, entier si m est pair. D_m détermine par exponentiation une représentation du groupe SU(2) qui devient une représentation de SO(3) lorsque m est pair, mais non lorsque m est impair. D_1 est la représentation de $\mathfrak{g}$ par les matrices 2×2 antisymétriques de $\mathfrak{S}l(2;\mathbb{C})$.

L'espace de la représentation est l'espace des spineurs complexes de dimension 2. Interroger le système par l'opérateur $J_3(\alpha = 1)$, c'est-à-dire mesurer le moment angulaire relativement à l'axe $0z$ donne alors pour résultat soit $\frac{1}{2}$ soit $-\frac{1}{2}$: ce sont les deux « états » possibles de ce moment angulaire, dit de spin $\frac{1}{2}$. Une rotation de 2Π fait passer de l'un à l'autre (cf. DR).

Nous laissons au lecteur le soin de vérifier que l'opérateur de Casimir de la représentation θ s'écrit :

$$C_\theta = \frac{1}{2\alpha\rho}\left\{\frac{\rho}{\alpha}\,\theta(H)^2 + \theta(X)\theta(Y) + \theta(Y)\theta(X)\right\}.$$

Il commute avec tous les opérateurs de la représentation. Dans D_m, l'opérateur C_θ est donc le scalaire :

$$C_\theta \cdot v = \frac{1}{2\alpha\rho}\left\{\frac{\rho}{\alpha}\,\frac{m^2\alpha^2}{4} + \rho\,\frac{m\alpha}{2}\right\}v = \frac{m(m+2)}{8},$$

indépendant du choix des paramètres α et ρ!

Notations des physiciens : de façon standard, on note $j = \frac{m}{2}$ soit $m = 2j$. D'autre part, on a alors $C_\theta = \frac{1}{2}\,\theta(J^2)$ ce qui donne pour le scalaire $\theta(J^2)$ la valeur $j(j+1)$.

Le moment angulaire s'appelle *spin* lorsqu'il se réfère à une propriété intrinsèque d'une particule élémentaire. Il peut alors prendre des valeurs demi-entières ou entières. Par contre lorsqu'il s'agit de mouvement effectif de rotation, le moment angulaire est dit *orbital* et ne peut prendre que des valeurs entières.

La démarche que l'on vient de suivre est entièrement algébrique et elle a permis de mettre en évidence les moments angulaires à valeurs demi-entières. En ce qui concerne les moments orbitaux, le point de vue de l'analyse permet de prendre en compte le caractère d'opérateurs différentiels des composantes du moment angulaire, qui s'écrivent en cordonnées polaires :

$$\mathrm{J}_+ = e^{i\varphi}\Big(\frac{\partial}{\partial\theta} + i\cot\theta\,\frac{\partial}{\partial\varphi}\Big) \quad \mathrm{J}_- = e^{-i\varphi}\Big(-\frac{\partial}{\partial\theta} + i\cot\theta\,\frac{\partial}{\partial\varphi}\Big).$$

$$\mathrm{J}_z = -i\,\frac{\partial}{\partial\varphi} \quad \mathrm{J}^2 = -\frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\varphi^2} - \frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\Big(\sin\theta\,\frac{\partial}{\partial\theta}\Big).$$

L'étude des fonctions propres de J^2 et J_z sera faite au §VII.11 : on obtient les « harmoniques sphériques » dont les divers espaces où opère directement le groupe des rotations SO(3) correspondent aux D_{2p}.

VI.6 – Produits scalaires invariants sur les espaces de spineurs

Nous nous contenterons dans ce dernier paragraphe de montrer la marche à suivre pour la détermination de ces produits scalaires.

L'idée conductrice est la suivante. Soit V un espace vectoriel sur un corps commutatif K. Un « produit scalaire » sur V, c'est-à-dire une forme bilinéaire, symétrique ou antisymétrique non-dégénérée sur V que nous noterons ici $[x \mid y]$, $\forall x, y \in \mathrm{V}$, détermine sur l'algèbre $\mathcal{L}(\mathrm{V})$ des opérateurs linéaire de V une antiinvolution, l'adjonction : $a \to a^*$, par : $[ax \mid y] = [x \mid a^*y]\ \forall x, y \in \mathrm{V}$

Si ρ est l'application linéaire de V sur V^* associée à droite de la forme bilinéaire par $[x \mid y] = \langle x, \rho y\rangle$ on a :

$$[ax \mid y] = \langle ax, \rho y\rangle = \langle x, {}^t a\rho y\rangle = [x \mid \overset{-1}{\rho}\,{}^t a\rho y]$$

d'où $a^* = \overset{-1}{\rho}\,{}^t a\rho$ et l'on vérifie immédiatement, puisque :

$$[x \mid ay] = \varepsilon[ay \mid x] = \varepsilon[y \mid a^*x] = \varepsilon^2[a^*x \mid y]$$

(où $\varepsilon \pm 1$), que l'on a aussi $[x \mid ay] = [a^*x \mid y]$.

Si λ est un scalaire non nul, la fonction $(x,y) \to \lambda[x \mid y]$ est encore un produit scalaire sur E, dont l'application associée à droite est $\lambda\rho$. Il détermine la même adjonction sur $\mathcal{L}(\mathrm{E})$ donc le même groupe d'invariance et les mêmes sous-espaces d'opérateurs symétriques $(a^* = a)$ ou antisymétriques $(a^* = -a)$.

Réciproquement, supposons que deux produits scalaires sur V, $[|]$ et $\{|\}$, d'applications associées à droite ρ et ρ', définissent la même adjonction sur V.

On a donc : $\overset{-1}{\rho}\,{}^t a\rho = \overset{-1}{\rho}{}'\,{}^t a\rho'$, soit $\rho'\,\overset{-1}{\rho}\,{}^t a = {}^t a\rho'\,\overset{-1}{\rho}$, quel que soit $a \in \mathcal{L}(\mathrm{V})$ donc quel que soit ${}^t a \in \mathcal{L}(\mathrm{V}^*)$. $\rho'\,\overset{-1}{\rho}$ appartient nécessairement au centre de $\mathcal{L}(\mathrm{V}^*)$ formé des multiples scalaires de l'identité et $\rho' = \lambda\rho$, $\{x \mid y\} = \lambda[x \mid y]$, $\forall x, y \in \mathrm{V}$.

Il est d'ailleurs nécessaire d'envisager une situation plus générale que la précédente afin de tenir compte, lorsqu'intervient le corps complexe, d'éventuels produits scalaires hermitiens. On va donc considérer un espace vectoriel V sur un corps commutatif Z, ce dernier étant muni d'une involution : $\lambda \to \overline{\lambda}$ dont le sous-corps des invariants est K.

Si K' est le sous-espace K-vectoriel de Z que forment les éléments antisymétriques pour l'involution : $\lambda \in \mathrm{K}' \Longrightarrow \overline{\lambda} = -\lambda$, $\mathrm{Z} = \mathrm{K} \oplus \mathrm{K}'$, et si $\jmath$ est un élément non nul de K', on a $\mathrm{K} = \jmath\mathrm{K}'$, $\mathrm{K}' = \jmath\mathrm{K}$, et $\mathrm{Z} = \mathrm{K} \oplus \jmath\mathrm{K}$.

Un produit scalaire hermitien, resp. antihermitien sur V est une forme sesquilinéaire non-dégénérée : $(x,y) \to [x \mid y]$ linéaire en le premier argument x, antilinéaire en le second y, satisfaisant à la symétrie hermitienne : $[y \mid x] = \overline{[x \mid y]}$, resp. à la symétrie antihermitienne : $[y \mid x] = -\overline{[x \mid y]}$.

Si $\overline{\jmath} = -\jmath \neq 0$, et si $[|]$ est un produit scalaire hermitien $\jmath[|]$ est un produit scalaire antihermitien, et réciproquement. Les deux structures sont donc équivalentes et, il suffit de considérer la première. L'application associée à droite ρ de V sur V^* définie par $[x \mid y] = \langle x, \rho(y)]$ est une bijection antilinéaire. L'adjonction : $a \to a^*$ de $\mathcal{L}(\mathrm{V})$ définie par :

$[ax \mid y] = [x \mid a^*y]$, $\forall x, y \in \mathrm{V}$, soit $a^* = \overset{-1}{\rho}\,{}^t a\rho$ est une anti-involution antilinéaire de $\mathcal{L}(\mathrm{V})$:

$$(ab)^* = b^*a^*,\ a^{**} = a,\ (\lambda a)^* = \overset{-1}{\rho}\,{}^t(\lambda a)\rho = \overline{\lambda}\overline{\rho}^{1t}a\rho = \overline{\lambda}a^*.$$

C'est un isomorphisme de $\mathcal{L}(\mathrm{V})$ sur l'algèbre opposée conjuguée $\overline{\mathcal{L}(\mathrm{V})^0}$.

Si $[||]$ et $\{|\}$ sont deux produits scalaires hermitiens sur V définissant la même adjonction, on a comme précédemment $\overset{-1}{\rho}\,{}^t a\rho = \overset{-1}{\rho'}\,{}^t a\rho'$ et $\rho'\,\overset{-1}{\rho}$ est un opérateur linéaire de $\mathfrak{L}(\mathrm{V}^*)$ appartenant au centre, donc multiple scalaire de l'identité. Il en résulte que $\rho' = \lambda\rho$ mais comme $[||]$ et $\{|\}$ ont tous deux la symétrie hermitienne, λ est un élément de K. Bien entendu, les produits scalaires $[||]$ et $j[||]$, où $\overline{j} = -j \neq 0$ ont eux aussi même adjonction.

Problème : Etant donné que les algèbres de Clifford des espaces quadratiques réguliers possèdent des antiinvolutions structurelle : τ et ν (§VI.2), on peut se demander si elles ne sont pas les adjonctions de produits scalaires qu'elles détermineraient ainsi sur les espaces de spineurs ce qui ne pourrait être, comme on l'a vu, qu'à un scalaire près.

Exemple VI.6.A : On a vu à l'exemple 4 du §VI.1. que l'algèbre de Clifford de l'espace euclidien E_3 est l'algèbre réelle sous-jacente à l'algèbre complexe $\mathbb{C}(2)$ ce qui permet de prendre pour espace des spineurs S le plan complexe $\mathbb{C}^2$.

L'antiinvolution principale τ est antilinéaire pour la structure complexe de $\mathrm{C}(\mathrm{E}_3)$: $\tau\mathrm{J} = \tau(e_1e_2e_3) = e_3e_2e_1 = -\mathrm{J}$.

Il est facile de vérifier que c'est l'adjonction de la métrique hermitienne standard sur $\mathbb{C}^2$: $[s \mid t] = [\varepsilon_1 s^1 + \varepsilon_2 s^2 \mid \varepsilon_1 t^1 + \varepsilon_2 t^2]$ $= s^1\overline{t}_1 + s^2\overline{t}_2$.

$$x = \begin{vmatrix} x^3 & x^1 - ix^2 \\ x^1 + ix^2 & -x^3 \end{vmatrix} = \sigma_1 x^1 + \sigma_2 x^2 + \sigma_3 x^3.$$

Les produits de deux matrices de Pauli : (définition VI.1.C) : $\sigma_j\sigma_k = i\sigma_l$ sont bien antihermitiennes et $\tau(\sigma_j\sigma_k) = -\sigma_j\sigma_k$, ainsi que $\sigma_1\sigma_2\sigma_3 = i\mathrm{I}$. On a donc quel que soit l'élément u de $\mathrm{C}(\mathrm{E}_3)$: $[us \mid t] = [s \mid u^\tau t]$.

Par contre la conjugaison ν est $\mathbb{C}$-linéaire : $\nu\mathrm{J} = \nu(e_1e_2e_3)$ $= (-e_3)(-e_2)(-e_1) = -\tau\mathrm{J} = \mathrm{J}$.

On vérifie que ν est l'adjonction du produit scalaire symplectique sur $\mathbb{C}^2$: $\{s \mid t\} = \{\varepsilon_1 s^1 + \varepsilon_2 s^2 \mid \varepsilon_1 t^1 + \varepsilon_2 t^2\} = s^1t^2 - s^2t^1$. Calculons en effet la matrice A^* de l'opérateur a^* adjoint de a :
si $\mathrm{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$, $\mathrm{A}^* = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}\begin{vmatrix} a & c \\ b & d \end{vmatrix}\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = \begin{vmatrix} d & -b \\ -c & a \end{vmatrix} = \check{\mathrm{A}}$.

On vérifie que si $x \in \mathrm{E}$, $\mathrm{X}^* = -\mathrm{X}$, et que de la même façon, les produits de deux matrices de Pauli sont antisymétriques : $\mathrm{A}^* = -\mathrm{A}$. On a donc quel que soit $u \in \mathrm{C}(\mathrm{E}_3)$: $\{us \mid t\} = \{s \mid u^v t\}$.

Théorème VI.6.A : Soit V un espace vectoriel sur le corps K muni d'un produit scalaire d'un type quelconque noté $[|]$, T l'adjonction correspondante sur l'algèbre $\mathfrak{L}_{\mathrm{K}}(\mathrm{V})$.

Pour un élément inversible g de $\mathfrak{L}_{\mathrm{K}}(\mathrm{V})$, les trois propriétés suivantes sont équivalentes :

a) l'automorphisme intérieur : $a \to \overset{-1}{g}\, ag$ de $\mathfrak{L}_{\mathrm{K}}(\mathrm{V})$ commute avec l'antiinvolution T.

b) g est un automorphisme « conforme » de V : il multiplie le produit scalaire $[|]$ par un scalaire non nul.

c) g conserve la relation d'orthogonalité dans V.

Les éléments g satisfaisant à ces propriétés forment un groupe.

Preuve : $(ga\overset{-1}{g})^{\mathrm{T}} = \overset{-1}{g}{}^{\mathrm{T}} a^{\mathrm{T}} g^{\mathrm{T}} = ga^{\mathrm{T}}\overset{-1}{g}$ si et seulement si $g^{\mathrm{T}} g a^{\mathrm{T}} = a^{\mathrm{T}} g^{\mathrm{T}} g$ soit si $g^{\mathrm{T}} g = \lambda \cdot \mathrm{I}$. Or $[gs \mid gt] = [s \mid g^{\mathrm{T}} gt]$.

Si $g^{\mathrm{T}} g = \lambda\mathrm{I}$: $[gs \mid gt] = \lambda[s \mid t]$ et $[s \mid t] = 0$ implique $[gs \mid gt] = 0$.

Réciproquement, soit u un opérateur inversible tel que $[us \mid ut] = 0$ chaque fois que $[s \mid t] = 0$. Si D est une droite portant s non nul, $u(\mathrm{D}^{\perp}) = \mathrm{D}^{\perp}$ d'où $u(\mathrm{D}^{\perp\perp}) = u(\mathrm{D}) = \mathrm{D}$: u conserve toutes les droites de V. Si $us = \alpha s$, $ut = \beta t$, $u(s+t) = \gamma(s+t)$, puisque u est linéaire $\alpha = \beta = \gamma$ et $u = \lambda\mathrm{I}$.

Exemple VI.6.B : Soient $(\mathrm{E}; q)$ un espace quadratique régulier, $\mathrm{C}(\mathrm{E}; q)$ son algèbre de Clifford, $\mathrm{G}(\mathrm{E}; q)$ son groupe de Clifford, S un espace de spineurs. S'il existe sur S des produits scalaire $[|]$ et $\{|\}$ dont les adjonctions sont dans $\mathrm{C}(\mathrm{E}; q)$ les antiinvolutions τ et ν (§VI.2), le groupe de Clifford conserve ces produits scalaires à un scalaire multiplicatif non nul près et opère donc sur S par des automorphismes « conformes » relativement à chacun d'eux.

En effet, un élément g de $\mathrm{G}(\mathrm{E}; q)$ est un produit $x_1 x_2 \ldots x_r$ de vecteurs réguliers et $g^{\tau} g = q(x_1)q(x_2)\ldots q(x_r) \cdot \mathrm{I} = \lambda\mathrm{I}$ et de même $g^{\nu} g = (-1)^r \lambda\mathrm{I}$. *Le groupe spinoriel (définition VI.3.D) conserve ces produits scalaires :* si $g \in \mathrm{Spin}(\mathrm{E}; q)$, $g^{\tau} g = g^{\nu} g = \mathrm{I}$.

Dans l'exemple VI.6.A ci-dessus le groupe spinoriel doit donc être contenu dans l'intersection du groupe unitaire $\mathrm{U}(2)$ et du groupe $\mathrm{Sl}(2; \mathbb{C})$, qui est le groupe $\mathrm{SU}(2)$.

En fait, $\mathrm{Spin}(\mathrm{E}_3)$ groupe multiplicatif des quaternions de norme un s'identifie à $\mathrm{SU}(2)$.

Terminons ce paragraphe en mentionnant quelques théorèmes généraux.

Théorème VI.6.B : Soit A une algèbre associative unitaire centrale simple de centre Z. Soient S et T deux antiinvolutions sur A dont les restrictions à Z déterminent la même involution de Z. Alors, S est la composée de T et d'un automorphisme intérieur U de A par un élément inversible, u, qui est soit T et S-symétrique, soit T et S-antisymétrique. Réciproquement, si u est un élément inversible de A, T-symétrique ou antisymétrique, la composée de T et de l'automorphisme intérieur par u est une antiinvolution de A dont la restriction à $\mathbb{Z}$ est la même que celle de T.

Preuve : Convenons de faire opèrer à droite sur A les automorphismes et les antiautomorphismes, et de noter a^{α} l'image de $a \in \mathrm{A}$ par α. Ces conventions entraînent que $a^{\alpha\beta} = (a^{\alpha})^{\beta}$.

L'application composée ST est évidemment un automorphisme de A qui induit l'identité sur Z. D'après le théorème de Skolem-Noether, il existe un élément inversible v de A tel que :

$$a^{\mathrm{ST}} = va \overset{-1}{v} .$$

En appliquant l'antiinvolution T aux deux membres et puisque T^2 est l'identité :

$$(a^{\mathrm{ST}})^{\mathrm{T}} = a^{\mathrm{S}} = (va \overset{-1}{v})^{\mathrm{T}} = \overset{-1}{v}{}^{\mathrm{T}} a^{\mathrm{T}} v^{\mathrm{T}} = \overset{-1}{u} a^{\mathrm{T}} u = a^{\mathrm{TU}}.$$

avec $u = v^{\mathrm{T}}$ d'où $\mathrm{S} = \mathrm{TU}$.

Mais il reste à exprimer que S est involutif : $\mathrm{S}^2 = \mathrm{TUTU} = \mathrm{I}$ soit $\overset{-1}{\mathrm{U}} = \mathrm{TUT}$ ce qui s'écrit :

$$ua \overset{-1}{u} = a^{\mathrm{TUT}} = (\overset{-1}{u} a^{\mathrm{T}} u)^{\mathrm{T}} = u^{\mathrm{T}} a \overset{-1}{u}{}^{\mathrm{T}} \text{ soit } \overset{-1}{u} u^{\mathrm{T}} a = a \overset{-1}{u} u^{\mathrm{T}}$$

quel que soit a. $\overset{-1}{u} u^{\mathrm{T}}$ appartient nécessairement au centre C de A et $u^{\mathrm{T}} = \lambda u$. Si $\lambda = \pm 1$, u est bien symétrique ou antisymétrique relativement à T. Sinon, l'élément $w = u + u^{\mathrm{T}} = (1 + \lambda)u$ est inversible et symétrique, et on a : $a^{\mathrm{S}} = \overset{-1}{u}\, a^{\mathrm{T}} u = (1+\lambda)\, \overset{-1}{w}\, a(1+\lambda)^{-1} w = \overset{-1}{w}\, a^{\mathrm{T}} w$ c.q.f.d.

Réciproquement, si l'on définit S par $a^{\mathrm{S}} = \overset{-1}{u}\, a^{\mathrm{T}} u$, on a bien $a^{\mathrm{S}^2} = a\, \overset{-1}{u}\, u^{\mathrm{T}} a\, \overset{-1}{u}{}^{\mathrm{T}} u$ lorsque $u^{\mathrm{T}} = \pm u$.

Application : Soient A une algèbre centrale simple complexe, M un A-module, τ une antiinvolution antilinéaire de A. Choisissons une base $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$ de l'espace vectoriel complexe M. Elle détermine sur M le produit scalaire hermitien standard : $\{s|t\} = \{\sum \varepsilon_j s^j \mid \sum \varepsilon_k t^k\} = \sum s^j \overline{t}^j$ auquel est associé sur $\mathrm{A} = \mathfrak{L}_{\mathbb{C}}(\mathrm{M})$ l'adjonction $u \to u^*$, soit matriciellement : $u^* = {}^t\, \overline{\mathrm{U}}$. D'après le théorème précédent, l'antiinvolution τ peut s'exprimer à l'aide de $*$: $u^\tau = \overset{-1}{f}\, u^* f$ soit matriciellement $\mathrm{U}^\tau = \overset{-1}{\mathrm{F}}\, {}^t\overline{\mathrm{U}}\mathrm{F}$ avec $f^* = f$, ${}^t\overline{\mathrm{F}} = \mathrm{F}$, car si l'on avait $f^* = -f$, ${}^t\overline{\mathrm{F}} = -\mathrm{F}$, on remplacerait F par $\mathrm{G} = i\mathrm{F}$ pour lequel ${}^t\overline{\mathrm{G}} = \mathrm{G}$. Mais un produit scalaire $[|]$ de symétrie hermitienne, de matrice $\mathrm{H} = {}^t\, \overline{\mathrm{H}}$ dans la base ε s'écrit $[s \mid t] = {}^t\, \mathrm{SH}\overline{\mathrm{T}}$ et l'adjonction ν correspondante est donnée par : $\mathrm{U}^v = {}^t\, \overset{-1}{\mathrm{H}}\, {}^t\overline{\mathrm{U}}{}^t\mathrm{H}$. Il suffit donc de prendre $\mathrm{H} = {}^t\, \mathrm{F}$ et l'on a ainsi prouvé que toute antiinvolition antilinéaire sur A est l'adjonction d'un produit scalaire « pseudohermitien » sur S.

A titre d'application soit $\mathrm{E}_{p,q}$ un espace pseudoeuclidien, espace vectoriel réel $\mathbb{R}^{p+q}$ muni de la forme quadratique $q(x) = (x^1)^2 + \ldots + (x^p)^2 - (x^{p+1})^2 - \ldots (x^{p+q})^2$, de dimension impaire $= p+q = 2r+1$. Le centre de l'algèbre de Clifford $\mathrm{C}(\mathrm{E}_{p,q})$ est formé des $\lambda \cdot 1 + \mu \mathrm{J}$ où $\mathrm{J} = e_1 e_2 \ldots e_n$. On a :

$$\begin{aligned}
\mathrm{J}^2 &= (-1)^r q(e_1) \ldots q(e_n) \cdot 1 = (-1)^{r+q} \cdot 1, \\
\mathrm{J}^\tau &= (-1)^r \mathrm{J}, \\
\mathrm{J}^\nu &= (-1)^{r+1} \mathrm{J}.
\end{aligned}$$

Il résulte alors de ce qui précède, le :

Théorème VI.6.C : L'algèbre de Clifford $\mathrm{C}(\mathrm{E}_{p,q})$ d'un espace pseudoeuclidien de dimension $p + q = 2r + 1$ impaire est une algèbre centrale simple complexe si et seulement si q et r sont de parités opposées ($r + q$ impair). Dans ce cas l'espace de spineurs

$S(E_{p,q})$ est un espace vectoriel complexe muni d'un produit scalaire pseudohermitien dont l'adjonction est soit τ (r impair) soit ν (r pair).

Remarque : On a $p-q = p+q-2q = 2(r-q)+1$. La condition r et q de parités opposées, soit $r - q = 2k + 1$ est équivalente à $p-q = 4k+3$ soit, congru à 3 ou à 7 mod 8. On retrouve le résultat de l'exemple 5 du §VI.4.

Les vecteurs d'un espace quadratique régulier $(E; q)$ devenant des opérateurs linéaires de l'espace des spineurs $S(E; q)$, leurs propriétés algébriques, conséquences de la métrique, ont d'importantes répercussions sur la structure de $S(E; q)$. Nous allons, pour les étudier, utiliser quelques propriétés des formes quadratiques.

Rappelons quelques définitions (cf. [DR]).

Définition VI.6.A :

1) Un *plan hyperbolique* P sur le corps K est un plan quadratique *régulier* qui satisfait à l'une des propriétés suivantes équivalentes :

a) P contient une droite isotrope,
b) P contient deux droites isotropes distinctes,
c) il existe une base (e_1, e_2) dans laquelle la forme quadratique s'écrit $q(x) = (x^1)^2 - (x^2)^2$,
d) il existe une base dans laquelle q s'écrit $q(x) = x^1x^2$.

2) Un *espace hyperbolique* H sur le corps K est un espace quadratique *régulier*, de dimension paire $n = 2p$, qui satisfait à l'une des propriétés équivalentes suivantes :

a) H est somme directe orthogonale de plans hyperboliques,
b) H est somme directe de deux sous-espaces totalement isotropes,
c) H contient un sous-espace totalement isotrope de dimension p.

3) Une forme quadratique régulière est dite *neutre* si c'est la forme quadratique d'un espace hyperbolique. C'est donc une forme de degré pair pour laquelle il existe des bases dans lesquelles elle peut s'écrire soit sous la forme $q(x) = (x^1)^2 + \dots (x^p)^2 - (x^{p+1})^2 - \dots - (x^{2p})^2$ soit sous la forme $q(x) = x^1x^2 + x^3x^4 + \dots x^{2p-1}x^{2p}$.

4) Un espace *pseudohermitien d'indice* (p, q) est un espace vectoriel complexe de dimension $p + q$ muni d'un produit qui dans une base normale s'écrit : $(x \mid y) = x^1\overline{y}^1 + \dots + x^p\overline{y}^p - x^{p+q}\overline{y}^{p+q}$

avec p signes plus et q signes moins. L'espace, et le produit scalaire, sont dits *neutres* lorsque $p = q$.

Théorème VI.6.D : Soient $(\mathrm{E}; q)$ un espace quadratique régulier, $\mathrm{C}(\mathrm{E}; q)$ son algèbre de Clifford, S un espace de spineurs de $(\mathrm{E}; q)$, S_+ un espace de spineurs de la sous-algèbre paire $\mathrm{C}_+(\mathrm{E}; q)$.

1) Un plan hyperbolique de $(\mathrm{E}; q)$ (s'il en existe) détermine canoniquement une décomposition de S et de S_+ en la somme directe de deux sous-espaces.

2) Si $[|]$ est un produit scalaire sur S ou S_+ pour lequel l'adjonction est l'une ou l'autre des antiinvolutions canoniques τ, ν les sous-espaces de S ou de S_+ associés par 1) à un plan hyperbolique sont totalement isotropes.

3) Si E contient un vecteur isotrope non nul, le produit scalaire $[|]$ est ou neutre ou symplectique.

Preuve :

1) Soit x un vecteur isotrope non nul de E : $x^2 = q(x) \cdot 1 = 0$ dans $\mathrm{C}(\mathrm{E}; q)$. En désignant par la même lettre x l'opérateur linéaire correspondant dans S le fait que x soit de carré nul est équivalent à $\operatorname{Im} x \subset \operatorname{Ker} x$. Si x est isotrope non nul, on peut trouver un autre vecteur isotrope non nul y tel que $(x \mid y) = \dfrac{1}{2}$ ou ce qui est équivalent $xy + yx = 1$ dans $\mathrm{C}_+(\mathrm{E}; q) \subset \mathrm{C}(\mathrm{E}; q)$. On a : $(xy)^2 = xy(1 - yx) = xy$ et $(yx)^2 = yx$ tandis que $(xy)(yx) = (yx)(xy) = 0$ xy et yx sont donc deux idempotents supplémentaires de $\mathrm{C}_+(\mathrm{E}; q)$ entièrement déterminés, à l'ordre près, par le plan hyperbolique, dont x et y forment une base isotrope.

Si $s \in \operatorname{Ker} x$, dans S, on a $s = (xy + yx)s = (xy)s \in \operatorname{Im} x$. Donc $\operatorname{Im} x = \operatorname{Ker} x = \operatorname{Im}(xy) = \operatorname{Ker}(yx)$ et $\dim(\operatorname{Im} x) = \dfrac{1}{2} \dim \mathrm{S}$.

Désignons par $\mathrm{S}(x) = \operatorname{Im} x = \operatorname{Ker} x$ le sous-espace de S associé au vecteur isotrope non nul x, donc à la droite isotrope qui le porte. Un plan hyperbolique de E détermine donc une décomposition de S : $\mathrm{S} = \mathrm{S}(x) \oplus \mathrm{S}(y)$ en la somme directe de deux sous-espaces de dimension $\dfrac{1}{2} \dim \mathrm{S}$. De même, si l'on note $\mathrm{S}_+(xy) = \operatorname{Im}(xy) = \operatorname{Ker}(yx)$, on a une décomposition de S_+ en somme directe : $\mathrm{S}_+ = \mathrm{S}_+(xy) \oplus \mathrm{S}_+(yx)$.

2) On a $x^\tau = x$, $x^v = -x$, $(xy)^\tau = (xy)^v = yx$. D'où dans S : $[xs \mid xt] = \pm[s \mid x^2 t] = 0$ et $\operatorname{Im} x$ est totalement isotrope.

Dans S_+ : $[xy\, s|xy\, t] = [s|yx\, xy\, t] = 0$ et $\operatorname{Im}(xy)$ est totalement isotrope. Puisque S ou S_+ est somme directe de deux sous-espaces totalement isotropes, nécessairement de dimension égale à la moitié de la dimension de l'espace, c'est que le produit scalaire est ou bien symplectique, ou bien neutre, c.q.f.d.

Exemple VI.6.C : On a vu au théorème VI.5.C que si $E_{p,q}$ est un espace pseudoeuclidien de dimension impaire $p+q = 2r+1$, l'espace des spineurs $S(E_{p,q})$ possède, associée à l'une des antiinvolutions ν ou τ, un produit scalaire pseudohermitien lorsque $r+q$ est impair.

Si *p et q* sont non nuls, ce produit scalaire est nécessairement neutre. Par contre, on a vu pour E_3, qui n'a pas de vecteurs isotropes non nuls, que le produit scalaire associé sur $S(E_3)$ à τ était hermitien.

VI.7 – L'opérateur de Dirac

De la même façon que l'application de Clifford φ d'un espace quadratique $(E; q)$ dans son algèbre de Clifford $C = C(E; q)$ permet d'obtenir une « racine carrée » de la forme quadratique q : $q(x) \cdot 1_C = \varphi(x)^2$, elle permet dans le cas d'un espace pseudoeuclidien $E_{r,s}$ d'obtenir un opérateur différentiel $\not{D}$, « racine carrée » du Laplacien Δ.

Si $e = (e_1, e_2, \ldots, e_n)$ est une base orthonormale de $E_{r,s}$, identifions les e_j à leurs images dans $C(E_{r,s})$. On a alors : $q(x) \cdot 1_C = (e_1 x^1 + e_2 x^2 + \ldots e_n x^n)^2$ et l'on définit *l'opérateur de Dirac* $\not{D}$ de E_n comme l'opérateur différentiel linéaire *à coefficients dans* C (coefficients constants mais « matriciels »)

$$\not{D} = e_1 \frac{\partial}{\partial x^1} + e_2 \frac{\partial}{\partial x^2} + \ldots + e_n \frac{\partial}{\partial x^n}.$$

Comme pour l'expression de q, il est immédiat de voir que $\not{D}$ ne dépend pas du choix de la base orthonormale e. $\not{D}$ opère sur *les champs différentiables de spineurs* sur $E_{r,s}$ fonctions différentiables de $E_{r,s}$ à valeurs dans un espace de spineurs S de $E_{r,s}$, par dérivation et « multiplication » par des éléments de C, et aussi sur les fonctions différentiables de $\mathrm{E}_{r,s}$ à valeurs dans son algèbre de Clifford $\mathrm{C}(\mathrm{E}_{r,s})$. Le carré $\not{D}^2$ est un opérateur scalaire, ou « diagonal » : $\not{D}^2 = \Delta$ est le Laplacien de $\mathrm{E}_{r,s}$.

Une représentation matricielle de $\mathrm{C}(\mathrm{E}_{r,s})$ détermine une expression matricielle de l'opérateur $\not{D}$.

Un champ de spineurs s sur $\mathrm{E}_{r,s}$ tel que $\not{D}s = 0$ est dit *harmonique.*

Exemples :

1) *L'opérateur de Dirac du plan euclidien* E_2.

$\mathrm{C}(\mathrm{E}_2) \equiv \mathbb{R}(2), \mathrm{S} \equiv \mathbb{R}^2$ représentation matricielle standard dans laquelle :

$$e_1 = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} ;\ e_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \text{ avec } \mathrm{J} = e_1 e_2 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \text{ et } \mathrm{J}^2 = -1,$$

J s'identifie à l'opérateur i lorsqu'on considère $\mathbb{R}^2$ comme la droite complexe C. Un champ de spineurs s est ici une fonction sur E_2 à valeurs dans $\mathbb{R}^2$ et :

$$\not{D}s = \left(e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y}\right) s = \begin{vmatrix} -\dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} \end{vmatrix} \begin{vmatrix} u \\ v \end{vmatrix} = \begin{vmatrix} -\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial y} \\ \dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x} \end{vmatrix}.$$

On voit que $\not{D}$ n'est autre que *l'opérateur de Cauchy Riemann* et les champs de spineurs harmoniques sont les fonctions analytiques $u + iv$ de la variable complexe $x + iy$.

Mais un champ de vecteurs de E_2 : $\mathrm{X} = ae_1 + be_2$ est un champ d'éléments de $\mathrm{C}(\mathrm{E}_2)$ c'est-à-dire une fonction à valeur matricielles sur laquelle $\not{D}$ opère également par :

$$\not{D}\mathrm{X} = \begin{vmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} \end{vmatrix} \begin{vmatrix} -a & b \\ b & a \end{vmatrix} = \begin{vmatrix} \dfrac{\partial a}{\partial x} + \dfrac{\partial b}{\partial y} & -\left(\dfrac{\partial b}{\partial x} - \dfrac{\partial a}{\partial y}\right) \\ \dfrac{\partial b}{\partial x} - \dfrac{\partial a}{\partial y} & \dfrac{\partial a}{\partial x} + \dfrac{\partial b}{\partial y} \end{vmatrix}$$

soit : $\not{D} = (\operatorname{div} \mathrm{X})\mathrm{I} + (\operatorname{rot} \mathrm{X})\mathrm{J}$.

2) *L'opérateur de Dirac de l'espace euclidien* E_3

$C(E_3) \equiv \mathbb{C}(2)$, $S = \mathbb{C}^2$. Dans la représentation matricielle standard :

$$e_1 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix};\ e_2 = \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix};\ e_3 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \quad \text{avec } J = e_1 e_2 e_3 = i\mathrm{I}.$$

Un champ de spineurs s est ici une fonction sur E_3 à valeurs dans $\mathbb{C}^2$.

$$\not{D} s = \begin{vmatrix} \dfrac{\partial}{\partial z} & \dfrac{\partial}{\partial x} - i\,\dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial x} + i\,\dfrac{\partial}{\partial y} & -\dfrac{\partial}{\partial z} \end{vmatrix} \begin{vmatrix} u \\ v \end{vmatrix} = \begin{vmatrix} \dfrac{\partial u}{\partial z} + \dfrac{\partial v}{\partial x} - i\,\dfrac{\partial v}{\partial y} \\ \dfrac{\partial u}{\partial x} + i\,\dfrac{\partial u}{\partial y} - \dfrac{\partial v}{\partial z} \end{vmatrix}.$$

Les spineurs harmoniques de E_3 peuvent être considérés comme une généralisation des fonctions analytiques.

3) *L'opérateur de Dirac de l'espace de Minkowski* $E_{1,3}$.

$C(E_{1,3}) \equiv \mathbb{H}(2)$ et $C(E_{1,3}) \otimes \mathbb{C} \equiv \mathbb{C}(4)$. On utilise essentiellement l'opérateur de Dirac complexe, opérant sur l'espace vectoriel complexe $S = \mathbb{C}^4$. Il s'écrit avec un ensemble de matrices de Dirac (Définition VI.1.D) :

$$\not{D} = \gamma_0\,\frac{\partial}{\partial x^0} + \gamma_1\,\frac{\partial}{\partial x^1} + \gamma_2\,\frac{\partial}{\partial x^2} + \gamma_3\,\frac{\partial}{\partial x^3}$$

Il a pour carré le Dalembertien :

$$\Box = \frac{\partial^2}{(\partial x^0)^2} - \frac{\partial^2}{(\partial x^1)^2} - \frac{\partial^2}{(\partial x^2)^2} - \frac{\partial^2}{(\partial x^3)^2}$$

Sur une variété riemannienne compacte (cf. Chapitre IX), l'opérateur de Dirac $\not{D}$ n'existe globalement que si la variété satisfait à une condition topologique particulière. Elle est dite dans ce cas spinorielle. (si elle ne l'est pas, elle possède toujours un revêtement d'ordre deux qui l'est). La courbure de la variété introduit alors un terme correctif entre $\not{D}$ et le laplacien Δ. On obtient en effet : $\not{D}^2 = \Delta + \dfrac{1}{4}\,\chi$, où χ est la courbure scalaire.

Remarque : Les éléments e_j sont impairs. Si l'espace des spineurs S se décompose en somme directe $S = S_+ \oplus S_-$ de deux sous-espaces de semi-spineurs échangés par Π (VI.2), il en est de même des champs de spineurs et $\not{D}$ les échange.

EXERCICES SUR LE CHAPITRE VI

EX. VI.1 :

Soit C l'algèbre de quaternions, algèbre de Clifford de l'espace quadratique $(\mathrm{E}; q)$ sur K où $q(e_1x+e_2y) = ax^2+by^2$. Montrer que C est isomorphe à l'algèbre de matrices K(2) si et seulement si il existe $(r,s) \in \mathrm{K}^2$ tels que $ar^2 + bs^2 = 1$. Montrer que dans le cas contraire C est un surcorps non-commutatif de K. (cf. DR).

EX. VI.2 :

Montrer à l'aide de la propriété multiplicative de la norme d'une algèbre de quaternions l'identité d'Euler :

$$\begin{aligned}&(a^2+b^2+c^2+d^2)(a'^2+b'^2+c'^2+d'^2)\\ &\quad = (aa'+bb'+cc'+dd')^2 + (ab'-ba'+cd'-dc')^2\\ &\quad + (ac'-bd'-ca'+db')^2 + (ad'+bc'-cb'-da')^2\end{aligned}$$

EX. VI.3 :

Soient e_1 un vecteur régulier de l'espace quadratique $(\mathrm{E}; q)$, $\mathrm{E}_1 = e_1^{\perp}$. En identifiant $\mathrm{C}_+(\mathrm{E})$ à l'algèbre de Clifford $\mathrm{C}(\mathrm{E}_1; q_1)$ comparer les conjugaisons ν de C(E) et ν_1 de $\mathrm{C}(\mathrm{E}_1)$. Déterminer l'automorphisme principal Π_1 et l'antiinvolution τ_1 de $\mathrm{C}(\mathrm{E}_1)$ en fonction de Π, τ, ν et de l'automorphisme intérieur de C(E) par e_1.

EX. VI.4 :

Soient C_4 et C'_4 les algèbres de Clifford respectives de l'espace euclidien C_4 et de l'espace antieuclidien E'_4.

Montrer que l'application de $\mathbb{R}^4$ dans C'_4 : $e_1 \to e'_2e'_3e'_4$; $e_2 \to e'_1e'_3e'_4$; $e_3 \to e'_1e'_2e'_4$; $e_4 \to e'_1e'_2e'_3$ détermine un isomorphisme d'algèbres unitaires de C_4 sur C'_4.

EX. VI.5 :

Soient $E_{n,(f)} = E_{r,s}$ l'espace pseudoeuclidien de dimension $n = r+s$ et d'indice $f = r-s$, $e = (e_1, e_2, \ldots, e_n)$ une base orthogonale avec $q(e_j) = +1$ pour $j = 1, 2, \ldots, r$: $q(e_k) = -1$ pour $k = r+1, \ldots n$. $J = e_1 e_2 \ldots e_n$ dans $C(E_{r,s})$. Montrer que l'application l'application $x \in E_{r,s} \to Jx \in C(E_{r,s})$ est une application de Clifford pour une forme quadratique q' sur l'espace vectoriel $E_{r,s}$. En déduire une comparaison des algèbres $C(E_{4m,0})$ et $C(E_{0,4m})$.

Montrer l'existence d'isomorphismes :

$$C_+(E_{n,(f)}) = C(E_{n-1,(f+1)}) = C(E_{n-1,(-f+1)})$$
$$C_+(E_{n,(f)}) = C_+(E_{n,(-f)})$$
$$C_+(E_{n,0}) = C_+(E_{0,n})$$

chaque fois que les termes sont définis.

EX. VI.6 :

Montrer à l'aide de représentations matricielles appropriées que $\mathrm{Spin}(4;\mathbb{C}) \equiv \mathrm{Sl}(2;\mathbb{C}) \times \mathrm{Sl}(2;\mathbb{C})$ $\mathrm{Spin}(4;\mathbb{R}) \equiv \mathrm{SU}(2) \times \mathrm{SU}(2)$. (cf. DR)

EX. VI.7 :

$E_{r,s}$ désignant l'espace pseudoeuclidien de type (r,s) montrer qu'il existe des isomorphismes d'algèbres unitaires, où $\otimes$ est le produit d'algèbres *ordinaire* (non gradué)

$$C(E_{n,0}) \otimes C(E_{0,2}) \equiv C(E_{0,n+2})$$
$$C(E_{0,n}) \otimes C(E_{2,0}) \equiv C(E_{n+2,0})$$
$$C(E_{r,s}) \otimes C(E_{1,1}) \equiv C(E_{r+1,s+1})$$
$$C(E_{n,0}) \otimes C(E_{8,0}) \equiv C(E_{n+8,0})$$
$$C(E_{n,0}) \otimes C(E_{0,8}) \equiv C(E_{0,n+8})$$

CHAPITRE VII

Tenseurs euclidiens et pseudoeuclidiens

On examine dans la première partie de ce chapitre quelques applications classiques des tenseurs euclidiens, en commençant par les tenseurs du second ordre symétriques, qui interviennent partout en géométrie et en mécanique. L'étude des courbes, surfaces et sous-variétés des espaces euclidiens conduit aux coordonnées « curvilignes » et à l'expression covariante ou tensorielle, de la dérivation des champs de vecteurs et des tenseurs. On est ainsi conduit de façon naturelle à la définition et aux propriétés des variétés, différentielles et riemanniennes, qui font l'objet du chapitre IX.

La fin de ce chapitre est consacrée aux propriétés spécifiques des espaces pseudoeuclidiens et à quelques applications à la physique.

VII.1 – Rappel de quelques propriétés élémentaires des espaces euclidiens

Un espace euclidien de dimension n, noté E_n est un espace vectoriel réel de dimension n muni d'une forme quadratique q non-dégénérée positive (on dit aussi strictement positive, ou « définie » positive). On peut ainsi y définir la norme ou longueur d'un vecteur $x : ||x|| = q(x)^{\frac{1}{2}}$. Un espace euclidien possède toujours des bases orthonormales : base orthogonales dont les vecteurs sont de norme un. Le choix d'une telle base est équivalent à un isomorphisme quadratique de E_n sur l'espace euclidien standard de dimension n : $\mathbb{R}^n$ muni du produit scalaire $(x|y) = \sum x^j y^j$ d'où $||x||^2 = (x|x)$ $= \sum (x^j)^2$.

Dans une base orthonormale, les composantes covariantes et contravariantes coïncident : $x_i = (e_i|x) = (e_i|\sum e_j x^j) = x^i$ et le produit scalaire s'écrit $(x|y) =^t \mathrm{XY}$.

Tous les sous-espaces d'un espace euclidien sont réguliers et euclidiens. Deux espaces euclidiens de même dimension sont quadratiquement isomorphes.

Soient maintenant ε une base orthonormale, $e = \varepsilon\mathrm{S}$ une base quelconque de E_n, $g_{ij} = (e_i|e_j)$, $\mathrm{G} = |g_{ij}|$. Si $x = e\mathrm{X} = \varepsilon\xi = \varepsilon\mathrm{SX}$, $(x|y) =^t \xi\eta =^t \mathrm{X}^t\mathrm{SSY} =^t \mathrm{XGY}$ d'où $\mathrm{G} =^t \mathrm{SS}$ et $\text{dét}\,\mathrm{G} = (\text{dét}\,\mathrm{S})^2 > 0$. Il en résulte immédiatement que si $x_1, x_2, \dots x_p$ sont p vecteurs de E_n, le déterminant de Gram : $\text{dét}|(x_i|x_j)|$, nul si les p vecteurs sont dépendants, est strictement positif s'ils sont linéairement indépendants ce qui généralise l'inégalité de Cauchy-Schwartz. L'algorithme d'orthogonalisation de Schmidt montre alors que $\text{dét}|(x_i|x_j)| \leqslant ||x_1|| \cdot ||x_2|| \dots ||x_p||$ l'égalité n'ayant lieu que si les vecteurs sont orthogonaux deux à deux (cf [DR]).

Rappelons que dans $\otimes\mathrm{E}_n$ et $\wedge^p\mathrm{E}_n$ l'extension du produit scalaire de E_n s'écrit sur des éléments décomposables :

$$(x_1 \otimes x_2 \otimes \dots x_p | y_1 \otimes y_2 \otimes \dots y_p) = (x_1|y_1)(x_2|y_2)\dots(x_p|y_p)$$
$$(x_1 \wedge x_2 \wedge \dots x_p | y_1 \wedge y_2 \wedge \dots y_p) = \text{dét}|(x_i|y_i)|$$

Si $e = (e_1, e_2, \dots e_n)$ est une base orthonormale de E_n, $\otimes^p e$ et $\wedge^p e$ sont également orthonormales : tous les produits scalaires étendus sont euclidiens.

Dans $\wedge^p \mathrm{E}_n$, la norme $||x_1 \wedge x_2 \wedge \dots x_p||$ d'un p-vecteur décomposable est appelée le *volume euclidien* des p vecteurs $(x_1, x_2, \dots x_p)$, soit :

$$\begin{aligned}\mathrm{Vol}(x_1, x_2, \dots x_p) &= ||x_1 \wedge x_2 \wedge \dots x_p|| \\ &= \text{dét}|(x_j|x_k)|^{1/2} (\leqslant ||x_1|| \cdot ||x_2|| \dots ||x_p||\end{aligned}$$

En particulier le carré du volume euclidien d'une base quelconque $e = (e_1, e_2, \dots, e_n)$ de E_n a pour valeur $\overline{g} = \text{dét}|(e_i|e_j)| = \text{dét}|g_{ij}|$. Si $e' = e\mathrm{S}$ est une autre base, $|g'_{ij}| =^t \mathrm{S}|g_{ij}|\mathrm{S}$ et $\overline{g}' = \overline{g}(\text{dét}\,\mathrm{S})^2$ ce qui montre que $\overline{g}$ est un pseudoscalaire covariant de variance $(+2)$. Sa racine carrée positive $\overline{g}^{1/2}$ se transforme par un changement de base en $\overline{g}'^{1/2} = |\text{dét}\,\mathrm{S}|\overline{g}^{1/2}$.

Si le changement de base conserve l'orientation de l'espace $\overline{g}^{1/2}$ se comporte comme un pseudoscalaire covariant de variance $(+1)$. C'est en fait un pseudoscalaire orienté (cf § III.8). Si $x_1, x_2, \dots x_n$ sont n vecteurs de E_n, de l'égalité $||x_1 \wedge x_2 \wedge \dots x_n|| = |\text{dét}\,\mathrm{X}| \cdot ||e_1 \wedge e_2 \wedge \dots e_n||$, il résulte que $\mathrm{Vol}(x_1, x_2, \dots x_n) = |\text{dét}\,\mathrm{X}| \cdot \overline{g}^{1/2}$. L'expression $\overline{g}^{1/2}\text{dét}\,\mathrm{X}$ est un *scalaire orienté* qui, s'il est non nul, prend donc la même valeur (positive) dans toutes les bases ayant même orientation que $(x_1, x_2, \dots x_n)$ et cette valeur changée de signe dans les bases ayant une orientation opposée. D'une manière analogue $\overline{g}^{(-1/2)}$ est un pseudoscalaire orienté contravariant de variance (-1). *La mesure des volumes euclidiens est la forme* $\overline{g}^{1/2} e^{*1} \wedge \dots e^{*n}$.

La matrice A^* de l'adjoint a^* de l'opérateur linéaire a, défini par $(ax|y) = (x|a^*y)$, $\forall x, y \in \mathrm{E}_n$, est dans une base *orthonormale*, la transposée ${}^t\mathrm{A}$ de la matrice de a.

Les opérateurs symétriques : $a^* = a$, resp. antisymétriques, $a^* = -a$ sont donc ceux qui, dans une base orthonormale, ont une matrice symétrique ${}^t\mathrm{A} = \mathrm{A}$, resp. antisymétrique ${}^t\mathrm{A} = -\mathrm{A}$.

Le produit scalaire de l'espace euclidien E_n peut se prolonger de deux façons à son complexifié $\mathrm{E}_n \otimes \mathbb{C}$:

a) par un produit scalaire hermitien :

$$(z_1|z_2) = (x_j + iy_1|x_2 + iy_2) = (x_1|x_2) + (y_1|y_2) + i\{(y_1|x_2) - (x_1|y_2)\}$$

d'où : $(z|z) = (x + iy|x + iy) = (x|x) + (y|y) = ||x||^2 + ||y||^2$

b) par un produit scalaire bilinéaire complexe :

$$[z_1|z_2] = [x_1 + iy_1|x_2 iy_2] = (x_1|x_2) - (y_1|y_2) + i\{(y_1|x_2) + (x_1|y_2)\}$$

pour lequel un vecteur complexe $z = x + iy$ est isotrope si et seulement si $||x|| = ||y||$ et $(x|y) = 0$.

On vérifie immédiatement que si $a \in \mathfrak{L}(\mathrm{E}_n)$ le complexifié de son adjoint $(a^*)_{\mathbb{C}}$ est identique à l'adjoint de son complexifié $(a_{\mathbb{C}})^*$ et ceci, aussi bien pour (I) que pour [I].

Si a est symétrique, $a_{\mathbb{C}}$ est symétrique pour chacun des produits scalaires sur $\mathrm{E}_n \otimes \mathbb{C}$. En particulier, puisque les valeurs propres de $a_{\mathbb{C}}$ sont celles de a, si λ est l'un d'elles et z un vecteur propre correspondant :

$$(a_{\mathbb{C}} z|z) = \lambda(z|z) = (z|a_{\mathbb{C}} z) = \overline{\lambda}(z|z)$$

d'où $\lambda = \overline{\lambda}$. On a ainsi une nouvelle propriété spécifique des espaces euclidiens. Les valeurs propres d'un opérateur symétrique sont réelles.

Deux vecteurs propres x et y associés à des valeurs propres distinctes λ et μ d'un opérateur symétrique a de E_n sont nécessairement orthogonaux. En effet $(ax|y) = \lambda(x|y) = (x|ay) = \mu(x|y)$ entraîne $(\mu - \lambda)(x|y) = 0$ d'où $(x|y) = 0$. Nous avons ainsi démontré le :

Théorème VII.1.A. Tout opérateur symétrique d'un espace euclidien est diagonalisable dans une base orthogonale, autrement dit, possède une base orthogonale de vecteurs propres.

On peut donner de ce théorème une démonstration directe, ne faisant pas intervenir le complexifié, et qui a l'avantage de donner des valeurs propres une interprétation géométrique intéressante. Soit a un opérateur linéaire quelconque de E_n.

La fonction sur la sphère unité $\mathrm{S}_{n-1} : x \rightarrow (x|ax)$ est continue et atteint son maximum λ_1 en un point au moins x_1 de S_{n-1} qui est compacte. Soit $y \in x_1^{\perp} \cap \mathrm{S}_{n-1}$ un point quelconque du grand cercle de S_{n-1} orthogonal à x_1. La fonction : $f(t) = (x_1 \cos t + y \sin t | a x_1 \cos t + a y \sin t)$ est différentiable et par le choix de x_1 atteint son maximum pour $t = 0$. Or $f(t) = (x_1|ax_1)\cos^2 t + (y|ay)\sin^2 t + \sin t \cos t\{(x_1|ay) + (y|ax_1)\}$ et $f'(0) = (x_1|ay) + (y|ax_1) = 0$.

Si a est symétrique, cette égalité devient $2(ax_1|y) = 0$. Devant être satisfaite quel que soit le choix de y, il en résulte nécessairement que $ax_1 \in (x_1^\perp)^\perp = \mathbb{R} \cdot x_1$, donc que x_1 est un vecteur propre de a pour la valeur propre λ_1.

Or si un opérateur symétrique a de E_n laisse stable un sous-espace F, il laisse évidemment stable son orthogonal $\mathrm{F}^\perp \cdot a$ laisse donc invariant le sous-espace $\mathrm{E}_{n-1} = x_1^\perp$ et l'on poursuit de la même façon.

On voit que les valeurs propres de a sont les maxima successifs de $(ax|x)$ sur $\mathrm{S}_{n-1}, \mathrm{S}_{n-2}, \ldots \mathrm{S}_1$.

Si q et q' sont deux formes quadratiques sur un même espace vectoriel, il n'existe pas en général de base de l'espace qui diagonalise simultanément q et q' (c'est-à-dire orthogonale à la fois pour q et q'). Supposons q régulière et soient ρ et ρ' les applications linéaires de E dans E^* associées aux formes bilinéaire symétriques b et b' définies par q et q'. Si la base e est orthogonale à la fois pour q et q' et si e^* est sa base duale, on a :$\rho e_j = \alpha_j e^{*j}$ et $\rho' e_j = \alpha'_j e^{*j}$ d'où $\overset{-1}{\rho}\, \rho' e_j = \lambda_j e_j$ avec $\lambda_j = \alpha'_j / \alpha_j$. Il est donc nécessaire que e soit une base de vecteurs propres (diagonalisante) pour l'opérateur linéaire $a = \overset{-1}{\rho}\, \rho'$ et c'est aussi suffisant. En effet, l'opérateur a est celui qui permet d'exprimer b' en fonction de b :$b'(x,y) = \langle x, \rho' y\rangle = b(x, \overset{-1}{\rho}\, \rho' y) = b(x, ay)$ et puisque b' est symétrique, a est un opérateur symétrique $b(x, ay) = b(ax, y)$. Si e_1 est un vecteur propre de a : $ae_1 = \lambda_1 e_1$, l'orthogonal $e_1^\perp = \{x; b(x, e_1) = 0\}$ est donc stable par a. Si $x \in e_1^\perp$ $b(ax, e_1) = b(x, ae_1) = \lambda_1 b(x; e_1) = 0$.

On peut ainsi par récurrence construire une base orthogonale pour b qui soit une base de vecteurs propres pour a et qui diagonalisent aussi b'.

Le théorème précédent VII.1.A. a donc pour conséquence le :

Théorème VII.1.B. Pour toute forme quadratique sur un espace euclidien E_n on peut trouver une base orthogonale de E_n qui la diagonalise.

Nous allons maintenant considérer les opérateurs antisymétriques. Les valeurs propres d'un tel opérateur sont imaginaires pures et ne peuvent donc être réelles qu'en étant nulles. En effet, si $a^* = -a$, $a^*_{\mathbb{C}} = -a_{\mathbb{C}}$. Si λ est une valeur propre de a donc de $a_{\mathbb{C}}$,

$(a_{\mathbb{C}}z|z) = \lambda(z|z) = (z|a^*_{\mathbb{C}}z) = -(z|\lambda z) = -\overline{\lambda}(z|z)$ d'où $\lambda = -\overline{\lambda}$ (où $(|)$ est l'extension hermitienne du produit scalaire.

Théorème VII.1.C. Les opérateurs antisymétriques d'un espace euclidien E_n sont ses « rotations infinitésimales », ou plus précisément les dérivées en l'identité des rotations de E_n. Ils forment une algèbre de Lie qui est l'algèbre de Lie du groupe orthogonal $\mathrm{O}(\mathrm{E}_n)$.

Preuve : Soit $t \in [0,1] \rightarrow r(t) \in 0_+(\mathrm{E}_n)$ un arc continu dans le groupe des rotations tel que $r(o)$ soit l'identité de E_n et que $r(t)$ soit dérivable pour $t = 0$. On a quel que soit t et quels que soient x, y : $(r(t)x|r(t)y) = (x|y)$ d'où, en dérivant pour $t = 0$: $(r'(\mathrm{O})x|y) + (x|r'(\mathrm{O})y) = 0$ qui montre que l'opérateur dérivée à l'origine $r'(\mathrm{O})$ est antisymétrique. Réciproquement, si t est un nombre réel et a un opérateur linéaire quelconque, la dérivée par rapport à t de la fonction : $t \in \mathbb{R} \rightarrow \exp(at) \in \mathrm{G}l(\mathrm{E}_n) \subset \mathfrak{L}(\mathrm{E}_n)$ est $a\exp(at)$. Si a est antisymétrique, la dérivée de la fonction numérique : $(\exp(at)x|\exp(at)y)$ est donc :

$$(a\exp(at)x|\exp(at)y) + (\exp(atx|a\ \exp(at)y) = 0.$$

Cette fonction est constante : $(\exp(at)x|\exp(at)y) = (x|y)$ ce qui prouve que $\exp(at)$ est un arc différentiable du groupe des rotations dont la dérivée en l'identité est a.

D'autre part, si a et b sont antisymétriques, leur crochet $[a,b] = ab - ba$ l'est aussi puisque :

$$[a,b]^* = b^*a^* - a^*b^* = ba - ab = -[a,b].$$

Les opérateurs antisymétriques de E_n forment donc une algèbre de Lie qui est celle des « vecteurs tangents » à l'origine du groupe des rotations.

Exemple : la rotation de vitesse constante ω du plan euclidien :

$$r(t) = \begin{vmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{vmatrix}$$

a pour dérivée à l'origine la rotation infinitésimale :

$$\Omega = r'(\mathrm{O}) = \begin{vmatrix} \mathrm{O} & -\omega \\ \omega & 0 \end{vmatrix}$$

et réciproquement $r(t) = \exp(\omega t)$.

Théorème VII.1.D. Soit $t \in [0, \tau] \to g(t) \in \mathrm{SO(E}_n)$ un arc différentiable dans le groupe $\mathrm{SO(E}_n)$ des rotations de l'espace euclidien E_n définissant un mouvement de rotation de cet espace, $x(t) = g(t)x_0$, t étant considéré comme le temps.

Le champ des vecteurs-vitesses à un instant donné t_0 est obtenu par l'action d'un opérateur antisymétrique $\Omega(t_0)$ de E_n appelé la *rotation instantanée* à l'instant t_0 : $x'(t_0) = \Omega(t_0)x(t_0)$.

Preuve : $x(t_0) = g(t_0)x_0$, $x(t) = g(t)\overset{-1}{g}(t_0)x(t_0)$ et $x'(t_0) = g'(t_0)\overset{-1}{g}(t_0)x(t_0)$. Or, si $t = t_0 + h$: $x(t_0 + h) = g(t_0 + h))\overset{-1}{g}(t_0)x(t_0)$.

L'application $h \to r(h) = g(t_0 + h))\overset{-1}{g}(t_0)$ est un chemin différentiable dans le groupe $\mathrm{SO(E}_n)$ issu de l'identité : $r(\mathrm{O}) = \mathrm{I}$, dont la dérivée en I : $r'(\mathrm{O}) = g'(t_0))\overset{-1}{g}(t_0) = \Omega(t_0)$ est un opérateur (antisymétrique d'après le théorème précédent, et $x'(t_0) = \Omega(t_0)x(t_0)$, cqfd.

Remarques :

1) Dire qu'un opérateur linéaire a de E_n est antisymétrique, revient à dire que la forme bilinéaire associée : $(x, y) \to (x|ay)$ est antisymétrique donc alternée : $(x|ax) = 0$, $\forall x$, autrement dit que a transforme un vecteur x en un vecteur orthogonal.

2) Le carré de a est un opérateur symétrique dont les valeurs propres sont négatives. En effet, si $a^2x = \lambda x$, on a : $(ax|ax) = ||ax||^2 = -(x|a^2x) = -\lambda||x||^2$ ce qui impose $\lambda \leqslant 0$.

3) Si F est un sous-espace de E_n stable par a, l'orthogonal $\mathrm{F}^\perp$ est également stable par a.

Théorème VII.1.E. Soit a un opérateur antisymétrique de l'espace euclidien $\mathrm{E}_n \cdot \mathrm{E}_n$ est une somme directe orthogonale de plans et de droites stables par a. La restriction de a à chacune des plans est une « rotation infinitésimale ». La restriction de a à chacune des droites est nulle.

Preuve : Puisque a^2 est symétrique, ses valeurs propres sont réelles. Si l'une d'elles est non nulle, soient e_1 un vecteur propre unitaire correspondant et $ae_1 = \alpha e_2$ avec $||e_2|| = 1$ et $\alpha \neq 0$.

D'après la remarque précédente, e_2 est orthogonal à e_1 et d'après le choix de e_1, $ae_2 = \frac{\lambda}{\alpha} e_1$. Le plan P_1 de base orthonormale (e_1, e_2) est donc stable par a et λ est déterminé par :

$$(ae_1|e_2) = \alpha(e_2|e_2) = -(e_1|ae_2) = -\frac{\lambda}{\alpha}(e_1|e_1) \text{ d'où } \lambda = -\alpha^2.$$

La restriction de a au plan (e_1, e_2) a donc pour matrice :

$$\begin{vmatrix} 0 & -\alpha \\ \alpha & 0 \end{vmatrix}.$$

On poursuit par récurrence dans le sous-espace $P_1^\perp$ de E_n jusqu'à épuisement des valeurs propres non nulles de a^2. Dans le sous-espace restant, stable par a, a^2 est nul donc, quel que soit x : $||ax||^2 = (ax|ax) = -(x|a^2x) = 0$ d'où $ax = 0$ et $a = 0$.

Exemple : dans l'espace euclidien E_3, pour tout opérateur antisymétrique non nul, c'est-à-dire pour toute rotation infinitésimale, on peut trouver une base orthonormale dans laquelle sa matrice s'écrit :

$$A = \begin{vmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \text{ et } \exp At = \begin{vmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

VII.2 – Le produit vectoriel

Rappelons (§ V.5) que dans l'algèbre extérieure $\wedge E_n$ de l'espace euclidien E_n les opérateurs linéaires adjoints des multiplications extérieures à gauche ou à droite, sont les produits intérieurs ; définis par :

$$(t \wedge u|v) = (t|u \lrcorner\, v) \text{ et } (v|u \wedge t) = (v \llcorner\, u|t).$$

Les bases orthonormales de même orientation de E_n déterminent la même base $\overline{e}$ de $\wedge^n E_n$ élément de volume orienté de E_n.

L'automorphisme canonique de dualité, correspondant α, est défini par :

$$\alpha(u) = u \lrcorner\, \overline{e} \text{ soit } t \wedge u = (t|\alpha(u))\overline{e} \text{ et sur } \wedge^p E_n : \alpha^2 = (-1)^{p(n-p)}.$$

Si $e = (e_1, e_2, \dots e_n)$ est une base orthonormale positive de E_n on calcule immédiatement les images des vecteurs de la base $\wedge e$ de $\wedge \mathrm{E}_n$ par α. Avec les notations usuelles : $(e'_\mathrm{I} \wedge e_\mathrm{I}|\overline{e}) = \varepsilon(\mathrm{I}')$ où $\varepsilon(\mathrm{I}')$ est le signe de la permutation qui permet de passer de $(1, 2, \dots n)$ à $(\mathrm{I}', \mathrm{I})$ tandis que $(e_\mathrm{J} \wedge e_\mathrm{I}|\overline{e}) = 0$ si $\mathrm{J} \neq \mathrm{I}'$. D'où : $\alpha(e_\mathrm{I}) = e_\mathrm{I} \lrcorner\ \overline{e} = \varepsilon(\mathrm{I}')e'_\mathrm{I}$. Rappelons encore la propriété :

Proposition VII.2

L'automorphisme de dualité α de l'algèbre extérieure d'un espace euclidien est une isométrie. (*cf* Corollaire V.5).

Preuve : Sur $\wedge^p \mathrm{E}_n$: $t \wedge u = (t|\alpha(u))\overline{e} = (-1)^{p(n-p)} u \wedge t = (u|(-1)^{p(n-p)}\alpha(t))\overline{e} = (\overset{-1}{\alpha}(t)|u)\overline{e}$. D'où : $(\alpha(t)|\alpha(u)) = (t|\overset{-1}{\alpha}\alpha(u)) = (t|u)$ c.q.f.d.

On a défini au § IV.10 (définition IV.10.A) le produit vectoriel de $(n-1)$ vecteurs : $x_1, x_2, \dots x_{n-1}$ d'un espace vectoriel de dimension n : c'est une pseudoforme linéaire sur E qui pour chaque base e fait correspondre le contracté sur les $(n-1)$ derniers indices du pseudotenseur ε de Levi-Civita avec les $(n-1)$ vecteurs. Ses composantes dans e sont $\sum \varepsilon_{j i_1 i_2 \dots i_{n-1}} x_1^{i_1} x_2^{i_2} \dots x_{n-1}^{i_{n-1}}$. Son contracté avec un vecteur z est le déterminant des composantes de z, $x_1, \dots x_{n-1}$ dans la base e, ou le volume algébrique de ces vecteurs : $\text{dét}_e(z, x_1, x_2, \dots x_{n-1}$. Puisque l'espace E est un espace euclidien E_n on a : $\mathrm{Vol}(z, x_1, x_2, \dots x_{n-1}) = \text{dét}_e(z, x_1, x_2, \dots x_{n-1})\overline{g}^{1/2}$ qui est un scalaire orienté égal à $\overline{g}^{1/2} \sum (-1)^{i-1} z^i \text{dét}(\mathrm{X}^{1,2,\dots\widehat{i},\dots n})$ $= \sum z^i u_i$.

Les $u_i = (-1)^{i-1}\overline{g}^{1/2}\text{dét}|\mathrm{X}^{1,2,\dots\widehat{i},\dots n}|$ sont les composantes covariantes d'un *vecteur* orienté de E_n que l'on note $x_1 \times x_2 \times \dots x_{n-1}$ et que l'on continue d'appeler produit vectoriel. Il est donc défini par la propriété que son contracté avec un vecteur quelconque z, c'est-à-dire son produit scalaire avec z est égal au volume algébrique de $(z, x_1, x_2, \dots x_{n-1})$. Dans une base *orthonormée* e, on a

$$\begin{aligned} z \wedge x_1 \wedge x_2 \wedge \dots x_{n-1} &= \text{dét}_e(z, x_1, x_2, \dots x_{n-1})\overline{e} \\ (z \wedge x_1 \wedge x_2 \wedge \dots x_{n-1}|\overline{e}) &= \text{dét}_e(z, x_1, x_2, \dots x_{n-1}) \\ &= (z|x_1 \times x_2 \times \dots x_{n-1}) \end{aligned}$$

d'où : $(z|(x_1 \wedge x_2 \wedge \dots x_{n-1}) \lrcorner\ \overline{e} = (z|x_1 \times x_2 \times \dots x_{n-1})$.

Le produit vectoriel de $(n-1)$ vecteurs est le dual de leur produit extérieur :

$$x_1 \times x_2 \times \ldots x_{n-1} = \alpha(x_1 \wedge x_2 \wedge \ldots x_{n-1}) = (x_1 \wedge x_2 \wedge \ldots x_{n-1}) \lrcorner\, \overline{e}.$$

Puisque α est une isométrie, on a donc :

$$||x_1 \times x_2 \times \ldots x_{n-1}|| = ||x_1 \wedge x_2 \wedge \ldots x_{n-1}||.$$

On déduit immédiatement de ces égalités les propriétés de $x_1 \times x_2 \times \ldots x_{n-1}$:

a) il est non nul si et seulement si $x_1, x_2, \ldots x_{n-1}$ sont linéairement indépendants

b) en remplaçant z par chaque x_j, on voit que le produit vectoriel $x_1 \times x_2 \times \ldots x_{n-1}$ est orthogonal à chacun d'eux, donc à l'hyperplan qu'ils sous-tendent, et qu'il forme avec eux une base de E_n.

c) en remplaçant z par le produit vectoriel lui-même on obtient :

$$||x_1 \times x_2 \times \ldots x_{n-1}||^2 = \text{dét}_e(x_1 \times x_2 \times \ldots x_{n-1}, x_1, x_2, \ldots, x_{n-1})$$

ce qui montre d'abord que le second membre est positif et que la base $(x_1 \times x_2 \times \ldots x_{n-1}, x_1, x_2, \ldots x_{n-1})$ a même orientation que la base fixée initialement. Le second membre, vu b), est le produit de la longueur $||x_1 \times x_2 \times \ldots x_{n-1}||$ du produit vectoriel par le volume euclidien des $(n-1)$ vecteurs ce qui redonne l'isométrie :

$$||x_1 \times x_2 \times \ldots x_{n-1}|| = \mathrm{Vol}\cdot(x_1, x_2, \ldots x_{n-1}) = ||x_1 \wedge x_2 \wedge \ldots x_{n-1}||.$$

On retrouve ainsi les propriétés usuelles du produit vectoriel de E_3. Puisque la dualité α est une isométrie, on a dans E_n :

$$\begin{aligned}(x_1 \times x_2 \times \ldots x_{n-1} | y_1 \times y_2 \times \ldots y_{n-1}) \\ = (x_1 \wedge x_2 \wedge \ldots x_{n-1} | y_1 \wedge y_2 \wedge \ldots y_{n-1}) = \text{dét}|(x_i|y_i)|.\end{aligned}$$

Dans E_3 on retrouve la formule connue du produit scalaire de deux produits vectoriels :

$$\boxed{(a \times b | c \times d) = (a|c)(b|d) - (a|d)(b|c)}.$$

Le produit vectoriel de deux vecteurs dans l'espace euclidien E_3 possède une quantité de propriétés remarquables. On appelle souvent «*produit mixte*» de trois vecteurs x, y, z de l'espace euclidien

E_3 *orienté*, et on note $[x, y, z]$ leur volume algébrique relativement à l'unité de volume $\overline{e}$, soit dans une base orthonormée positive e :

$$[x,y,z] = \text{dét}_e(x,y,z) = (x \wedge y \wedge z|\overline{e}) = \text{dét}\begin{vmatrix} x^1 & y^1 & z^1 \\ x^2 & y^2 & z^2 \\ x^3 & y^3 & z^3 \end{vmatrix}$$

$[x,y,z] = (x|y \times z)$, définition du produit vectoriel. $[x,y,z]$ est invariant par toute permutation circulaire de ses termes, ce qui donne les égalités : $(x|y \times z) = (y|z \times x) = (z|x \times y)$.

Il en résulte la relation fort simple :

$$\boxed{(x|y \times z) = (x \times y|z)}\,.$$

En l'utilisant, l'égalité du produit scalaire de deux produits vectoriels :

$$(a \times b|c \times d) = ((a \times b) \times c|d) = (a|c)(b|d) - (a|d)(b|c)$$

donne la formule bien connue du double produit vectoriel :

$$\boxed{(a \times b) \times c = (a|c)b - (b|c)a}$$

ou aussi

$$\boxed{a \times (b \times c) = ((a|c)b - (a|b)c)}\,.$$

En y remplaçant c par $c \times d$:

$$(a \times b) \times (c \times d) = (a|c \times d)b - (b|c \times d)a$$

on obtient la formule du triple produit vectoriel :

$$\boxed{(a \times b) \times (c \times d) = [a,c,d]b - [b,c,d]a}\,.$$

D'une façon générale, dans un espace euclidien orienté E_n, tout élément $a \in \wedge^2 E_n$ détermine un opérateur linéaire de E_n par : $x \in E_n \to a \llcorner x$. Dans le cas d'un bivecteur décomposable par exemple, on a (§ IV.10) : $(u \wedge v) \llcorner x = (u|x)v - (v|x)u$ ce qui est égal à $(v \otimes u - u \otimes v)x$ (*il faut bien noter l'ordre de u et v dans les deux termes*.

On vérifie immédiatement que l'opérateur ainsi associé à a est antisymétrique : $(a \llcorner x|y) = (a|x \wedge y) = -(a|y \wedge x) = -(a \llcorner y|x) = -(x|a \llcorner y)$.

Réciproquement, tout opérateur antisymétrique de E_n peut être représenté de cette façon. En effet, soit e une base orthonormale de E_n. Un opérateur linéaire de E_n est représenté par un tenseur d'ordre deux : $u = \sum a^{ij} e_i \otimes e_j$ de matrice $|a^{ij}|$: $ux = \sum a^{ij} e_i(e_j|x)$ et $(ux|y) = \sum a^{ij}(e_i|y)(e_j|x)$.

Ainsi qu'on l'a déjà vu, u est antisymétrique si et seulement si $a^{ij} = -a^{ji}$ et $u = \sum\limits_{i<j} a^{ij}(e_i \otimes e_j - e_j \otimes e_i)$.

Si $a = \sum\limits_{i<j} a^{ij} e_i \wedge e_j$, on a : $ux = -a \llcorner x$.

Revenons à E_3. La dualité α applique ici chacun des espaces $\wedge^2\mathrm{E}_3$ et E_3 sur l'autre et à $a \in \wedge^2\mathrm{E}_3$ correspond son vecteur dual $\alpha(a)$: si

$$a = a^1 e_2 \wedge e_3 + a^2 e_3 \wedge e_1 + a^3 e_1 \wedge e_2, \ \alpha(a) = a^1 e_1 + a^2 e_2 + a^3 e_3.$$

Puisque α est une isométrie, on a quel que soit $y \in \mathrm{E}_3$:

$$(a \llcorner x|y) = (a|x \wedge y) = (\alpha(a)|x \times y) = (\alpha(a) \times x|y).$$

Il en résulte que dans E_3, l'action d'un opérateur antisymétrique $a \in \wedge^2\mathrm{E}_3$ s'effectue au moyen du produit vectoriel par son vecteur dual : $a \llcorner x = \alpha(a) \times x$.

La matrice de l'opérateur antisymétrique $x \to a \llcorner x$ s'écrit dans une base orthonormée : $a = a^1(e_3 \otimes e_2 - e_2 \otimes e_3) + a^2(e_1 \otimes e_3 - e_3 \otimes e_1) + a^3(e_2 \otimes e_1 - e_1 \otimes e_2)$ soit :

$$\mathrm{A} = \begin{vmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{vmatrix}$$

et l'image de $x = x^1 e_1 + x^2 e_2 + x^2 e_3$ s'obtient par les règles classiques :

$$\begin{vmatrix} a^1 \\ a^2 \\ a^3 \end{vmatrix} \times \begin{vmatrix} x^1 \\ x^2 \\ x^3 \end{vmatrix} = \begin{vmatrix} a^2x^3 & -a^3x^2 \\ a^3x^1 & -a^1x^3 \\ a^1x^2 & -a^2x^1 \end{vmatrix} \text{ ou } \begin{vmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{vmatrix} \begin{vmatrix} x^1 \\ x^2 \\ x^3 \end{vmatrix}.$$

Terminons ce paragraphe par une formule fort utile exprimant le composé de deux opérateurs antisymétriques de E_3 : Si a et b sont deux vecteurs de E_3, $\overline{a}$ et $\overline{b}$ les opérateurs antisymétriques associés : $x \to \alpha(a) \mathsf{L}\, x$ et $x \to \alpha(b) \mathsf{L}\, x$, on a

$$\overline{b} \circ \overline{a} = a \otimes b - (a|b)\mathrm{I}$$

qui n'est qu'un aspect de la formule du double produit vectoriel :

$$b \times (a \times x) = a(b|x) - (a|b)x.$$

VII.3 – Tenseurs euclidiens du second ordre symétriques

L'identification des tenseurs de même ordre et de diverses variances permet de considérer un tenseur du second ordre t d'un espace euclidien E_n comme une forme bilinéaire b sur E_n ou comme un opérateur linéaire a de E_n, liés par les relations : $x \cdot t \cdot y = b(x, y) = (x|ay)$, $\forall x, y \in \mathrm{E}_n$.

t symétrique est équivalent à b ou a symétrique, et dans ce cas, t est entièrement déterminé par la forme quadratique $q(x) = b(x, x) = (x|ax)$, donc aussi par le couple de quadriques à centre Q_+ d'équation $q(x) = 1$ et Q_- d'équation $q(x) = -1$ qui en donnent une représentation géométrique.

Afin d'éviter toute confusion, nous dirons que deux vecteurs x et y sont *conjugués* par rapport t lorsque $x \cdot t \cdot y = (x|ay) = b(x, y) = 0$ (au lieu de dire que x et y sont « orthogonaux » vis-à-vis de la forme b).

Si $x \to q(x)$ est une forme quadratique sur un espace vectoriel réel quelconque, de dimension finie, c'est une fonction polynomiale, donc différentiable en tout point. Si $b(x, y) = \dfrac{1}{2}(q(x + y) - q(x) - q(y))$ on déduit directement la différentielle de q en un point x_0 de l'égalité : $q(x_0 + h) = q(x_0) + 2b(x_0, h) + q(h)$. C'est donc la forme différentielle $h \to dq_{x_0}(h) = 2b(x_0, h)$. Son noyau est l'hyperplan conjugué de x_0.

L'hyperplan affine tangent en x_0 à la quadrique Q_0 d'équation $q(x) = q(x_0)$ est l'hyperplan parallèle au noyau de la différentielle en ce point : $b(x_0, x - x_0) = 0$, ensemble des vecteurs affines issus de x_0 parallèles aux vecteurs conjugués de x_0.

Si a est l'opérateur symétrique permettant d'exprimer b à l'aide du produit scalaire : $b(x, y) = (x|ay)$, la demi-différentielle de q en x_0 est la forme linéaire produit scalaire par ax_0 :

$$\frac{1}{2} dq_{x_0}(h) = b(x_0, h) = (ax_0|h)$$

et l'hyperplan tangent à la quadrique Q_0 en x_0 est donc l'hyperplan affine passant par x_0 et orthogonal à ax_0 : $(ax_0|x - x_0) = 0$. Il en résulte que l'hyperplan tangent à Q_0 en x_0 est orthogonal au rayon vecteur x_0 si et seulement si $ax_0 = \lambda x_0$, x_0 est un vecteur propre de a.

Définition VII.3

Soient t un tenseur symétrique du second ordre de l'espace euclidien E_n, a l'opérateur linéaire symétrique qui lui correspond. Un sous-espace F de E_n est dit *principal relativement à t* s'il est appliqué en lui-même par a. En particulier les droites ou directions principales, ou axes principaux, sont les droites propres de a.

Puisque a est symétrique, $a\mathrm{F} \subset \mathrm{F}$ implique $a\mathrm{F}^{\perp} \subset \mathrm{F}^{\perp}$. L'orthogonal d'un sous-espace principal l'est aussi. Si $a\mathrm{F} \subset \mathrm{F}$, la restriction de a à l'espace euclidien F est symétrique et F possède donc une base orthogonale de vecteurs propres de a. Les sous-espaces principaux de t sont donc les sous-espaces engendrés par un certain nombre de vecteurs propres de a.

Si u est l'involution orthogonale de E_n qui laisse fixes les éléments de F et change en leurs opposés ceux de $\mathrm{F}^{\perp}$ (on dit abusivement que u est la « symétrie » orthogonale par rapport à F), u laisse t invariant donc aussi chaque quadrique associée $q(x) = \alpha$. En effet, puisque $\overset{-1}{u} = u$, cette invariance s'exprime sur l'opérateur a par la propriété $uau = a$ soit $au = ua$, qui est évidente sur tout vecteur propre de a.

Rapportée à des axes principaux, l'équation de la quadrique Q_α s'écrit : $\lambda_1(x^1)^2 + \lambda_2(x^2)^2 + \dots \lambda_n(x^n)^2 = \alpha$. Les valeurs propres de a : $\lambda_1, \lambda_2, \dots \lambda_n$ sont les *coefficients principaux de t*.

VII.4 – Trois théorèmes d'Apollonius sur les coniques et les quadriques

Soient q une forme quadratique *régulière* de signature (r, s) sur l'espace euclidien E_n, b la forme bilinéaire symétrique associée, $\overset{-1}{b}$ la forme bilinéaire inverse β l'isomorphisme de E_n sur E_n^* associé à b. Les produits doublement contractés des tenseurs $\overset{-1}{b}$ et g ou des tenseurs $\overset{-1}{g}$ et b sont des scalaires indépendants du choix de la base permettant de les calculer.

Si ρ est l'isomorphisme de E_n sur E_n^* associé au tenseur métrique g, ce sont d'ailleurs les traces des opérateurs linéaires de E_n : $\overset{-1}{\beta} \circ \rho$ et $\overset{-1}{\rho} \circ \beta$, soient $\sum \overset{-1}{b}{}^{ij} g_{ij}$ et $\sum \overset{-1}{g}{}^{ij} b_{ij}$. Si Q est l'ensemble des x de E_n tels que $q(x) = \pm 1$, c'est ou bien un ellipsoïde si r ou $s = 0$, ou bien un couple d'hyperboloïdes Q_+ et Q_-.

Commençons par diagonaliser b par une base e de E de vecteurs conjugués. La matrice B de b dans e est diagonale avec $b_{ii} \neq 0, \forall i$; celle de $\overset{-1}{b}$ est la matrice diagonale inverse $(\overset{-1}{b})^{ii} = (b_{ii})^{-1}$. On peut supposer $b_{ii} > 0$ pour $i = 1, 2, \dots r, b_{ii} < 0$ pour $i > r$.

Dans cette base e : $\overset{-1}{b} : g = \sum\limits_i (b_{ii})^{-1} g_{ii} = \sum\limits_i b_{ii}^{-1} (e_i | e_i)$.

La demi-droite portant le vecteur de base e_i coupe Q en un point $\lambda_i e_i$, où $\lambda_i > 0$ satisfait à $q(\lambda_i e_i) = b(\lambda_i e_i, \lambda_i e_i) = \lambda_i^2 b_{ii} = \pm 1$. On a donc : $\lambda_i^2 = (b_{ii})^{-1}$ (signe de b_{ii}), où le signe de b_{ii} est $(+)$ pour $i = 1, 2, \dots r$, et $\lambda_i \cdot e_i$ appartient à Q_+ et $(-)$ pour $i = r+1, \dots, r+s = n$, $\lambda_i \cdot e_i$ appartenant à Q_-.

Le carré de la longueur de ce rayon vecteur est :

$$\begin{aligned} l_i^2 = (\lambda_i e_i | \lambda_i e_i) = \lambda_i^2 (e_i | e_i) &= (\text{signe de } b_{ii}) b_{ii}^{-1} (e_i | e_i) \\ &= (\text{signe de } b_{ii}) \frac{g_{ii}}{b_{ii}} \end{aligned}$$

D'où : $\overset{-1}{b} : g = \sum \dfrac{g_{ii}}{b_{ii}} = \sum\limits_{i=1}^{r} l_i^2 - \sum\limits_{j=1}^{s} l_{r+j}^2$.

L'indépendance vis-à-vis de la base choisie donne le :

1er Théorème d'Apollonius. La somme des carrés des longueurs de n rayons conjugués d'un ellipsoïde, ou d'un couple d'hyperboloïdes (Q_+, Q_-) de E_n, affectés dans ce dernier cas d'un signe $(+)$ pour les rayons de Q_+ et d'un signe $(-)$ pour ceux de Q_-, est constante.

On va maintenant diagonaliser g, c'est-à-dire prendre une base e orthogonale et considérer les demi droites portant les vecteurs de base. Comme précédemment la demi-droite portant e_i coupe Q en un point $\lambda_i e_i$ avec $\lambda_i > 0$ et

$$\lambda_i^2 g_{ii} = (\text{signe de } b_{ii})\frac{g_{ii}}{b_{ii}}, \text{ soit } \frac{1}{l_i^2} = (\text{signe de } b_{ii})\frac{b_{ii}}{g_{ii}}$$

D'où : $\overset{-1}{g} : b = \sum \dfrac{b_{ii}}{g_{ii}} = \sum\limits_{i=1}^{r} \dfrac{1}{l_i^2} - \sum\limits_{j=1}^{s} \dfrac{1}{l_{r+j}^2}.$

L'indépendance vis-à-vis de la base choisie donne le :

2ème Théorème d'Apollonius. La somme des inverses des carrés des longueurs de n rayons orthogonaux d'un ellipsoïde ou d'un couple d'hyperboloïdes (Q_+, Q_-) de E_n, affectés dans ce dernier cas d'un signe $(+)$ pour les rayons de Q_+ et d'un signe $(-)$ pour ceux de Q_- est constante.

Nous terminerons ce paragraphe avec le :

3ème Théorème d'Apollonius. Le volume de n rayons conjugués d'un ellipsoïde ou d'un couple d'hyperboloïdes (Q_+, Q_-) de E_n est constant.

Preuve : Soient $e = (e_1, e_2, \ldots e_n)$ une base quelconque de E_n, $\overline{g} = \text{dét}|g_{ij}|$. Le volume euclidien de e est $\overline{g}^{1/2} = (\text{dét}|g_{ij}|)^{1/2}$. Si $(x_1, x_2, \ldots x_n)$ sont n vecteurs quelconques, on a $x_1 \wedge x_2 \wedge \ldots x_n = (\text{dét}\, X) \cdot e_1 \wedge e_2 \wedge \ldots e_n$ où $X = |x_j^i|$ est la matrice de leurs composantes, et leur volume euclidien est donc : $\|x_1 \wedge x_2 \wedge \ldots x_n\| = |\text{dét}\, X|\overline{g}^{1/2}$.

Si e est une base conjuguée relativement à q, le volume des rayons-vecteurs correspondants est, d'après les calculs précédents :

$$\begin{aligned} V &= \|\lambda_1 e_1 \wedge \lambda_2 e_2 \wedge \ldots \lambda_n e_n\| = |\lambda_1 \lambda_2 \ldots \lambda_n|\overline{g}^{1/2} \\ &= (\overline{g} \cdot \overset{-1}{\overline{b}}_{11} \overset{-1}{\overline{b}}_{22} \ldots \overset{-1}{\overline{b}}_{nn}|)^{1/2} = (\text{dét}|g_{ij}|\text{dét}|\overset{-1}{b}{}^{ij}|^{1/2}. \end{aligned}$$

Le terme entre parenthèse, produit de deux pseudoscalaires de variances $(+2)$ et (-2) et un scalaire, c.q.f.d.

VII.5 – Une incursion en mécanique classique : le tenseur ou opérateur d'inertie

Dans tout ce paragraphe où il s'agit de mécanique classique, nous nous plaçons dans l'espace euclidien E_3 qui sera simplement appelé E sans l'indice 3. Le *moment* d'un vecteur V issu d'un point M de E relativement à un point I est, par définition, le produit vectoriel $IM \times V$. Si $t \to x(t)$ est la trajectoire d'un point matériel de masse m, *l'impulsion* de ce point à l'instant t est le vecteur $p = m\frac{dx}{dt}$ produit de la masse par la vitesse, son *moment angulaire* L à l'instant t relativement à l'origine 0 est le moment de p par rapport à 0 : $L = x \times p = m \cdot x \times \frac{dx}{dt}$. A l'instant $t = t_0$, $x(t_0) = x_0$, $L_0 = m \cdot x_0 \times (\frac{dx}{dt})_0 = m\frac{d}{dt}(x_0 \times x(t))_0$. Or, soit $r(t) = ||x(t)||$.

Définissons un vecteur unitaire par $e(t) = \frac{x(t)}{||x(t)||}$, d'où $x(t) = r(t)e(t)$ et $\frac{dx}{dt} = r'(t)e(t) + r(t)\frac{de}{dt}$. On a $(e|\frac{de}{dt}) = \frac{1}{2}\frac{d}{dt}(||e(t)||^2) = 0$. Si $\omega = e \times \frac{de}{dt}$, la formule du double produit vectoriel exprime le vecteur-vitesse de $e(t)$ par : $\frac{de}{dt} = \omega \times e$. Le vecteur ω de rotation instantanée du vecteur unitaire $e(t)$ est orthogonal au plan $(e, \frac{de}{dt})$ et a pour mesure $\frac{d\theta}{dt}$, vitesse angulaire instantanée de rotation de $e(t)$. Le moment angulaire du point matériel à l'instant t s'écrit donc finalement : $L(t) = m \cdot x(t) \times \frac{dx}{dt} = mr(t)e(t) \times (r'e(t) + r\frac{de}{dt}) = mr^2 e(t) \times \frac{de}{dt} = mr^2\omega$.

Or $\frac{1}{2}x(t_0) \times x(t)$ représente, par un vecteur orthogonal à leur plan, l'aire du triangle bâti sur $x(t_0)$ et $x(t)$ ou « aire balayée » par un rayon-vecteur allant de $x(t_0)$ à $x(t)$. Sa dérivée en t_0 :

$\frac{1}{2}\mathrm{L}(t_0) = \frac{1}{2}mr^2\left(\frac{\overrightarrow{d\theta}}{dt}\right)_0$ est donc un vecteur orthogonal au plan $(x(t_0), (\frac{dx}{dt})_0)$ représentant la dérivée en t_0 de l'aire balayée par le rayon $x(t)$, ou *dérivée aréolaire* en t_0. L'« élément d'aire » est ici $\frac{1}{2}r^2 d\theta$.

Si une force F est appliquée au point matériel, le mouvement de ce dernier est déterminé par la loi de Newton $m\frac{d^2x}{dt^2} = \frac{dp}{dt} = \mathrm{F}$. Quant à son moment angulaire par rapport à 0 :

$$\frac{d\mathrm{L}}{dt} = \frac{dx}{dt} \times p + x \times \frac{dp}{dt} = x \times \mathrm{F}.$$

Si la force F est constamment portée par une droite passant par 0, donc colinéaire à x (force centrale), $\frac{d\mathrm{L}}{dt} = 0$, L = constante. Puisque $\mathrm{L}(t) = mr^2\omega$, le vecteur ω, est porté par une droite fixe. La trajectoire du point matériel est plane et s'effectue suivant la loi des aires de Képler : la dérivée aréolaire $\frac{1}{2}r^2\frac{d\theta}{dt}$ est constante et l'aire balayée par le rayon vecteur est proportionnelle au temps.

Soit maintenant un corps solide *rigide* formé par exemple d'un nombre fini de points matériels x_i de masses m_i ou d'une distribution matérielle continue de densité ρ dans un volume V qui effectue un mouvement dans E autour du point *fixe* 0. A un instant donné les vitesses s'expriment alors à l'aide du vecteur de rotation instantanée ω (§ VII.2) par : $\frac{dx}{dt} = \omega \wedge x$. Le moment angulaire en 0 du corps solide à l'instant t est défini comme la somme des moments angulaires de ses composantes soit :

$$\mathrm{L}(t) = \sum m_i x_i \times \frac{dx_i}{dt} \text{ ou } \int_{\mathrm{V}} \rho(x)(x \times \frac{dx}{dt})dv$$

ce qui s'écrit, toujours à l'aide du double produit vectoriel :

$$\begin{aligned}\mathrm{L}(t) &= \sum m_i x_i \times (\omega \times x_i) = \sum m_i\{x_i|x_i)\omega - x_i(x_i|\omega)\} \\ &= \sum m_i(||x_i||^2 - (x_i \otimes x_i) \cdot \omega = \mathrm{R}(t)\omega(t)\end{aligned}$$

ou dans le cas d'une distribution continue :

$$\mathrm{L}(t) = \int_{\mathrm{V}} \rho(x)(||x||^2 - (x \otimes x)) \cdot \omega dv = \mathrm{R}(t)\omega(t).$$

Le moment angulaire est ainsi obtenu en appliquant au vecteur rotation instantanée ω un opérateur :

$$\mathrm{R} = \sum m_i(||x_i||^2 \mathrm{I} - x_i \otimes x_i) \text{ ou } \int_{\mathrm{V}} \rho(||x||^2 \cdot \mathrm{I} - (x \otimes x))dv$$

appelé *opérateur ou tenseur d'inertie* du corps solide au point O · R est visiblement symétrique et a pour matrice dans une base orthonormale :

$$\sum m_i \begin{vmatrix} (x_i^2)^2 + (x_i^3)^2 & -x_i^1 x_i^2 & -x_i^1 x_i^3 \\ -x_i^1 x_i^2 & (x_i^1)^2 + (x_i^3)^2 & -x_i^2 x_i^3 \\ -x_i^1 x_i^3 & -x_i^2 x_i^3 & (x_i^1)^2 + (x_i^2)^2 \end{vmatrix}.$$

La sommation montre que R est évidemment une fonction additive du corps solide !

Soient D une droite issue de 0, u un vecteur unitaire de D. Le scalaire obtenu par double contraction de R avec u, soit : $\mathrm{R(D)} = u \cdot \mathrm{R} \cdot u = (u|\mathrm{R}u) = (\mathrm{R}u|u)$ est appelé le moment d'inertie du corps solide par rapport à la droite D. On peut en donner une définition simple. En effet : $u \cdot \mathrm{R} \cdot u = \sum m_i(||x_i||^2 - (x_i|u)^2) = \sum m_i r_i^2$ ou r_i est la distance du point x_i à la droite D. On a ainsi l'interprétation des termes diagonaux de la matrice de R, égaux à $e_j \cdot \mathrm{R}e_j$, $j = 1, 2, 3$.

L'énergie cinétique T du corps, en mouvement autour de 0 peut également s'exprimer à l'aide de R. En effet :

$$\begin{aligned} \mathrm{T} &= \frac{1}{2} \sum m_i \left(\frac{dx_i}{dt} | \frac{dx_i}{dt} \right) = \frac{1}{2} \sum m_i(\omega \times x_i | \omega \times x_i) \\ &= \frac{1}{2} \sum m_i(||x_i||^2 \omega^2 - (\omega|x_i)^2) = \frac{1}{2}\omega \cdot \mathrm{R} \cdot \omega = \frac{1}{2}(\omega|\mathrm{L}) \end{aligned}$$

T est la moitié du double contracté de R par la rotation instantannée ω, soit $\frac{1}{2}(\sum m_i r_i^2) \cdot \omega^2$ où les r_i sont les distances des points x_i à la droite portant ω.

Soient G le centre de gravité du corps solide et g le vecteur OG. On a $\sum m_i x_i = (\sum m_i) g = \mathrm{M}g$ et si l'on décompose chaque x_i en $x_i = g + y_i$, on a donc : $\sum m_i y_i = 0$.

Le moment d'inertie du corps solide par rapport à 0 s'écrit :

$$\begin{aligned}
\mathrm{R} = \textstyle\sum m_i\{(g+y_i|g+y_i)\mathrm{I} - (g+y_i)\otimes(g+y_i)\} \\
= \mathrm{M}(||g||^2\mathrm{I} - g\otimes g) + \textstyle\sum m_i(||y_i||^2\cdot\mathrm{I} - y_i\otimes y_i) \\
+2(g|\textstyle\sum m_i y_i) - g\otimes(\sum m_i y_i) - (\sum m_i y_i)\otimes g
\end{aligned}$$

dont il ne reste que la première ligne, soit

$$\mathrm{R}_0 = \mathrm{M}(||g||^2\mathrm{I} - g\otimes g) + \mathrm{R}_\mathrm{G}.$$

Le tenseur d'inertie R_0 d'un corps solide relatif à un point 0 est égal au tenseur d'inertie R_G de ce corps relatif à son centre de gravité G plus le tenseur d'inertie relatif à 0 d'un point matériel situé en G ayant pour masse la masse totale du corps. En contractant avec un vecteur unitaire u, on obtient un théorème analogue relatif à des axes parallèles passant par 0 et par G distants de d :

$$u\mathrm{R}_0 u = \mathrm{M}(||g||^2 - (g|u)^2) + u\mathrm{R}_\mathrm{G}u = \mathrm{M}d^2 + u\mathrm{R}_\mathrm{G}u.$$

Puisque le tenseur R_0 est symétrique, il existe une base orthogonale de E dans laquelle il est diagonal. Les droites d'une telle base sont appelées axes principaux d'inertie du corps solide en 0. La quadrique $x\cdot\mathrm{R}_0\cdot x = 1$ est un *ellipsoïde* : l'ellipsoïde d'inertie du corps au point 0. En effet, son équation dans un système d'axes principaux s'écrit :

$$\begin{aligned}
x\cdot\mathrm{R}\cdot x = (e_1\cdot\mathrm{R}\cdot e_1)(x^1)^2 \\
+(e_2\cdot\mathrm{R}\cdot e_2)(x^2)^2 + (e_3\cdot\mathrm{R}\cdot e_3)(x^3)^2 = 1 \\
= \mathrm{R}_1(x^1)^2 + \mathrm{R}_2(x^2)^2 + \mathrm{R}_3(x^3)^2
\end{aligned}$$

dont les coefficients sont les moments d'inertie du corps relatifs aux axes de coordonnées donc > 0.

Application : les équations d'Euler et le théorème de Poinsot.

Terminons ce paragraphe par quelques considérations de mécanique naïve, c'est-à-dire newtonienne, en assimilant toujours, pour la simplicité de l'écriture, notre corps solide S à un ensemble fini de points matériels M_i de masses m_i. Soit F_i la force appliquée au point M_i. l'impulsion P du corps solide S est la somme des impulsions de ses composants : $P = \sum m_i \frac{dx_i}{dt}$ et son moment angulaire (ou cinétique) par rapport à 0 est $L = \sum m_i x_i \times \frac{dx_i}{dt}$. La loi de Newton $F_i = m_i \frac{d^2 x_i}{dt^2}$, $\forall i$ entraîne les égalités :

$$\frac{dP}{dt} = \sum F_i; \ \frac{dL}{dt} = \sum x_i \times F_i$$

Si T est l'énergie cinétique du corps en mouvement :

$$T = \frac{1}{2} \sum m_i \left(\frac{dx_i}{dt} | \frac{dx_i}{dt} \right)$$

$$\text{on a } \frac{dT}{dt} = \sum m_i \left(\frac{dx_i}{dt} | \frac{d^2 x_i}{dt^2} \right) = \sum \left(\frac{dx_i}{dt} | F_i \right)$$

Appliquons ces égalités au mouvement du corps solide S autour du point fixe 0 sous l'action de forces ayant un moment résultant par rapport à 0 : $\sum x_i \times F_i$ constamment nul (il en est ainsi des forces d'interaction intérieures au solide, donc, en particulier, dans le cas de forces extérieures nulles). On a alors $\frac{dL}{dt} = 0$.

Si ω est la rotation instantanée, le moment $L = R\omega$ reste fixe.

D'autre part, puisque $\frac{dx_i}{dt} = \omega \times x_i$, $\frac{dT}{dt} = \sum \left(\frac{dx_i}{dt} | F_i \right)$ $= \sum (\omega \times x_i | F_i) = \sum (\omega | x_i \times F_i) = (\omega | \sum x_i \times F_i) = 0$: l'énergie cinétique $T = \frac{1}{2}(\omega | L)$ est constante, ce qui est équivalent au fait que la projection orthogonale de ω sur le vecteur fixe $L = R\omega$ reste fixe.

La constance de l'énergie cinétique s'exprime également par le fait que l'extrémité du vecteur ω de rotation instantanée demeure sur l'ellipsoïde fixe $\mathfrak{E}$ d'équation $\omega \cdot R \cdot \omega = 2T$ homothétique de l'ellipsoïde d'inertie. Nous avons ainsi trouvé des quantités qui restent invariantes, au cours du mouvement, ou « intégrales premières », mais il reste à décrire son évolution. Par exemple, si le vecteur moment angulaire L reste fixe dans un repère fixe, on voudrait savoir comment il se déplace dans un système d'axes liés au corps solide S ainsi d'ailleurs que le vecteur ω de rotation instantanée.

Or considérons un mouvement quelconque d'un solide S relativement à un repère fixe de E. On peut le décrire par une isométrie $t \to g(t)$, fonction du temps de telle sorte que la position du corps solide à l'instant t soit $g(t)\mathrm{S}$. Si M est un point fixement lié à S, sa position à l'instant t est $g(t)\mathrm{M}$ et sa vitesse $g'(t)\mathrm{M}$. Si maintenant M est un point mobile par rapport à S : $t \to \mathrm{M}(t)$ *dans* S, la position de M, qui est aussi entraîné par le mouvement de S, est à l'instant t : $g(t)\mathrm{M}(t)$. Sa vitesse : $g'(t)\mathrm{M}(t) + g(t)\mathrm{M}'(t)_s$ est la somme de sa vitesse d'entraînement et de sa vitesse relative. Appliquons ce théorème, au mouvement que nous étudions, en considérant un système d'axes lié à S, par exemple un système A d'axes principaux de l'ellipsoïde d'inertie de S. Il est entraîné par le mouvement de S. A l'instant t, ce système d'axes est $g(t)\mathrm{A}$. La vitesse d'entraînement d'un point M est $\omega \times \mathrm{OM}$. Les composantes de sa vitesse relative $\mathrm{M}'(t)_\mathrm{A}$ par rapport aux axes A sont aussi celles de $g(t)\mathrm{M}'(t)_\mathrm{A}$ par rapport aux axes $g(t)\mathrm{A}$.

Considéré *relativement* à S, le vecteur moment L se déplace dans le système d'axes A et sa vitesse relative est $\left(\dfrac{d\mathrm{L}}{dt}\right)_\mathrm{A}$.

L'équation d'Euler exprime, dans le système d'axes A liés au corps S, que la vitesse absolue de L est nulle, soit :

$$\left(\frac{d\mathrm{L}}{dt}\right)_\mathrm{A} + \omega \times \mathrm{L} = 0$$

Mais relativement à A, le tenseur d'inertie R est constant. Si les axes sont principaux, la matrice de R est diagonale : $\mathrm{R} = (\mathrm{R}_1, \mathrm{R}_2, \mathrm{R}_3)$. En projetant l'équation précédente *sur les axes* A, on obtient les équations d'Euler classiques :

$$\begin{aligned}
&\mathrm{R}_1 \frac{d\omega_1}{dt} + (\mathrm{R}_3 - \mathrm{R}_2)\omega_2\omega_3 = 0 \\
&\mathrm{R}_2 \frac{d\omega_2}{dt} + (\mathrm{R}_1 - \mathrm{R}_3)\omega_3\omega_1 = 0 \ (\text{avec } \mathrm{L}_j = \mathrm{R}_j\omega_j;\ j = 1,2,3) \\
&\mathrm{R}_3 \frac{d\omega_3}{dt} + (\mathrm{R}_2 - \mathrm{R}_1)\omega_1\omega_2 = 0
\end{aligned}$$

qui possèdent les deux intégrales premières : $||\mathrm{L}||^2 = \mathrm{R}_1^2\omega_1^2 + \mathrm{R}_2^2\omega_2^2 + \mathrm{R}_3^2\omega_3^2$ et $2\mathrm{T} = \mathrm{R}_1\omega_1^2 + \mathrm{R}_2\omega_2^2 + \mathrm{R}_3\omega_3^2$.

L'extrêmité de L parcourt la courbe intersection de la sphère $\sum L_j^2 = ||L||^2$ et de l'ellipsoïde $\sum \dfrac{L_j^2}{R_j} = 2T$.

Le plan tangent H à l'ellipsoïde $\mathfrak{E} : \omega \cdot R \cdot \omega = (\omega|R\omega) = 2T$ au point extrêmité de ω est précisément le plan orthogonal au vecteur $R\omega$ en son extrêmité (§ VII.2) qui est donc fixe. Puisque ω est la rotation instantanée, la vitesse du point lié à S qui, à un instant donné, coïncide avec son extrêmité, est nulle, ce qui signifie que l'ellipsoïde $\mathfrak{E}$ *roule* sur le plan H au cours du mouvement (théorème de Poinsot).

Précession : Si le corps S présente une symétrie de révolution autour de l'axe Oz, $R_1 = R_2$ et la troisième équation d'Euler montre que ω_3, projection de ω sur Oz est constant.

En multipliant la seconde équation par i et en ajoutant à la première, on obtient :

$$\frac{d(\omega_1 + i\omega_2)}{dt} = i\frac{R_3 - R_1}{R_1}(\omega_1 + i\omega_2)$$

qui prouve que, si S ne présente pas de symétrie sphérique ($R_3 \neq R_1$), la projection de ω sur l'équateur Oxy tourne à vitesse constante :

$$\omega_1 + i\omega_2 = a \exp i\frac{R_3 - R_1}{R_1}\omega t.$$

C'est le mouvement de précession.

VII.6 – Tenseurs de déformation. (« Strain tensors »)

Dans tout ce paragraphe, nous nous plaçons dans un espace euclidien E de dimension n quelconque. Soit S un corps dans E qui se déplace en fonction du temps t. En général, le corps changera de position mais aussi de forme, subissant une déformation. L'analyse géométrique du déplacement de S au voisinage d'un point consiste à séparer ce qui correspond à un mouvement de corps rigide de ce qui provient d'une déformation.

Nous rapporterons toujours l'espace E à un *repère orthonormal.* Les coordonnées du point x du corps $S(t)$ à l'instant t qui se trouvait à l'instant $t = 0$ en $x_0 = \mathrm{S}$ sont des fonctions des coordonnées de x_0 et de t, que nous supposerons différentiables : $x = \Phi(x_0; t) = \Phi_t(x_0)$ soient $x^j = x^j(x_0^1; x_0^2; \dots x_0^n; t)$ pour $j = 1, 2, \dots n$. t étant fixé, on passe de $\mathrm{S}_0 = \mathrm{S(O)}$ à $\mathrm{S}(t)$ par la transformation Φ_t. La différentielle au point $x_0 : u \to d\Phi_t(x_0; u)$ est un opérateur linéaire de E représentant en ce point la partie « du premier ordre » de Φ, soit :

$$\begin{aligned} x + \Delta x &= \Phi_t(x_0 + \Delta x_0) \\ &= \Phi_t(x_0) + d\Phi_t(x_0; \Delta x_0) + ||\Delta x_0||\mathrm{P}(x_0, \Phi x_0, t) \end{aligned}$$

où $||\mathrm{P}(x_0; \Phi x_0; t)||$ tend vers 0 avec $||\Delta x_0||$, d'où : $\Delta x = d\Phi_t(x_0; \Delta x_0)$ + un terme d'ordre > 1 en $||\Delta x_0||$.

En termes de coordonnées : $\Delta x^j = \sum \left(\dfrac{\partial x^j}{\partial x_0^k}\right)(x_0; t)\Delta x_0^k$.

La limite de l'opérateur $d\Phi_t(x_0)$ lorsque t tend vers zéro par valeurs > 0 est *l'opérateur ou tenseur de déplacement instantané en* x_0.

$$\mathrm{D} = \lim_{t \to 0} d\Phi_t(x_0) = d\Phi_0(x_0).$$

La matrice de ses composantes est : $\left| \left(\dfrac{\partial x^j}{\partial x_0^k}\right)(x_0; 0)\right|$.

Il s'agit de décrire comment se comportent les grandeurs euclidiennes au cours du déplacement. Ci-dessous nous comparerons le carré de la longueur $(\Delta x_0|\Delta x_0)$ à celle de son image $(\mathrm{D}(\Delta x_0)|\mathrm{D}(\Delta x_0)) = (\Delta x_0|\mathrm{D}^*\mathrm{D}(\Delta x_0))$. Si n est un vecteur unitaire, la racine carrée positive du scalaire $(\mathrm{D}(n)|\mathrm{D}(n)) = n \cdot \mathrm{D}^* \cdot \mathrm{D} \cdot n$, double contracté de $\mathrm{D}^*\mathrm{D}$ par n est le *coefficient d'allongement* lors du déplacement dans la direction de n. Le déterminant de D est le facteur par lequel sont multipliés les volumes. Il mesure la dilatation ou la contraction au cours de la déformation.

Lorsque le déplacement est un mouvement rigide, sans déformation, les longueurs sont conservées, autrement dit Φ_t est une isométrie de E. Φ_t s'écrit d'une façon et d'une seule comme composée d'une rotation a_t suivie d'une translation T_{z_t} : $\Phi_t(x_0 + \Delta x_0) = z_t + a_t(\Delta x_0)$. (On pourrait bien sûr choisir de faire d'abord une translation suivie d'une rotation). La différentielle de Φ est alors la rotation a_t : $\mathrm{D} = a_0 \in \mathrm{O(E)}$, $\mathrm{D}^*\mathrm{D} = \mathrm{I}$ et $\operatorname{dét} \mathrm{D} = 1$.

Dans le cas général, une séparation de D entre une partie isométrique et une partie déformation peut être obtenue à l'aide de la *décomposition polaire* des opérateurs linéaires des espaces euclidiens (cf [DR]).

Soient g un opérateur linéaire inversible de E, g^* son adjoint : $r = g^*g$ est un opérateur symétrique. Puisque $||gx||^2 = (x|g^*gx) = (x|rx)$ ses valeurs propres sont strictement positives, et il possède une « racine carrée » p. De façon précise, il existe un opérateur p, symétrique, uniquement déterminé par la condition d'avoir pour vecteurs propres ceux de $r = g^*g$ et pour valeurs propres les racines carrées positives des valeurs propres de r. Alors : $g^*g = r = p^2$. Si $a = g\overset{-1}{p}$, $a^* = \overset{-1}{p}g^*$, et $a^*a = \overset{-1}{p}p^2\overset{-1}{p} = \mathrm{I}$: a est orthogonal. On a donc décomposé g en produit ap d'un opérateur symétrique positif p et d'un opérateur orthogonal a, et cette décomposition est unique.

Soit $q = ap\overset{-1}{a}$. C'est encore un opérateur symétrique positif et $g = qa$, décomposition analogue à la précédente, mais dans l'ordre inverse, et, comme elle, unique. Remarquons que $q^2 = gg^*$ et que p et q ont mêmes valeurs propres.

Définition VII.6

Si D est l'opérateur de déplacement instantané en x_0, les *opérateurs ou tenseurs de déformation* associés sont les opérateurs symétriques positifs $\mathrm{R} = \mathrm{P}^2 = \mathrm{D}^*\mathrm{D}$ et $\mathrm{S} = \mathrm{Q}^2 = \mathrm{DD}^*$. Le déplacement instantané D en x_0 peut alors être considéré comme le produit d'une déformation pure d'opérateur symétrique P suivie d'une rotation $\mathrm{A} : \mathrm{D} = \mathrm{AP}$ ou en changeant l'ordre des opérateurs, comme une rotation A suivie d'une déformation pure d'opérateur symétrique $\mathrm{Q} = \mathrm{AP}\overset{-1}{\mathrm{A}} : \mathrm{D} = \mathrm{QA}$. On appelle *invariants de déformation en* x_0, les valeurs propres, qui sont strictement positives, de P ou de Q ou coefficients d'allongement principaux et plus généralement des fonctions de ces valeurs propres, comme par exemple leurs carrés, valeurs propres de $\mathrm{R} = \mathrm{D}^*\mathrm{D}$ et de $\mathrm{S} = \mathrm{DD}^*$ ou leurs diverses fonctions symétriques. Les axes principaux du tenseur symétrique P ou du tenseur symétrique Q sont des axes principaux de déformation en x_0.

Dans une base orthonormale de E, la matrice de D^* est la transposée de celle de D. Les matrices de $\mathrm{R} = \mathrm{D}^*\mathrm{D}$ et $\mathrm{S} = \mathrm{DD}^*$

sont donc :

$$\mathrm{R}_{jk} = \sum_i \frac{\partial x^i}{\partial x_0^j}\frac{\partial x^i}{\partial x_0^k} \qquad \mathrm{S}_{jk} = \sum_i \frac{\partial x^j}{\partial x_0^i}\frac{\partial x^k}{\partial x_0^i}.$$

Cependant, en théorie de l'élasticité, on n'utilise guère les variables x^i. En effet la transformation $x = \Phi_t(x_0)$ se réduit à l'identité pour $t = 0$, et Φ_t est plus clairement analysée en considérant au lieu de Φ_t, la différence $\Phi_t - \mathrm{I}$, soit : $u = x - x_0 = (\Phi_t - \mathrm{I})(x_0)$. On remplace alors l'opérateur D par (D − I) et $\mathrm{R} = \mathrm{D}^*\mathrm{D}$ par *l'opérateur, ou tenseur de déformation de Lagrange* : $\Gamma = \frac{1}{2}(\mathrm{D}^*\mathrm{D} - \mathrm{I})$. Ainsi qu'on l'a vu ci-dessus, on a : $(\mathrm{D}(\Delta x)|\mathrm{D}(\Delta x)) - (\Delta x|\Delta x) = (\mathrm{P}(\Delta x)|\mathrm{P}(\Delta x)) - (\Delta x|\Delta x) = ((\mathrm{D}^*\mathrm{D} - \mathrm{I})\Delta x|\Delta x)$ soit : $\frac{1}{2}\Delta x \cdot \Gamma \cdot \Delta x = \frac{1}{2}(||\mathrm{D}(\Delta x)||^2 - ||\Delta x||^2)$. En un point x_0, 2Γ est donc le tenseur de déformation des carrés des longueurs. Nous allons calculer ses composantes dans une base orthonormale en fonction des nouvelles variables $u^j = x^j - x_0^j$.

On a $\mathrm{D}_{ij} = \dfrac{\partial x^i}{\partial x_0^j} = \dfrac{\partial u^i}{\partial x_0^j} + \delta_{ij}$.

$$\Gamma = \frac{1}{2}(\mathrm{R}_{jk} - \delta_{jk}) = \frac{1}{2}\sum_i \left(\frac{\partial u^i}{\partial x_0^j} + \delta_{ij}\right)\left(\frac{\partial u^i}{\partial x_0^k} + \delta_{ik}\right) - \delta_{jk}$$
$$= \frac{1}{2}\left(\frac{\partial u^k}{\partial x_0^j} + \frac{\partial u^j}{\partial x_0^k} + \sum_i \frac{\partial u^i}{\partial x_0^j}\frac{\partial u^i}{\partial x_0^k}\right).$$

Pour les déformations de très faible amplitude qui sont le domaine de l'élasticité linéaire, on néglige la somme du second membre ci-dessus pour ne considérer que le *tenseur de déformation infinitésimale* E de composantes :

$$\mathrm{E}_{jk} = \frac{1}{2}\left(\frac{\partial u^j}{\partial x_0^k} + \frac{\partial u^k}{\partial x_0^j}\right).$$

Ce tenseur s'exprime directement à partir de l'opérateur de déplacement instantané par : $\mathrm{E} = \frac{1}{2}(\mathrm{D} + \mathrm{D}^*) - \mathrm{I}$. C'est le symétrisé de $\mathrm{D}-\mathrm{I} : \mathrm{E} = \frac{1}{2}(\mathrm{D}-\mathrm{I}) + \frac{1}{2}(\mathrm{D}-\mathrm{I})^*$. L'antisymétrisé de $\mathrm{D}-\mathrm{I}$: $\Omega = \frac{1}{2}(\mathrm{D}-\mathrm{D}^*)$ est un opérateur antisymétrique donc une rotation instantanée, ce qui donne la décomposition, additive, cette fois, de $\mathrm{D}-\mathrm{I}$ en une partie de déformation E et une partie de rotation $\Omega : \mathrm{D}-\mathrm{I} = \mathrm{E}+\Omega$.

Enfin, le déterminant de D s'écrit avec les variables $u^i = x^i - x_{i;0}$; $\text{dét}\,\mathrm{D} = \text{dét}|\delta_{ij} + \frac{\partial u^i}{\partial x_0^j}| = 1 + \sum \frac{\partial u^i}{\partial x_0^j} +$ des termes d'ordre supérieur.

Le scalaire : $\mathrm{T}r\,\mathrm{E} = \sum \mathrm{E}_{ii} = \sum \frac{\partial u^i}{\partial x_0^i}$ est la *dilatation infinitésimale* en x_0.

VII.7 – Tenseur des contraintes (« Stress tensor »). Loi de Hookes

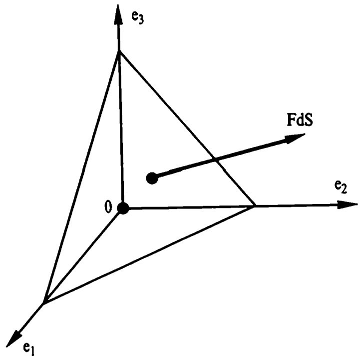
On a vu rapidement comme exemple, au § II.2 (et sans justification) une définition du tenseur des contraintes.

Nous allons maintenant procéder à une étude plus précise. Soit V un élément de volume régulier, limité par une surface S, d'un corps, solide ou liquide, en équilibre statique ou dynamique.

Les forces agissant sur V sont de deux espèces : les forces « *volumiques* » : poids, forces d'inertie, etc..., et les forces *superficielles* d'origine interne, exercées sur la surface S de V par le reste du corps dont il fait partie ou dans lequel il est plongé : pression, tensions superficielles, etc... Soit $d\mathrm{S}$ un élément de surface *orienté* par le choix d'un vecteur normal n dirigé vers l'extérieur du volume V. La force exercée sur $d\mathrm{S}$ par la partie extérieure du corps peut être représentée par un vecteur $\mathrm{F}d\mathrm{S}$: F est la force par unité de surface. La composante normale de F correspond à une pression ou à une tension, la composante de F sur le plan tangent à $d\mathrm{S}$ correspond à un cisaillement (« shear stress »).

Les deux raisonnements classiques qui vont suivre (« du tétraèdre » et du « cube ») vont simplement exprimer que la somme des forces et la somme des moments des forces agissant sur V sont nulles, comme conséquence de l'équilibre supposé.

Considérons un tétraèdre élémentaire τ dont trois côtés sont portés par les axes d'un repère orthonormé (e_1, e_2, e_3) d'origine 0.



Si dS est la surface du triangle τ_0 opposé à 0, orientée par le choix d'un vecteur unitaire normal n dirigé vers l'extérieur de τ, l'aire du triangle τj, orthogonal à e_j est la projection orthogonale de l'aire de τ_0. Sa valeur est : $(n|e_j)d\mathrm{S} = n^j d\mathrm{S}$. Mais comme les vecteurs unitaires du repère e_1, e_2, e_3 sont orientés vers l'intérieur et non vers l'extérieur de τ, l'élément de surface orienté par $(-e_j)$ du triangle τ_j a pour mesure algébrique $d\mathrm{S}_j = n^j d\mathrm{S}$.

Soit $\mathrm{F}_j d\mathrm{S}_j$ la force s'exerçant sur la surface du triangle τ_j, représentant l'action sur l'élément de volume dV du tétraèdre τ de la partie du corps située à l'extérieur de la face τ_j. En équilibre, statique ou dynamique, la somme de toutes les forces qui s'exercent sur dV est nulle :

$$\Phi d\mathrm{V} + \mathrm{F}d\mathrm{S} + \mathrm{F}_1 d\mathrm{S}_1 + \mathrm{F}_2 d\mathrm{S}_2 + \mathrm{F}_3 d\mathrm{S}_3 = 0$$

où $\Phi d\mathrm{V}$ est la somme des forces « volumiques », d'où : $\Phi d\mathrm{V} + (\mathrm{F} - \mathrm{F}_1 n^1 - \mathrm{F}_2 n^2 - \mathrm{F}_3 n^3)d\mathrm{S} = 0$.

Lorsqu'on fait tendre le tétraèdre vers un point, les aires tendent vers zéro comme le carré des longueurs, les volumes comme le cube. On a donc nécessairement : $\mathrm{F} = \mathrm{F}_1 n^1 + \mathrm{F}_2 n^2 + \mathrm{F}_3 n^3$, quelle que soit la face τ_0.

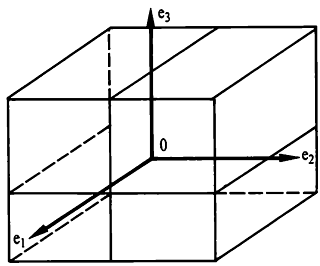
En projetant sur chacun des axes de coordonnées :

$$\mathrm{F}^j = \mathrm{F}^j_1 n^1 + \mathrm{F}^j_2 n^2 + \mathrm{F}^j_3 n^3.$$

Nous avons ainsi démontré que les forces superficielles en un point d'un corps en équilibre sont déterminées par un tenseur du second ordre, *le tenseur des contraintes*, qui s'exprime par un opérateur $(n \rightarrow \mathrm{T}(n))$: à chaque élément de surface orienté

dS de vecteur unitaire normal n, T fait correspondre la force FdS $=$ T$(n)d \cdot$ S définie par : $\mathrm{F}^j = \mathrm{T}(n)^j = \mathrm{F}_1^j n^1 + \mathrm{F}_2^j n^2 + \mathrm{F}_3^j n^3$.

La nullité du moment des forces agissant sur un élément de volume va maintenant avoir pour conséquence la symétrie du tenseur des contraintes T. Considérons pour cela un élément de volume ayant la forme du cube de côté 2ε. Choisissons un repère orthonormé (e_1, e_2, e_3) ayant pour origine 0 le centre du cube, d'axes Ox, Oy, Oz parallèles aux arêtes. Parmi les forces agissant sur le cube, les forces « volumiques » et les forces de contraintes normales aux faces ont un moment nul par rapport à O.



Par projection orthogonale sur Oz du moment en O des forces, toutes celles qui sont colinéaires avec Oz donnent évidemment zéro, en particulier les forces agissant sur les faces « horizontales » $z = \pm\varepsilon$. Ne peuvent donner une projection non nulle sur Oz que les moments des forces superficielles « horizontales », projections sur le plan (e_1, e_2) de $\mathrm{T}(e_1)_+$ et $\mathrm{T}(e_1)_-$ agissant sur les faces $x = \pm\varepsilon$, $\mathrm{T}(e_2)_+$ et $\mathrm{T}(e_2)_-$ agissant sur $y = \pm\varepsilon$.

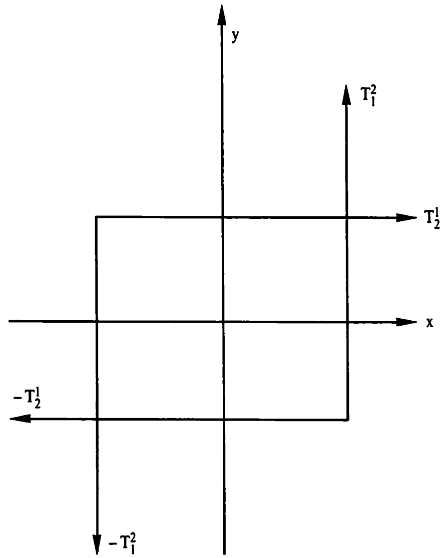
Or, puisque la somme de toutes les forces agissant sur le cube est nulle, on a :

$$\mathrm{T}(e_1)_+^2 = -\mathrm{T}(e_1)_-^2 \text{ et } \mathrm{T}(e_2)_+^1 = -\mathrm{T}(e_2)_-^1.$$

La nullité de la projection du moment sur Oz donne finalement :

$$2\varepsilon \mathrm{T}(e_1)^2 - 2\varepsilon \mathrm{T}(e_2)^1 = 0 \text{ soit } \mathrm{T}_1^2 = \mathrm{T}_2^1.$$

Le même raisonnement pouvant être effectué pour les axes Ox et Oy, on a ainsi démontré que le tenseur des contraintes est symétrique : $\mathrm{T}_{j,i} = \mathrm{T}_{i,j},\ \forall i, j$.



Terminons ce paragraphe par ce qu'il est d'usage d'appeler la « loi de Hookes » : dans un corps élastique en équilibre statique ou dynamique, il existe entre le tenseur des contraintes T et le tenseur de déformation infinitésimale E, une relation linéaire qui s'exprime par un tenseur C d'ordre quatre, appelé tenseur d'élasticité du corps :

$$\mathrm{E}_{ij} = \sum_{k,l} \mathrm{C}_{ijkl}\mathrm{T}_{kl}.$$

VII.8 – Courbes et surfaces dans les espaces euclidiens

Les propriétés classiques des courbes et des surfaces servent d'exemples et de préliminaires à l'étude ultérieure des variétés riemanniennes.

Rappelons d'abord quelques définitions :

– un *chemin* d'un espace métrique E est une application continue : $t \to c(t)$ d'un intervalle (a, b) de $\mathbb{R}$ dans E.

Si φ est une application continue strictement croissante ou strictement décroissante de (a_1, b_1) sur (a, b), le chemin : $u \in a_1, b_1) \to c_1(u) = c(\varphi(u))$ est dit $\mathfrak{C}_0$-équivalent à c.

On a $c(t) = c_1(\overset{-1}{\varphi}(t))$ et c_1 est dit de même orientation ou d'orientation opposée à c selon que φ est croissante ou décroissante.

Si φ est r fois continûment différentiable, avec $\varphi' > 0$, ou $\varphi' < 0$, c_1 est dit $\mathfrak{C}_r$-équivalent à $\mathfrak{C}$.

– une *courbe* C de classe $\mathfrak{C}_r$ de l'espace euclidien E_n est une classe de $\mathfrak{C}_r$-équivalence de chemins r fois continûment différentiables.

Un chemin de la classe est appelé un paramétrage de la courbe. On le note : $t \to \mathrm{M}(t)$. La courbe C est dite *simple* si l'un des paramétrages est bijectif, auquel cas ils le sont tous.

Reprenons le cas général d'un chemin : $t \to c(t)$ d'un espace métrique E, défini sur un intervalle fermé borné $[a, b]$ de $\mathbb{R}$. A chaque subdivision σ : $[a = t_0 < t_1 < \ldots < t_n < t_{n+1} = b]$ associons le nombre $l(\sigma) = \sum\limits_{i=0}^{m} d(c(t_i); c(t_{i+1}))$. Si les $l(\sigma)$ ont une borne supérieure finie l, le chemin est dit rectifiable de longueur l. Un théorème analogue au théorème de Darboux montre alors que les $l(\sigma)$ tendent vers leur borne supérieure lorsque $\mathrm{Sup}_{,i} |t_{i+1} - t_i|$ tend vers zéro. Il en résulte que toute partie d'un chemin rectifiable est rectifiable et que la longueur est additive. En choisissant $t_0 \in [a, b]$ d'image $\mathrm{M}_0 = c(t_0)$, on appelle *abscisse curviligne* d'origine M_0 sur le chemin, orienté dans le sens des t croissants, la fonction $s(t) = l(c[t_0, t])$ si $t \geqslant t_0$, $s(t) = -l(c[t_0, t])$ si $t < t_0$ d'où : $l(c|t_1, t_2]) = |s(t_2) - s(t_1)|$.

La fonction $(-s(t))$ est l'abscisse curviligne d'origine M_0 du chemin orienté dans le sens opposé. La fonction s est une application continue croissante de $[a, b]$ sur un intervalle de $\mathbb{R}$.

Si E est un espace euclidien E_n et si le chemin : $t \in [a, b] \to c(t)$ est continûment différentiable, le vecteur dérivé : $c'(t_0) = \left(\dfrac{dc}{dt}\right)_0 = \lim\limits_{h \to 0} \dfrac{1}{h}\{c(t_0 + h) - c(t_0)\}$ est le vecteur tangent au chemin en t_0. Si φ définit un paramétrage $\mathfrak{C}_1$-équivalent : $t = \varphi(u)$, $c_1(u) = c(\varphi(u))$, on a $\dfrac{dc_1}{du} = \dfrac{dc}{dt}\dfrac{dt}{du}$. La droite tangente en M_0 est indépendante du paramétrage et ne dépend que de la courbe C que définit le chemin. Rappelons qu'un chemin continûment différentiable de E_n est rectifiable et que sa longueur est donnée par l'intégrale $\int_a^b ||c'(t)||dt$. En effet, en prenant une base orthonormée de E_n, et une subdivision σ de $[a,b]$ la formule des accroissements finis donne :

$$l(\sigma) = \sum_{i=0}^{m} \left[\sum_{k=1}^{n} (x^k(t_{i+1}) - x^k(t_i))^2 \right] 1/2$$
$$= \sum_{i=0}^{m} (t_{i+1} - t_i) \left[\sum_{k=1}^{n} x'^k(\theta_i^k)^2 \right]^{1/2}$$

Chaque fonction continue x'^k et la fonction $[\sum(x'^k)^2]^{1/2}$ étant intégrables on en déduit que les $l(\sigma)$ ont pour limite l'intégrale : $\int_a^b \left[\sum_{k=1}^{n} x'^k(t)^2 \right]^{1/2} dt = \int_a^b ||c'(t)||dt$. Un paramétrage $\mathfrak{C}_1$-équivalent revient à un changement de variable dans l'intégrale et lui donne donc la même valeur, ce qui était géométriquement évident. La longueur ne dépend par conséquent que de la courbe C définie par le chemin, ainsi que l'abscisse curviligne, une fois l'origine fixée.

Pour un paramétrage : $t \longrightarrow \mathrm{M}(t)$ avec $\mathrm{M}_0 = \mathrm{M}(t_0)$, $\mathrm{M} = \mathrm{M}(t_1) : s(\mathrm{M}) = \int_{t_1,t_0} ||\mathrm{M}'(t)||dt$

La fonction $s(t) = \int_{t,t_0} ||\mathrm{M}'(t)||dt$ est continûment dérivable de dérivée $\geqslant 0$: $\dfrac{ds}{dt} = ||\mathrm{M}'(t)||$. Sa différentielle $ds = ||\mathrm{M}'(t)||dt$ est *l'élément d'arc du chemin*, soit : $ds = \left[\sum_{k=1}^{n} x'^k(t)^2 \right]^{1/2} dt$.

La courbe C définie par un chemin continûment différentiable : $t \longrightarrow \mathrm{M}(t)$ est dite *régulière* si quel que soit t, le vecteur tangent $\mathrm{M}'(t)$ est non nul. La fonction $t \longrightarrow s(t)$ a dans ce cas une dérivée $||\mathrm{M}'(t)||$ strictement positive ainsi que la fonction inverse $s \longrightarrow t(s)$ qui est également continument différentiable. On obtient ainsi, parmi les paramétrages $\mathfrak{C}_1$-équivalents à $t \longrightarrow \mathrm{M}(t)$, un paramétrage canonique de C par l'abscisse curviligne, une fois l'origine choisie : $s \longrightarrow \mathrm{M}(s) = \mathrm{M}(t(s))$.

Le vecteur tangent correspondant à ce paramétrage est $\mathrm{M}'(s) = \mathrm{M}'(t)\dfrac{dt}{ds} = \dfrac{\mathrm{M}'(t)}{||\mathrm{M}(t)||}$ d'où $||\mathrm{M}'(s)|| = 1$.

$u(s) = \mathrm{M}'(s)$ est le vecteur unitaire de la tangente orientée à C dans le sens des s croissants. L'extrêmité du vecteur $u(s)$ décrit sur la sphère unité de E_n un arc de courbe qui est l'indicatrice de C. Si l'on suppose C de classe $\mathfrak{C}_2$, le vecteur dérivé $u'(s) = \mathrm{M}''(s)$ est orthogonal à $u(s)$ puisque $(u(s)|u(s)) = 1$. Sa norme $\rho = ||u'(s)||$, qui est une mesure de la rapidité avec laquelle $u(s)$ change de direction, est la courbure de C au point M. Si $\rho \neq 0$, $\mathrm{R} = \dfrac{1}{\rho}$ est le rayon de courbure en ce point, la droite portant $\mathrm{M}''(s) = u'(s)$ est la *normale principale*, orientée par le vecteur unitaire $n = \dfrac{u'(s)}{||u(s)||}$.

On a donc : $\mathrm{M}''(s) = u'(s) = \rho n = \dfrac{n}{\mathrm{R}}$. Le plan (u, x) est dit *osculateur* en M à C. En revenant à un paramètrage quelconque $t \to \mathrm{M}(t)$, on a $\mathrm{M}'(t) = \mathrm{M}'(s)\dfrac{ds}{dt}$ avec $|\dfrac{ds}{dt}| = ||\mathrm{M}'(t)||$, $\mathrm{M}''(t) = u'(s)\left(\dfrac{ds}{st}\right)^2 + u(s)\dfrac{d^2s}{dt^2}$. Dans $\wedge^2\mathrm{E}_n$, le bivecteur $\mathrm{M}'(t) \wedge \mathrm{M}''(t) = u \wedge u'\left(\dfrac{ds}{dt}\right)^3 = \dfrac{1}{\mathrm{R}}u \wedge n\left(\dfrac{ds}{dt}\right)^3$.

En prenant les normes, il en résulte que :

$$\mathrm{R} = \frac{||\mathrm{M}'(t)||^3}{||\mathrm{M}'(t) \wedge \mathrm{M}''(t)||}.$$

En particulier dans l'espace euclidien de dimension trois :

$$\begin{aligned}\mathrm{R} &= \frac{||\mathrm{M}'(t)||^3}{||\mathrm{M}'(t) \times \mathrm{M}''(t)||} \\ &= \frac{(x'^2 + y'^2 + z'^2)^{3/2}}{[(y'z'' - z'y'')^2 + (z'x'' - x'z'')^2 + (x'y'' - y'x'')^2]^{1/2}}.\end{aligned}$$

Si la courbe est de classe $\geqslant 3$, on peut poursuivre la dérivation afin d'obtenir des invariants géométriques d'ordre supérieur au second. Si $n'(s)$ n'appartient pas au plan (u, n), il détermine, au signe près, un vecteur unitaire b par la condition que b soit orthogonal au plan (u, n). La droite portant b est la *binormale*. Lorsque l'espace est de dimension trois, on a un choix naturel de b par la condition que (u, n, b) soit positivement orienté soit : $b = u \times n$. Dès lors, puisque n' est orthogonal à n, on a : $n' = \lambda u + \mu b$.

De $(u|n) = 0$, il résulte $(u'|n) + (u|n') = \frac{1}{\mathrm{R}} + \lambda = 0$ et $\lambda = -\frac{1}{\mathrm{R}}$. le vecteur b étant choisi, on peut par analogie, écrire μ sous la forme $\mu = \frac{1}{\mathrm{T}}$: μ est la torsion, et T le rayon de torsion. En poursuivant la dérivation, b' est naturellement orthogonal à b et s'écrit : $b' = \alpha u + \beta n + e$, où e est orthogonal au sous-espace (u, n, b) nul dans le cas de E_3. De $(b|u) = 0$ et $(b|n) = 0$, il résulte que $\alpha = 0$ et $\beta = -\frac{1}{\mathrm{T}}$. Les trois égalités obtenues ont été écrites pour la première fois par Serret et Frenet (1851). Dans E_3 :

$$\frac{du}{ds} = \frac{n}{\mathrm{R}};\ \frac{dn}{ds} = -\frac{u}{\mathrm{R}} + \frac{b}{\mathrm{T}};\ \frac{db}{ds} = -\frac{n}{\mathrm{T}}$$

qui exprime que la rotation instantanée, par rapport à s, du trièdre (u, n, b) est $\omega = \frac{u}{\mathrm{T}} + \frac{b}{\mathrm{R}}$.

Nous allons maintenant examiner la notion de surface. Une application r fois continûment différentiable, $r \geqslant 3$, d'un domaine connexe D de R^2 dans un espace vectoriel réel E est dite *régulière* si quel que soit $(u^1, u^2) \in \mathrm{D}$, les vecteurs $\frac{\partial \mathrm{M}}{\partial u^1}$ et $\frac{\partial \mathrm{M}}{\partial u^2}$ sont linéairement indépendants. Cette condition, si $\mathrm{E} = \mathrm{R}^2$, s'exprime par la non-nullité du Jacobien :

$$\frac{\mathrm{D}(x^1, x^2)}{\mathrm{D}(u^1, u^2)} = \text{dét} \begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \frac{\partial x^1}{\partial u^2} \\ \frac{\partial x^2}{\partial u^1} & \frac{\partial x^2}{\partial u^2} \end{vmatrix}.$$

Dans ce cas particulier, puisque le jacobien est une fonction continue à valeurs réelles non nulle sur le domaine connexe D, c'est qu'il est partout, ou bien strictement positif, et l'application conserve alors l'orientation, ou bien strictement négatif, et l'application change l'orientation en son opposée.

Un élément de surface S de E_n est déterminé par une application $\mathfrak{C}_r$, $r \geqslant 3$, régulière, d'un domaine connexe D de $\mathbb{R}^2$ dans E_n : $(u^1, u^2) \to (\mathrm{M}(u^1, u^2)$, appelé *système de coordonnées de* S. Toute application φ d'un domaine D' de $\mathbb{R}^2$ sur D : $(v^1, v^2) \to (\varphi^1(v^1, v^2), \varphi^2(v^1, v^2)$ bijective, de classe $r \geqslant 3$, régulière, définit par composition :

$$(v^1 v^2) \to \mathrm{M}(\varphi^1(v^1, v^2), \varphi^2(v^1, v^2))$$

un nouveau système de coordonnées de S dit *équivalent* au premier, de même orientation si $\dfrac{\mathrm{D}(\varphi^1, \varphi^2)}{\mathrm{D}(v^1, v^2)} > 0$, d'orientation opposée si ce jacobien est négatif.

Aucun système de coordonnées ne devant être privilégié dans sa définition, un élément de surface S peut être considéré comme une classe d'équivalence de systèmes de coordonnées. Puisque :

$$\frac{\partial \mathrm{M}}{\partial v^1} = \frac{\partial \mathrm{M}}{\partial u^1}\frac{\partial u^1}{\partial v^1} + \frac{\partial \mathrm{M}}{\partial u^2}\frac{\partial u^2}{\partial v^1} \text{ et } \frac{\partial \mathrm{M}}{\partial v^2} = \frac{\partial \mathrm{M}}{\partial v^1}\frac{\partial u^1}{\partial v^2} + \frac{\partial \mathrm{M}}{\partial u^2}\frac{\partial u^2}{\partial v^2}$$

le plan défini par les vecteurs $\dfrac{\partial \mathrm{M}}{\partial u^1}$ et $\dfrac{\partial \mathrm{M}}{\partial u^2}$ est indépendant du choix du système de coordonnées : c'est le *plan tangent* $\mathrm{T_M}$ *en* M *à* S. L'ordre des vecteurs $\dfrac{\partial \mathrm{M}}{\partial u^1}, \dfrac{\partial \mathrm{M}}{\partial u^2}$ définit une orientation de ce plan, conservée par tous les changements de coordonnées à jacobien positif.

Tout système de coordonnées détermine ainsi une base $\left(\dfrac{\partial \mathrm{M}}{\partial u^1}, \dfrac{\partial \mathrm{M}}{\partial u^2}\right)$ du plan tangent $\mathrm{T_M}$ et tout changement de coordonnées détermine un changement de base linéaire :

$$\left(\frac{\partial \mathrm{M}}{\partial v^1}, \frac{\partial \mathrm{M}}{\partial v^2}\right) = \left(\frac{\partial \mathrm{M}}{\partial u^1}, \frac{\partial \mathrm{M}}{\partial u^2}\right) \begin{vmatrix} \dfrac{\partial u^1}{\partial v^1} & \dfrac{\partial u^1}{\partial v^2} \\ \dfrac{\partial u^2}{\partial v^1} & \dfrac{\partial u^2}{\partial v^2} \end{vmatrix},$$

dont il est très commode d'écrire la matrice inverse puisque cela revient à échanger les u et le v.

Les fonctions coordonnées sur $\mathrm{T_M}$ relatives à la base $\left(\dfrac{\partial \mathrm{M}}{\partial u^1}, \dfrac{\partial \mathrm{M}}{\partial u^2}\right)$ sont les différentielles (du^1, du^2). Elles forment la base duale de l'espace cotangent $\mathrm{T^*_M}$. La matrice de passage des coordonnées (du^1, du^2) aux coordonnées (dv^1, dv^2) est contragrédiente de la matrice de changement de base (cf. II.2).

Une courbe située sur l'élément de surface S est définie dans un système de coordonnées, par les coordonnées de ses points, c'est-à-dire par une application : $t \in [a,b] \subset \mathbb{R} \to (u^1(t), u^2(t)] \in \mathrm{D} \to \mathrm{M}(u^1(t), u^2(t)) \in \mathrm{S}$. Nous supposons toujours que les courbes sont r fois continûment différentiables, $r \geqslant 3$.

Le vecteur tangent à une telle courbe en un point M :

$$\frac{d\mathrm{M}}{dt} = \frac{\partial \mathrm{M}}{\partial u^1}\frac{du^1}{dt} + \frac{\partial \mathrm{M}}{\partial u^2}\frac{du^2}{dt}$$

appartient au plan tangent $\mathrm{T_M}$ de S en ce point.

Son élément d'arc :

$$\begin{aligned} ds &= ||\mathrm{M}'(t)||dt \\ &= \left(\frac{\partial \mathrm{M}}{\partial u^1}\frac{du^1}{dt} + \frac{\partial \mathrm{M}}{\partial u^2}\frac{du^2}{dt}\middle| \frac{\partial \mathrm{M}}{\partial u^1}\frac{du^1}{dt} + \frac{\partial \mathrm{M}}{\partial u^2}\frac{du^2}{dt}\right)^{1/2} dt \\ &= \left[\left\|\frac{\partial \mathrm{M}}{\partial u^1}\right\|^2 + 2\left(\frac{\partial \mathrm{M}}{\partial u^1}\middle|\frac{\partial \mathrm{M}}{\partial u^2}\right) du^1 du^2 + \left\|\frac{\partial \mathrm{M}}{\partial u^2}\right\|^2 (du^2)^2\right]^{1/2} \end{aligned}$$

fait apparaître la forme quadratique sur le plan tangent $\mathrm{T_M}$:

$$ds^2 = \sum g_{ij} du^i du^j = ||d\mathrm{M}||^2 \text{ où } g_{ij} = \left(\frac{\partial \mathrm{M}}{\partial u^i}\middle|\frac{\partial \mathrm{M}}{\partial u^j}\right)$$

Elle est appelée la métrique de S, ou « première forme fondamentale ».

C'est évidemment la restriction à $\mathrm{T_M}$ de la métrique euclidienne de E_n et elle définit sur $\mathrm{T_M}$ une structure d'espace euclidien.

Définition VII.8.A.

Les tenseurs attachés à l'espace euclidien tangent $\mathrm{T_M}$ à la surface S au point M, sont appelés les *tenseurs sur* S *en* M.

$$\text{Notons } \overline{g}_u = \left\|\frac{\partial \mathrm{M}}{\partial u^1} \wedge \frac{\partial \mathrm{M}}{\partial u^2}\right\|^2 = g_{11}g_{22} - g_{12}^2$$

$\overline{g}_u$ est le carré de l'aire de la base $\left(\frac{\partial \mathrm{M}}{\partial u^1}, \frac{\partial \mathrm{M}}{\partial u^2}\right)$ de $\mathrm{T_M}$. La 2-forme extérieure sur $\mathrm{T_M}$: $\overline{g}_u^{1/2} du^1 \wedge du^2 = \left\|\frac{\partial \mathrm{M}}{\partial u^1} \wedge \frac{\partial \mathrm{M}}{\partial u^2}\right\| du^1 \wedge du^2$

est un invariant de la surface orientée. Par un changement de coordonnées conservant l'orientation (jacobien > 0), on a en effet :

$$\overline{g}_v^{1/2} dv^1 \wedge dv^2 = \overline{g}_u^{\frac{1}{2}} du^1 \wedge du^2.$$

C'est *l'élément d'aire de* S.

Etude particulière des surfaces de E_3. Soit S un élément de surface de E_3, défini et orienté par des coordonnées (u^1, u^2). Soit $N = \dfrac{\dfrac{\partial M}{\partial u^1} \times \dfrac{\partial M}{\partial u^2}}{\left\| \dfrac{\partial M}{\partial u^1} \times \dfrac{\partial M}{\partial u^2} \right\|}$ le vecteur unitaire normal à S en M. Les vecteurs N forment un champ de vecteurs sur S. Si $\dfrac{dM}{dt}$ est le vecteur tangent d'une courbe régulière C de S en un point M, la droite orientée du plan tangent T_M définie par le produit vectoriel $N \times \dfrac{dM}{dt}$, est appelée la *normale géodésique de* C. Si l'on paramètre la courbe C à l'aide de l'arc, et si ρ est la courbure de C en M la normale principale de C appartient au plan orthogonal à $\dfrac{dM}{ds}$ et on a : $\dfrac{d^2M}{ds^2} = \rho n = \rho_n N + \rho_g N \times \dfrac{dM}{ds}$.

La composante ρ_n est appelée la courbure normale, et ρ_g est la courbure géodésique de C en M. On a donc : $\rho_n = \left(N \middle| \dfrac{d^2M}{ds^2} \right)$.

Revenons à un paramètrage quelconque pour C. Le vecteur accélération $\dfrac{d^2M}{dt^2}$ se décompose en :

$$M''_{t^2} = \frac{\partial^2 M}{(\partial u^1)^2} \left(\frac{du^1}{dt} \right)^2 + 2 \frac{\partial^2 M}{\partial u^1 \partial u^2} \frac{du^1}{dt} \frac{du^2}{dt} + \frac{\partial^2 M}{(\partial u^2)^2} \left(\frac{du^2}{dt} \right)^2 + \frac{\partial M}{\partial u^1} \frac{d^2 u^1}{dt^2} + \frac{\partial M}{\partial u^2} \frac{d^2 u^2}{dt^2}.$$

La seconde partie appartient évidemment au plan tangent T_M. Nous allons examiner la décomposition entre T_M et N de la première, et pour cela écrire les équations dites de Gauss :

$$\boxed{\frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j} = \mathrm{L}_{ij}\mathrm{N} + \sum \Gamma^k_{ij} \frac{\partial \mathrm{M}}{\partial u^k}} \text{ avec } \mathrm{L}_{ij} = \mathrm{L}_{ji} \text{ et } \Gamma^k_{ij} = \Gamma^k_{ji}.$$

La composante normale de M''_{t^2} s'écrit donc :

$$(\mathrm{N}|\mathrm{M}''_{t^2}) = \sum \mathrm{L}_{ij} \frac{du^i}{dt}\frac{du^j}{dt}.$$

La forme quadratique sur le plan tangent : $\sum \mathrm{L}_{ij} du^i du^j$ est un invariant, d'après sa définition. Elle est appelée la *seconde forme fondamentale de* S. Si l'on paramètre C par l'arc :

$$\frac{d\mathrm{M}}{ds} = \frac{d\mathrm{M}}{dt}\frac{dt}{ds}, \ \frac{d^2\mathrm{M}}{ds^2} = \frac{d^2\mathrm{M}}{dt^2}\left(\frac{dt}{ds}\right)^2 + \frac{d\mathrm{M}}{dt}\frac{d^2t}{ds^2} \text{ d'où}$$

$$\rho_n = \left(\mathrm{N}\left|\frac{d^2\mathrm{M}}{ds^2}\right.\right) = \left(\mathrm{N}\left|\frac{d^2\mathrm{M}}{dt^2}\right.\right)\left(\frac{dt}{ds}\right)^2 = \frac{\sum \mathrm{L}_{ij}\frac{du^i}{dt}\frac{du^j}{dt}}{\sum g_{ij}\frac{du^i}{dt}\frac{du^j}{dt}}$$

ρ_n quotient des deux formes fondamentales, a la même valeur pour toutes les courbes de S ayant même tangente.

Les directions conjuguées communes aux deux formes quadratiques s'obtiennent en écrivant que le système de deux équations homogènes :

$$\sum(\mathrm{L}_{jk} - r g_{jk})\mathrm{X}^k = 0$$

a une solution non nulle soit :

$$\begin{aligned} 0 &= \text{dét}|\mathrm{L}_{jk} - r g_{jk}| \\ &= \text{dét}|\mathrm{L}_{jk}| - r(g_{11}\mathrm{L}_{22} - 2g_{12}\mathrm{L}_{12} + g_{22}\mathrm{L}_{11}) + r^2 \text{dét}|g_{jk}|. \end{aligned}$$

Si les racines sont distinctes, il existe deux axes orthogonaux principaux diagonalisant L : ce sont les directions principales, et les courbures normales correspondantes, solutions de l'équation ci-dessus sont les courbures principales, extrêma de ρ_n lorsque la tangente tourne autour de 0.

La demi-somme des racines est la courbure moyenne de S en M, le produit est la courbure de Gauss G_M :

$$G_M = \frac{\text{dét}|L_{jk}|}{\text{dét}|g_{jk}|}.$$

La composante tangentielle de M''_{t^2} s'écrit avec les notations choisies :

$$\begin{aligned}(M''_{t^2})_T &= \frac{\partial M}{\partial u^1}\frac{d^2u^1}{dt^2} + \frac{\partial M}{\partial u^2}\frac{d^2u^2}{dt^2} + \sum \Gamma^k_{ij}\frac{\partial M}{\partial u^k}\frac{du^i}{dt}\frac{du^j}{dt} \\ &= \sum \frac{\partial M}{\partial u^k}\left(\frac{d^2u^k}{dt^2} + \sum \Gamma^k_{ij}\frac{du^i}{dt}\frac{dt^j}{dt}\right).\end{aligned}$$

On remarque d'abord que, contrairement à la composante normale de M''_{t_2}, la composante tangentielle contient les dérivées secondes $\dfrac{d^2u^k}{dt^2}$. Sa direction dans T_M dépend du paramètrage. C'est donc du vecteur accélération du chemin : $t \to M(t)$, plutôt que de la courbe, qu'il faut parler. On n'obtient un vecteur ne dépendant que de la courbe qu'en paramétrant par l'arc, ce qui donne une expression de la courbure géodésique :

$$\rho_g = \sum \frac{\partial M}{\partial u^k}\left(\frac{d^2u^k}{ds^2} + \sum \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds}\right).$$

De toutes façons l'expression de la composante tangentielle du vecteur accélération paraît compliquée. Nous allons éclairer sa signification géométrique en remarquant que l'étude du vecteur accélération $\dfrac{d^2M}{dt^2}$ est celle du champ de vecteurs vitesse le long du chemin $t \to M(t)$. Or, considérons plus généralement un champ de vecteurs $V(t)$ tangents à la surface S le long d'un tel chemin. Dans un système de coordonnées (u^1, u^2) :

$$V(t) = \frac{\partial M}{\partial u^1}X^1 + \frac{\partial M}{\partial u^2}X^2 = \sum \frac{\partial M}{\partial u^i}X^i.$$

Supposons $\mathrm{V}(t)$ différentiable :

$$\frac{d\mathrm{V}}{dt} = \sum \frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j} \mathrm{X}^i \frac{du^j}{dt} + \sum \frac{\partial \mathrm{M}}{\partial u^i} \frac{dx^i}{dt}.$$

La composante tangentielle de $\dfrac{d\mathrm{V}}{dt}$ ou *dérivée tangentielle de* V se calcule simplement en remplaçant les $\dfrac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j}$ par leurs composantes tangentielles introduites plus haut. On obtient :

$$\left(\frac{d\mathrm{V}}{dt}\right)_{\mathrm{T}} = \sum \frac{\partial \mathrm{M}}{\partial u^k} \left(\frac{d\mathrm{X}^k}{dt} + \sum \Gamma^k_{ij} \mathrm{X}^i \frac{du^j}{dt} \right)$$

qui généralise l'expression de la composante tangentielle de M''_{t^2}.

Si l'on change le paramètre du chemin $t \to \mathrm{M}(t)$ par $t = \varphi(\tau)$, le vecteur dérivée de V est multiplié par $\dfrac{dt}{d\tau}$ et ne change pas de direction. La nullité de la composante tangentielle de $\dfrac{d\mathrm{V}}{dt}$ est donc une propriété indépendante du paramétrage :

Définition VII.8.B. (E_3)

Un champ de vecteurs V tangents à une surface S, défini le long d'une courbe régulière C de S est dit *parallèle sur* S *le long de* C s'il est différentiable et si sa dérivée tangentielle est nulle en tout point de C.

Cette définition peut être utilement étendue aux courbes *régulières «par morceaux»*. Une telle courbe est une courbe continue mais qui n'est régulière qu'à l'exception d'un nombre fini de points où elle possède un vecteur dérivé à gauche et un vecteur dérivé à droite.

Si l'on se donne en un point M_0 d'une courbe C de S, régulière par morceaux, un vecteur V_0 tangent à S, il existe alors le long de C un champ de vecteurs parallèles V, égal à V_0 en M_0, uniquement déterminé comme solution du système d'équations différentielles linéaires :

$$\frac{d\mathrm{X}^k}{dt} + \sum \Gamma^k_{ij} \frac{du^j}{dt} \mathrm{X}^i = 0 \qquad k = 1, 2.$$

V est dit le *transporté parallèle* de V_0 le long de C.

Si V et W sont deux champs de vecteurs parallèles le long de la courbe C :

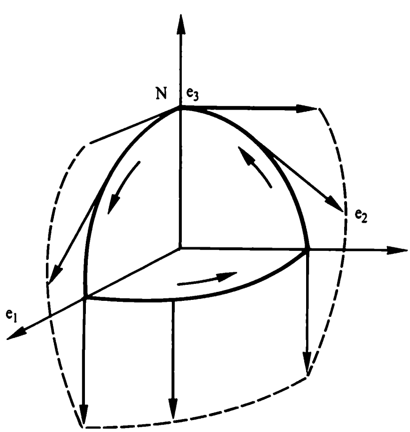
$$\begin{aligned}\frac{d}{dt}(V|W) &= \left(\frac{dV}{dt}\middle|W\right) + \left(V\middle|\frac{dW}{dt}\right)\\ &= \left(\left(\frac{dV}{dt}\right)_T\middle|W\right) + \left(V\middle|\left(\frac{dW}{dt}\right)_T\right) = 0 \text{ d'où :}\end{aligned}$$

Proposition VII.8.

Le produit scalaire et en particulier la longueur de vecteurs parallèles sur une surface S le long d'une courbe C sont constants.

Autrement dit, le transport parallèle conserve le produit scalaire.

Remarque : Si l'on transporte parallèlement un vecteur le long d'une courbe fermée, même petite, le vecteur ne revient pas coïncider avec le vecteur initial. C'est une manifestation de la *courbure* de la surface S, comme l'illustre la figure ci-dessous où un vecteur égal à e_1 partant du pôle Nord de la sphère revient après avoir été transporté parallèlement le long de trois « quarts de grands cercles » successifs à son point de départ égal au vecteur e_2 ayant subi une rotation de $+\frac{\pi}{2}$.



Un chemin $t \to \mathrm{M}(t)$ de la surface S est dit *géodésique* si le champ que forment ses vecteurs-vitesse est parallèle sur S. Dans ce cas le paramétrage ne peut pas être arbitraire.

En effet de $\left\| \dfrac{d\mathrm{M}}{dt} \right\| =$ constante, il résulte que $\dfrac{dt}{ds}$ est constant et que $t = as + b$ est une fonction affine de l'arc. Néanmoins l'arc définissant un paramétrage canonique, la propriété d'être géodésique est une propriété de la courbe C :

Définition VII.8.C.

On appelle *géodésique* d'une surface S de l'espace euclidien E_3 toute courbe C qui, lorsqu'elle est paramétrée par l'arc, a ses vecteurs vitesses parallèles sur S le long de C.

Les calculs faits précédemment ont pour conséquence la :

Proposition VII.8.B.

Une courbe C d'une surface S de E_3 est une géodésique si et seulement si sa courbure géodésique est nulle, autrement dit, si et seulement si son plan osculateur est normal à S en tout point.

Nous allons maintenant étudier les coefficients Γ_{ij}. On a :

$$\left(\frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j} \Big| \frac{\partial \mathrm{M}}{\partial u^l} \right) = \sum \Gamma_{ij}^k g_{kl} = \Gamma_{ijl}.$$

Les Γ_{ij}^k et Γ_{ijl} sont appelés les *symboles de Christoffel* respectivement de seconde et de première espèce. Ils sont symétriques en i et j.

Nous allons montrer qu'ils dérivent de la métrique g et peuvent donc être considérés comme intrinsèquement attachés à S. En effet :

$$\frac{\partial g_{jl}}{\partial u^i} = \frac{\partial}{\partial u^i} \left(\frac{\partial \mathrm{M}}{\partial u^j} \Big| \frac{\partial \mathrm{M}}{\partial u^l} \right) = \left(\frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j} \Big| \frac{\partial \mathrm{M}}{\partial u^l} \right) + \left(\frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^l} \Big| \frac{\partial \mathrm{M}}{\partial u^j} \right)$$
$$= \Gamma_{ijl} + \Gamma_{ilj}.$$

En permutant les trois indices i, j, l, on obtient de la même façon :

$$\frac{\partial g_{ij}}{\partial u^l} = \Gamma_{lij} + \Gamma_{lji}; \ \frac{\partial g_{li}}{\partial u^j} = \Gamma_{jli} + \Gamma_{jil}$$

d'où l'on tire : $\Gamma_{ijl} = \dfrac{1}{2}\left(\dfrac{\partial g_{jl}}{\partial u^i} + \dfrac{\partial g_{il}}{\partial u^j} - \dfrac{\partial g_{ij}}{\partial u^l}\right)$

tandis que, en utilisant la matrice inverse $|g^{ij}|$ de $|g_{ij}|$ $\Gamma^k_{ij} = \sum g^{kl}\Gamma_{ijl}$.

Remarques :

1) On rencontre fréquemment dans la littérature les notations : $\Gamma^k_{ij} = \left\{{k \atop ij}\right\}$ et $\Gamma_{ijk} = [ij, k]$.

2) Il faut prendre garde que, lors d'un changement de coordonnées de S, les formules de transformation des symboles de Christoffel font intervenir les dérivées secondes. Malgré leur collection bien apparente d'indices, les Γ^k_{ij} et Γ_{ijk} *ne sont donc pas* des tenseurs sur S comme on le verra au paragraphe suivant.

VII.9 – Les variétés dans E_n. Coordonnées curvilignes. Dérivation covariante

Ce paragraphe n'est que la généralisation immédiate de ce qui vient d'être dit sur les surfaces.

Il constitue néanmoins une transition utile vers la généralisation suivante qui est la théorie des variétés riemanniennes, où tout espace euclidien ambiant a disparu (Chap. IX).

Définition VII.9.A.

Une application r fois continûment différentiable, $r \geqslant 1$, d'un domaine connexe D de $\mathbb{R}^p$ dans un espace vectoriel réel E de dimension $n \in \mathrm{D} \to \mathrm{M}(u) \in \mathrm{E}$ est dite *régulière* si quel que soit $u \in \mathrm{D}$, les p vecteurs : $\dfrac{\partial \mathrm{M}}{\partial u^j}$ sont linéairement indépendants. Lorsque $\mathrm{E} = \mathbb{R}^p$, cette condition s'exprime par la non-nullité du jacobien :

$$\frac{\mathrm{D}(x^1, x^2, \dots x^p)}{\mathrm{D}(u^1, u^2, \dots u^p)} = \text{dét}\left|\frac{\partial x^i}{\partial u^j}\right|.$$

Puisque ce jacobien est alors une fonction continue à valeurs réelles non nulle sur un domaine connexe, il est partout ou bien strictement positif et l'application conserve l'orientation, ou bien strictement négatif et l'application change l'orientation en son opposée.

Définition VII.9.B.

Un *élément de variété* S *de dimension* p d'un espace vectoriel réel de dimension $n \geqslant p$ de classe $\mathfrak{C}_r$ $(r \geqslant 3)$ est défini par une application d'un domaine connexe D de $\mathbb{R}^p : u \rightarrow \mathrm{M}(u)$, dans E, de classe $\mathfrak{C}_r$, *régulière bijective*, appelée *système de coordonnées de* S. Toute application bijective régulière φ de classe $\mathfrak{C}_r$ d'un domaine D$'$ de $\mathbb{R}^p$ sur D : $v \rightarrow u = \varphi(v)$ définit par composition : $v \rightarrow \mathrm{M}(\varphi(v))$ un nouveau système de coordonnées dit *équivalent* au premier, de même orientation si $\mathrm{J}(\varphi) > 0$, d'orientation opposée si $\mathrm{J}(\varphi) < 0$.

Si $n = p$, l'application $u \rightarrow \mathrm{M}(u)$ définit alors un système de *coordonnées* souvent appelées « *curvilignes* » dans un domaine S de E_n « difféomorphe » de D.

Lorsque E est euclidien, le sous-espace T_M de base $\left(\dfrac{\partial \mathrm{M}}{\partial u^1}, \dfrac{\partial \mathrm{M}}{\partial u^2}, \ldots \dfrac{\partial \mathrm{M}}{\partial u^p}\right)$ est muni de la métrique euclidienne induite :

$$(\mathrm{X}|\mathrm{Y}) = \left(\sum \frac{\partial \mathrm{M}}{\partial u^i}\mathrm{X}^i \Big| \sum \frac{\partial \mathrm{M}}{\partial u^j}\mathrm{Y}^j\right) = \sum \left(\frac{\partial \mathrm{M}}{\partial u^i}\Big|\frac{\partial \mathrm{M}}{\partial u^j}\right)\mathrm{X}^i\mathrm{Y}^j$$
$$= \sum g_{ij}\mathrm{X}^i\mathrm{Y}^j.$$

T_M est l'*espace tangent à* S *au point* M. Il est orienté par le système de coordonnées choisi. Si $\overline{g}_u = \text{dét}|g_{ij}| = \left\|\dfrac{\partial \mathrm{M}}{\partial u^1} \wedge \dfrac{\partial \mathrm{M}}{\partial u^2} \wedge \ldots \dfrac{\partial \mathrm{M}}{\partial u^p}\right\|^2$, la p-forme extérieure sur T_M : $\overline{g}_u^{1/2} du^1 \wedge du^2 \wedge \ldots du^p$ est un invariant orienté de T_M appelé l'*élément de volume de* S.

Les dérivées secondes des vecteurs tangents de base se décomposent en une composante normale à S, nulle si $n = p$, et une composante tangentielle :

$$\frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j} = \mathrm{N}_{ij} + \sum \Gamma_{ij}^k \frac{\partial \mathrm{M}}{\partial u^k}.$$

On calcule les coefficients Γ_{ij}^k en éliminant la composante normale par le produit scalaire avec les vecteurs de la base de T_M :

$$\left(\frac{\partial^2 M}{\partial u^i \partial u^j} \Big| \frac{\partial M}{\partial u^l}\right) = \sum \Gamma_{ij}^k \left(\frac{\partial M}{\partial u^k} \Big| \frac{\partial M}{\partial u^l}\right) = \sum \Gamma_{ij}^k g_{kl} = \Gamma_{ij,l}.$$

Les $\Gamma_{ij}^k = \left\{ {k \atop ij} \right\}$ et $\Gamma_{ij,l} = [ij, l]$ sont *les symboles de Christoffel, de seconde et de première espèces* respectivement. Leur calcul est exactement celui du paragraphe précédent :

$$\Gamma_{ij,l} = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial u^j} + \frac{\partial g_{jl}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^l}\right) \text{ et } \Gamma_{ij}^k = \sum \Gamma_{ij,l} g^{lk}$$

et montre qu'ils sont déterminés par la métrique g de S.

Si, par exemple, les dérivées en un point M des coefficients g_{ij} de la métrique de S sont nulle (la métrique de S est « tangente » en M à une métrique euclidienne), les $\Gamma_{ij,l}$ et les Γ_{ij}^k sont tous nuls.

Lors d'un changement de coordonnées $(v) \to (u = \varphi(v))$, un calcul direct permet d'obtenir les formules suivantes de transformation des symboles de Christoffel :

$$(\Gamma_{jk}^i)_v = \sum \Gamma_{mn}^l \frac{\partial v^i}{\partial u^l} \frac{\partial u^m}{\partial v^j} \frac{\partial u^n}{\partial v^k} + \sum \frac{\partial v^i}{\partial u^r} \frac{\partial^2 u^r}{\partial v^j \partial v^k}$$
$$(\Gamma_{ij,k})_v = \sum \Gamma_{lm,n} \frac{\partial u^l}{\partial v^i} \frac{\partial u^m}{\partial v^j} \frac{\partial u^n}{\partial v^k} + \sum g_{rs} \frac{\partial u^r}{\partial v^k} \frac{\partial^2 u^s}{\partial v^i \partial v^j}.$$

Elles font intervenir les dérivées secondes du changement de coordonnées. Les symboles de Christoffel ne se transforment donc pas de façon tensorielle lors du changement de base correspondant de l'espace tangent : *ce ne sont pas les composantes d'un tenseur*.

Considérons un chemin régulier sur S : $t \in [a, b] \to M(u(t)) = M(t)$, et un champ de vecteurs, tangents à S, le long du chemin : $t \in [a, b] \to V(t) \in T_M \subset E_n$. On suppose V différentiable.

Si $V = \sum \dfrac{\partial M}{\partial u^i} V^i$, sa dérivée dans l'espace ambiant E_n, est
$$\frac{dV}{dt} = \sum \frac{\partial M}{\partial u^i} \frac{dV^i}{dt} + \sum V^i \frac{\partial^2 M}{\partial u^i \partial u^j} \frac{du^j}{dt}.$$

Nous noterons sa composante tangentielle avec un D majuscule et l'appellerons désormais *dérivée covariante* de V dans S, terme qui écarte toute référence à un éventuel espace ambiant, et qui, ultérieurement, marquera le fait que cette dérivation ne dépend que de la structure de variété riemannienne de S, c'est-à-dire de sa métrique propre.

En faisant apparaître les composantes des $\dfrac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j}$:

$$\frac{\mathrm{DV}}{dt} = \sum \frac{\partial \mathrm{M}}{\partial u^k} \left(\frac{d\mathrm{V}^k}{dt} + \sum \mathrm{V}^i \Gamma^k_{ij} \frac{du^j}{dt} \right).$$

La « vitesse » de V est somme de sa vitesse relative et de sa vitesse d'entraînement.

Du point de vue tensoriel, la propriété essentielle est que les composantes de la dérivée covariante de V : $\left(\dfrac{\mathrm{DV}}{dt}\right)^k = \dfrac{d\mathrm{V}^k}{dt} + \sum \mathrm{V}^i \Gamma^k_{ij} \dfrac{du^j}{dt}$ sont les composantes d'un vecteur tangent à S, ce qui veut dire que, par un changement de coordonnées de S : $(v) \rightarrow u = \varphi(v)$, et *malgré* toutes les dérivations qui interviennent, *les composantes ci-dessus se transforment comme les composantes d'un vecteur* de l'espace au moyen de la transformation linéaire du changement de bases. Si $\mathrm{X} = \sum \mathrm{X}^j \dfrac{\partial \mathrm{M}}{\partial u^j} = \sum \overline{\mathrm{X}}^k \dfrac{\partial \mathrm{M}}{\partial v^k} = \sum \mathrm{X}^j \dfrac{\partial \mathrm{M}}{\partial v^k} \dfrac{\partial v^k}{\partial u^j}$ on a : $\overline{\mathrm{X}}^k = \sum \mathrm{X}^j \dfrac{\partial v^k}{\partial u^j}$. La propriété indiquée s'écrit donc : $\left(\dfrac{\overline{\mathrm{DV}}}{dt}\right)^k = \sum \left(\dfrac{\mathrm{DV}}{dt}\right)^j \dfrac{\partial v^k}{\partial u^j}$.

Le champ de vecteurs V est parallèle sur la variété S le long du chemin $\mathrm{M}(t)$ si $\dfrac{\mathrm{DV}}{dt} \equiv 0$. Puisque la dérivée covariante a été obtenue comme composante tangentielle de la dérivée de l'espace euclidien ambiant, on a, si V et W sont parallèles sur S le long de $\mathrm{M}(t)$:

$$\frac{d}{dt}(\mathrm{V}|\mathrm{W}) = \left(\frac{\mathrm{DV}}{dt}\Big|\mathrm{W}\right) + \left(\mathrm{V}\Big|\frac{\mathrm{DW}}{dt}\right) = 0$$

d'où il résulte que le transport parallèle sur S conserve le produit scalaire.

La propriété de parallèlisme le long du chemin $\mathrm{M}(t)$, régulier par morceaux, s'exprime par le système d'équations différentielles linéaires :

$$\frac{d\mathrm{V}^k}{dt} = -\sum \Gamma_{ij}^k \frac{d\mathrm{M}^j}{dt}\mathrm{V}^i = \sum \mathrm{A}_i^k(t)\mathrm{V}^i$$

d'où il résulte, comme précédemment, que la donnée d'un vecteur tangent V_0 au point $\mathrm{M}(t_0)$ détermine de façon unique et le long de la totalité du chemin $\mathrm{M}(t)$ un champ de vecteurs parallèles.

Une *géodésique* de S est une courbe qui possède un paramétrage tel que les vecteurs vitesses soient parallèles donc de longueur constante, et le paramètrage est alors nécessairement une fonction affine de l'arc. L'équation des géodésiques est comme précédemment :

$$\frac{\mathrm{D}}{dt}\left(\frac{d\mathrm{M}}{dt}\right) = \sum \frac{\partial \mathrm{M}}{\partial u^k}\left(\frac{d^2u^k}{dt^2} + \sum \Gamma_{ij}^k \frac{du^i}{dt}\frac{du^j}{dt}\right) = 0.$$

Il faut bien observer, que dans le cas particulier ou $n = p$, c'est-à-dire celui des « coordonnées curvilignes » dans E_n, le parallèlisme est le parallélisme ordinaire dans E_n et que l'équation des géodésiques est tout simplement l'équation générale des droites de E_n : $\dfrac{d^2\mathrm{M}}{dt^2} = 0$ en coordonnées curvilignes.

Supposons maintenant que V soit un champ de vecteurs tangents à S, défini et différentiable dans tout un voisinage du point M. La restriction de V à un chemin régulier $t \to \mathrm{M}(t)$ issu de M est alors un champ de vecteurs le long de ce chemin, et sa dérivée covariante s'écrit :

$$\begin{aligned}\frac{\mathrm{DV}}{dt} &= \sum \frac{\partial \mathrm{M}}{\partial u^k}\left(\sum \frac{\partial \mathrm{V}^k}{\partial u^l}\frac{du^l}{dt} + \sum \mathrm{V}^i\Gamma_{ij}^k \frac{du^j}{dt}\right)\\ &= \sum \frac{du^j}{dt}\left(\frac{\partial \mathrm{V}^k}{\partial u^j} + \sum \Gamma_{ij}^k \mathrm{V}^i\right)\frac{\partial \mathrm{M}}{\partial u^k}.\end{aligned}$$

Cette expression ne dépend que du vecteur tangent $\dfrac{d\mathrm{M}}{dt} = \sum \dfrac{\partial \mathrm{M}}{\partial u^j}\dfrac{du^j}{dt}$ de composantes $\dfrac{du^j}{dt}$ au chemin considéré.

Tous les chemins issus de M ayant même vecteur tangent déterminent la même dérivée du champ V. On peut donc définir la *dérivée covariante du champ de vecteurs* V *relativement à un vecteur tangent* X *à* S *en* M, par

$$\boxed{\mathrm{D_X V} = \sum \mathrm{X}^j \left(\frac{\partial \mathrm{V}^k}{\partial u^j} + \sum \Gamma_{ij}^k \mathrm{V}^i \right) \frac{\partial \mathrm{M}}{\partial u^k}}.$$

Puisque $\mathrm{X}^j = \langle du^j, \mathrm{X} \rangle$, la *différentielle covariante* du champ de vecteurs V est le champ d'opérateurs linéaires des espaces tangents ou champ de tenseurs d'ordre deux sur S définis en « composantes une fois contravariante, une fois covariante » par :

$$\begin{aligned} \mathrm{DV} &= \sum \left[\sum_k \left(\frac{\partial \mathrm{V}^k}{\partial u^j} + \sum \Gamma_{ij}^k \mathrm{V}^i \right) \frac{\partial \mathrm{M}}{\partial u^k} \right] \otimes du^j \\ &= \sum \frac{\partial \mathrm{M}}{\partial u^k} \otimes (d\mathrm{V}^k + \sum \Gamma_{ij}^k \mathrm{V}^i du^j). \end{aligned}$$

Il est d'usage de faire précéder d'une barre (ou d'une virgule) l'indice j lorsqu'on désire indiquer la dérivation covariante partielle relativement au j-ème vecteur tangent de base $\dfrac{\partial \mathrm{M}}{\partial u^j}$ soit :

$$\mathrm{V}_{|j} = \mathrm{D}_j \mathrm{V} = \frac{\mathrm{DV}}{du^j} = \mathrm{D}_{\frac{\partial \mathrm{M}}{\partial u^j}} \mathrm{V} = \sum \left(\frac{\partial \mathrm{V}^k}{\partial u^j} + \sum \Gamma_{ij}^k \mathrm{V}^i \right) \frac{\partial \mathrm{M}}{\partial u^k}$$

$$\boxed{\mathrm{V}_{|j}^k = \frac{\partial \mathrm{V}^k}{\partial u^j} + \sum \Gamma_{ij}^k \mathrm{V}^i}.$$

Il est maintenant aisé de calculer la dérivée covariante d'un champ de covecteurs A le long d'un chemin $\mathrm{M}(t)$.

Ce champ A est défini par ses composantes $\mathrm{A}_k(t)$, soit : $\mathrm{A}(t) = \sum \mathrm{A}_k(t) du^k$ dans l'espace cotangent $\mathrm{T}^*_{\mathrm{M}(t)}$. Prenons le long du chemin $\mathrm{M}(t)$ un champ de vecteurs contravariants parallèle : $\mathrm{X}(t) = \sum \mathrm{X}^k(t) \dfrac{\partial \mathrm{M}}{\partial u^k}$. Le crochet de dualité entre A et X est la fonction à valeurs réelles définie le long du chemin par les fonctions à valeurs réelles $\mathrm{A}_k(t)$ et $\mathrm{X}^k(t)$ par :

$$\langle \mathrm{A}, \mathrm{X} \rangle = \sum \mathrm{A}_k(t) \mathrm{X}^k(t)$$

et prend la même valeur quel que soit le système de coordonnées, de même que sa dérivée :

$$\frac{d}{dt}\langle \mathrm{A}, \mathrm{X}\rangle = \sum \frac{d\mathrm{A}_i}{dt}\mathrm{X}^i + \sum \mathrm{A}_k \frac{d\mathrm{X}^k}{dt} = \sum \left(\frac{d\mathrm{A}_i}{dt} - \sum \mathrm{A}_k \Gamma^k_{ij} \frac{du^j}{dt} \right) \mathrm{X}^i.$$

Cette expression étant indépendant du choix de la base, il en résulte que les termes entre parenthèses sont les composantes d'un covecteur appelé *dérivée covariante* du champ A :

$$\frac{\mathrm{DA}}{dt} = \sum \left(\frac{d\mathrm{A}_i}{dt} - \sum \mathrm{A}_k \Gamma^k_{ij} \frac{du^j}{dt} \right) du^i.$$

Si maintenant A est un champ différentiable de covecteurs sur S, sa *différentielle covariante* est un champ de tenseurs deux fois covariants $\mathrm{DA} = \sum (d\mathrm{A}_i - \sum \Gamma^k_{ij} \mathrm{A}_k du^j) \otimes du^i$.

On obtient dès lors immédiatement la dérivée covariante d'un champ de tenseurs de variance quelconque sur S.

Si par exemple T est un champ de tenseurs une fois contravariant, une fois covariant, de composantes T^i_j, sa différentielle covariante est :

$$\mathrm{DT} = \sum (d\mathrm{T}^i_j + \sum \mathrm{T}^m_j \Gamma^i_{mk} du^k - \sum \mathrm{T}^i_p \Gamma^p_{jq} du^q) \otimes \frac{\partial \mathrm{M}}{\partial u^i} \otimes du^j.$$

La différentielle covariante d'un champ de tenseurs deux foix covariants T *est le champ de tenseurs trois fois covariant :*

$$\mathrm{DT} = \sum (d\mathrm{T}_{pq} - \sum \mathrm{T}_{kq} \Gamma^k_{pj} du^j - \sum \mathrm{T}_{pk} \Gamma^k_{jq} du^j) \otimes du^p \otimes du^q.$$

Nous allons utiliser cette expression pour démontrer une propriété à laquelle on peut s'attendre. Le tenseur métrique g est un champ de tenseurs deux fois covariant dont les composantes sont définies en chaque point par :

$$g_{pq} = \left(\frac{\partial \mathrm{M}}{\partial u^p} \middle| \frac{\partial \mathrm{M}}{\partial u^q} \right)$$

$$\text{d'où } dg_{pq} = \left(\sum \frac{\partial^2 \mathrm{M}}{\partial u^p \partial u^j} \middle| \frac{\partial \mathrm{M}}{\partial u^q} \right) du^j + \left(\frac{\partial \mathrm{M}}{\partial u^p} \middle| \sum \frac{\partial^2 \mathrm{M}}{\partial u^q \partial u^j} \right) du^j$$

$$= \sum \Gamma_{pj,q} du^j + \sum \Gamma_{jq,p} du^j$$

$$\text{et} \quad dg_{pq} = \sum \Gamma^k_{pj} g_{kq} du^j + \sum \Gamma^k_{jq} g_{pk} du^j$$

En reportant dans l'expression de $\mathrm{D}g$, on obtient $\mathrm{D}g = 0$.

Dans le reste de ce paragraphe on se place dans l'espace euclidien E_n tout entier.

L'opérateur de différentiation, ou de gradient, qui s'écrit en coordonnées « rectilignes » $\mathrm{D} = e^*j \otimes \dfrac{\partial}{\partial x^j}$ devient dans un système de coordonnées curvilignes $(u^1, u^2, \dots u^n)$ de E_n : $\mathrm{D} = \sum du^j \otimes \mathrm{D}_j$ où D_j désigne la dérivée covariante partielle relative au j-ème vecteur de base $\dfrac{\partial \mathrm{M}}{\partial u^j}$. En effet, si T est un champ de tenseurs différentiable et X un vecteur tangent, on a : $\mathrm{X} = \sum \mathrm{X}^j \dfrac{\partial \mathrm{M}}{\partial u^j}$, et $\mathrm{D_X T} = \sum \mathrm{X}^j \mathrm{D}_j \mathrm{T} = \sum \langle \mathrm{X}, du^j \rangle \mathrm{D}_j \mathrm{T}$ d'où $\mathrm{D} = \sum du^j \otimes \mathrm{D}_j$. Si f est une fonction numérique différentiable, sa différentielle $\mathrm{D}f = df = \sum du^j \dfrac{\partial f}{\partial u^j}$ est un champ de covecteurs, ou forme différentielle de degré un, et l'on réserve plutôt le nom de *gradient de* f à son expression en composantes contravariantes : $\operatorname{grad} f = \sum g^{ij} \dfrac{\partial f}{\partial u^j} \dfrac{\partial \mathrm{M}}{\partial u^i}$ ce qui permet d'écrire la dérivée de f par rapport à un vecteur X comme un produit scalaire :

$$\frac{\partial f}{\partial \mathrm{X}} = \mathrm{D_X} f = \sum \frac{\partial f}{\partial u^j} \mathrm{X}^j = (\operatorname{grad} f | \mathrm{X}).$$

Le calcul de la *divergence* d'un champ de vecteurs V en coordonnées curvilignes est instructif. Div V est la fonction scalaire Trace DV, soit :

$$\begin{aligned}\operatorname{div} \mathrm{V} &= \sum \langle du^j, \mathrm{D}_j \mathrm{V} \rangle = \sum \langle du^j, \sum \left(\frac{\partial \mathrm{V}^k}{\partial u^j} + \sum \Gamma^k_{ij} \mathrm{V}^i \right) \frac{\partial \mathrm{M}}{\partial u^k} \rangle \\ &= \sum \frac{\partial \mathrm{V}^j}{\partial u^j} + \sum \left(\sum \Gamma^j_{ij} \right) \mathrm{V}^i.\end{aligned}$$

On est donc ramené au calcul de $\sum\limits_i \Gamma^j_{ij}$ qui ressemble à une trace, et qui en est effectivement une.

Les coefficients de Christoffel de seconde espèce Γ^k_{ij} d'un système de coordonnées curvilignes de E_n ont été définis par les équations :

$$\frac{\partial^2 \mathrm{M}}{\partial u^i \partial u^j} = \sum \Gamma_{ij}^k \frac{\partial \mathrm{M}}{\partial u^k} = \frac{\partial}{\partial u^i}\left(\frac{\partial \mathrm{M}}{\partial u^j}\right)$$

i étant fixé, les Γ_{ij}^k sont donc les coefficients de la matrice de l'opérateur linéaire $\dfrac{\partial}{\partial u^i}$ appliqué aux vecteurs de base et $\sum \Gamma_{ij}^j$ en est la trace. Or, on a :

$$\overline{g} = \left(\frac{\partial \mathrm{M}}{\partial u^1} \wedge \frac{\partial \mathrm{M}}{\partial u^2} \wedge \dots \frac{\partial \mathrm{M}}{\partial u^n} \middle| \frac{\partial \mathrm{M}}{\partial u^1} \wedge \frac{\partial \mathrm{M}}{\partial u^2} \wedge \dots \frac{\partial \mathrm{M}}{\partial u^n}\right).$$

D'après le § IV.13 sur les dérivations dans l'algèbre extérieure, la dérivation par $\dfrac{\partial}{\partial u^i}$ de $\dfrac{\partial \mathrm{M}}{\partial u^1} \wedge \dfrac{\partial \mathrm{M}}{\partial u^2} \wedge \dots \dfrac{\partial \mathrm{M}}{\partial u^n}$ revient à multiplier ce produit extérieur par la trace de l'opérateur linéaire correspondant, d'où :

$$\frac{\partial \overline{g}}{\partial u^i} = 2(\sum \Gamma_{ij}^j)\overline{g} \text{ et } \boxed{\sum \Gamma_{ij}^j = \frac{1}{2}\frac{\partial}{\partial u^i} \operatorname{Log} \overline{g} = \frac{1}{\sqrt{\overline{g}}}\frac{\partial \sqrt{\overline{g}}}{\partial u^i}}.$$

On obtient donc finalement :

$$\boxed{\operatorname{div} \mathrm{V} = \frac{1}{\sqrt{\overline{g}}} \sum \frac{\partial}{\partial u^i}(\sqrt{\overline{g}} \cdot \mathrm{V}^i)} \text{ (cf § VII.10 et définition IX.3.D).}$$

Le *Laplacien* d'une fonction numérique deux fois différentiable est la divergence de son gradient qui s'écrit donc en coordonnées curvilignes :

$$\boxed{\Delta f = \operatorname{div} \operatorname{grad} f = \frac{1}{\sqrt{\overline{g}}} \sum \frac{\partial}{\partial u^i}\left(\sqrt{\overline{g}} \cdot g^{ij} \frac{\partial f}{\partial u^j}\right)}.$$

Cette expression se réduit, bien entendu, dans une base orthonormale, à $\Delta f = \sum\limits_i \dfrac{\partial^2}{(\partial u^i)^2}$.

Dans l'espace euclidien usuel E_3, la mécanique et la physique s'écrivent à l'aide des trois opérateurs : gradient, divergence et rotationnel, ce dernier n'existant, sous sa forme habituelle, que dans E_3. Dans une base orthonormale $(\vec{i}, \vec{j}, \vec{k})$, les expressions de ces opérateurs apparaissent dans les différentielles des $0, 1$ et 2-formes extérieures ce qui leur donne une certaine unité. Il en résulte qu'on peut écrire ces trois opérateurs à l'aide d'un seul : l'opérateur de dérivation ∇ (del) qui s'écrit dans la *base orthonormée* $(\vec{i}, \vec{j}, \vec{k})$ comme un vecteur (covariant) : $\nabla = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}$.

L'avantage de cet opérateur synoptique est double. Outre l'unification des expressions qu'il permet, c'est une notation « légère » alors que les notations : grad, rot, div ne le sont guère. On peut ainsi effectuer le « produit » de ∇ par une fonction numérique (gradient), le « produit scalaire » de ∇ par un champ de vecteurs V (divergence) et le « produit vectoriel » de ∇ par V (rotationnel).

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

et
$$\operatorname{grad} f = \vec{i}\frac{\partial f}{\partial x} + \vec{j}\frac{\partial f}{\partial y} + \vec{k}\frac{\partial f}{\partial z} = \nabla f$$

$$\begin{aligned} d(\mathrm{X}dx + \mathrm{Y}dy + \mathrm{Z}dz) &= \left(\frac{\partial \mathrm{Z}}{\partial y} - \frac{\partial \mathrm{Y}}{\partial z}\right) dydz + \left(\frac{\partial \mathrm{X}}{\partial z} - \frac{\partial \mathrm{Z}}{\partial x}\right) dz \cdot dx \\ &\quad + \left(\frac{\partial \mathrm{Y}}{\partial x} - \frac{\partial \mathrm{X}}{\partial y}\right) dxdy \end{aligned}$$

$$\text{et}\ \operatorname{rot} \mathrm{V} = \vec{i}\left(\frac{\partial \mathrm{Z}}{\partial y} - \frac{\partial \mathrm{Y}}{\partial z}\right) + \vec{j}\left(\frac{\partial \mathrm{X}}{\partial z} - \frac{\partial \mathrm{Z}}{\partial x}\right) + \vec{k}\left(\frac{\partial \mathrm{Y}}{\partial x} - \frac{\partial \mathrm{X}}{\partial y}\right) = \nabla \times \mathrm{V}$$

$$d(\mathrm{X}dydz + \mathrm{Y}dzdx + \mathrm{Z}(dxdy)) = \left(\frac{\partial \mathrm{X}}{\partial x} + \frac{\partial \mathrm{Y}}{\partial y} + \frac{\partial \mathrm{Z}}{\partial z}\right) dxdydz$$

et
$$\operatorname{div} \mathrm{V} = \frac{\partial \mathrm{X}}{\partial x} + \frac{\partial \mathrm{Y}}{\partial y} + \frac{\partial \mathrm{Z}}{\partial z} = \nabla \cdot \mathrm{V}$$

On peut également exprimer avec ∇ l'opération de dérivation par rapport à un vecteur A soit $\mathrm{D_A}$. C'est tout simplement l'opérateur produit scalaire (dans l'ordre) : $(\mathrm{A}\cdot\nabla) = \mathrm{A}_x\frac{\partial}{\partial x} + \mathrm{A}_y\frac{\partial}{\partial y} + \mathrm{A}_z\frac{\partial}{\partial z}$.

On a $\nabla^2 f = \Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$ mais aussi :

$$\nabla^2\mathrm{V} = (\Delta\mathrm{X})\vec{\imath} + (\Delta\mathrm{Y})\vec{\jmath} + (\Delta\mathrm{Z})\vec{k}$$

qui s'écrit aussi : $\nabla^2\mathrm{V} = \mathrm{grad}\,\mathrm{div}\,\mathrm{V} - \mathrm{rot}\,\mathrm{rot}\,\mathrm{V}$.

Voici quelques identités, fréquemment utilisées dans les applications, et d'obtention facile :

$$\begin{aligned}
\mathrm{grad}(fg) &= \nabla(fg) = f\nabla g + g\nabla f \\
\mathrm{grad}(\mathrm{A}\cdot\mathrm{B}) &= \nabla(\mathrm{A}\cdot\mathrm{B}) = \\
&\qquad (\mathrm{B}\cdot\nabla)\mathrm{A} + (\mathrm{A}\cdot\nabla)\mathrm{B} + \mathrm{B}\times(\nabla\times\mathrm{A}) + \mathrm{A}\times(\nabla\times\mathrm{B}), \\
\mathrm{div}(f\cdot\mathrm{A}) &= \nabla(f\mathrm{A}) = (\nabla f)\cdot\mathrm{A} + f(\nabla\cdot\mathrm{A}), \\
\mathrm{div}(\mathrm{A}\times\mathrm{B}) &= \nabla\cdot(\mathrm{A}\times\mathrm{B}) = \mathrm{B}\cdot(\nabla\times\mathrm{A}) - \mathrm{A}\cdot(\nabla\times\mathrm{B}), \\
\mathrm{rot}(f\cdot\mathrm{A}) &= \nabla\times(f\mathrm{A}) = (\nabla f)\times\mathrm{A} + f(\nabla\times\mathrm{A}), \\
\nabla\times(\mathrm{A}\times\mathrm{B}) &= (\mathrm{B}\cdot\nabla)\mathrm{A} - (\mathrm{A}\cdot\nabla)\mathrm{B} + (\nabla\cdot\mathrm{B})\mathrm{A} - (\nabla\cdot\mathrm{A})\mathrm{B}. \\
\mathrm{rot}\,\mathrm{rot}\,\mathrm{A} &= \nabla\times\nabla\times\mathrm{A} = \nabla(\nabla\cdot\mathrm{A}) - \nabla^2\mathrm{A} \\
&= \mathrm{grad}\,\mathrm{div}\,\mathrm{A} - \nabla^2\mathrm{A}.
\end{aligned}$$

Un champ de vecteurs V défini dans un ouvert Ω de E_3 dont le rotationnel est identiquement nul est dit irrotationnel ou *conservatif* : le travail de ce vecteur le long d'un circuit fermé C déformable en un point dans Ω est nul. En effet, d'après le théorème de Stokes, si un circuit fermé C borde dans Ω un élément de surface S, le flux de rot V à travers S mesure le travail de V le long de C :

$$\int_{\mathrm{C}} \overrightarrow{\mathrm{V}}\cdot\overrightarrow{ds} = \int\int_{\mathrm{S}} (\nabla\times\overrightarrow{\mathrm{V}})\vec{n}d\sigma \quad \text{(cf Chap. IX).}$$

Le théorème de Gauss exprime le fait que la divergence d'un champ de vecteurs V mesure la variation du flux de ce champ. Si le volume Ω est limité par la surface S :

$$\iint_{\mathrm{S}} \overrightarrow{\mathrm{V}} \cdot \vec{n} d\tau = \iiint_{\Omega} \nabla \cdot \mathrm{V} d\omega.$$

VII.10 – Les formes différentielles extérieures d'un espace E_n ou pseudoeuclidien $\mathrm{E}_{r,s}$

Les développements qui suivent seront repris dans toute leur généralité dans le cadre des variétés pseudoriemanniennes au § IX.3. Nous nous limiterons donc, dans ce §, au strict nécessaire permettant d'étudier quelques applications comme les équations de Maxwell au § IX.13.

Désignons simplement par E l'espace métrique E_n ou $\mathrm{E}_{r,s}$ que nous supposerons *orienté*. Soit ρ l'isomorphisme de E sur E^* défini par la métrique g de E qui permet en particulier d'identifier les algèbres extérieures $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ (cf Chapitre V). Il en résulte, comme on l'a vu au § V.5, que la dualité naturelle entre $\wedge\mathrm{E}$ et $\wedge\mathrm{E}^*$ peut être exprimée comme une dualité interne dans $\wedge\mathrm{E}$ comme dans $\wedge\mathrm{E}^*$. Cette dualité s'appuie sur la forme « volume orienté » ω (Définition V.3.C) qui s'écrit dans une base positive quelconque e de E : $\omega = |\overline{g}|^{1/2} e^{*1} \wedge e^{*2} \wedge \dots e^{*n}$: ω est une n-forme différentielle extérieure, invariante par tout difféomorphisme conservant l'orientation : $\omega = |\overline{g}|^{1/2} dx^1 dx^2 \dots dx^n$ (ou l'on a omis les signes $\wedge$).

Le produit scalaire étendu à $\wedge\mathrm{E}$ et à $\wedge\mathrm{E}^*$ (§ V.4), les produits intérieurs et la dualité se transportent naturellement à l'algèbre des champs de p-vecteurs et à l'algèbre $\mathfrak{A}$ des formes différentielles extérieures sur E puisque $\mathfrak{A}$ peut-être considérée comme l'algèbre des fonctions différentiables (de classe $\mathfrak{C}_\infty$ pour simplifier), de E à valeurs dans $\wedge\mathrm{E}^*$. On dispose donc des opérations suivantes dans $\mathfrak{A}$:

a) le produit scalaire de deux p-formes $(\alpha|\beta)$, qui est ici une *fonction numérique* $\mathfrak{C}_\infty$ sur E. En particulier $(\omega|\omega) = |\overline{g}|\text{dét}|g^{ij}| = |\overline{g}| \overset{-1}{\overline{g}} = (-1)^s$.

b) le produit intérieur gauche $\lrcorner$ d'un champ de q-vecteurs u par une p-forme $\alpha (q \leqslant p)$: $u \lrcorner \alpha$. C'est le transposé de la multiplication extérieure par u : $\langle z \wedge u, \alpha \rangle = \langle z, u \lrcorner \alpha \rangle$, $\forall$ le champ de $(p-q)$ vecteurs z.

En particulier, un champ de vecteurs X détermine une dérivation de degré (-1) de l'algèbre $\mathfrak{A}$, notée $i(\mathrm{X})$ (cf § IV.13) : sur une p-forme α, $i(\mathrm{X})\alpha = (-1)^{p-1}\mathrm{X} \lrcorner\, \alpha$.

c) le produit intérieur gauche « interne » $\lrcorner$, adjoint de la multiplication extérieure à droite dans $\mathfrak{A}$: $(\gamma \wedge \beta|\alpha) = (\gamma|\beta \lrcorner\, \alpha)$.

L'image par $\wedge^p\rho$ d'un champ u de p-vecteurs est une p-forme, et puisque ρ fait passer du produit scalaire au crochet de dualité, on obtient par un calcul déjà vu (§ V.5) :

$$\begin{aligned}\langle z, u \lrcorner\, \alpha\rangle &= \langle z \wedge u, \alpha\rangle = (\wedge^p\rho(z \wedge u)|\alpha) \\ &= ((\wedge^{p-q}\rho \cdot z) \wedge (\wedge^q\rho \cdot u)|\alpha) \\ &= (\wedge^{p-q}\rho \cdot z|(\wedge^q\rho \cdot u) \lrcorner\, \alpha) = \langle z, (\wedge^q\rho \cdot u) \lrcorner\, \alpha\rangle,\end{aligned}$$

la relation entre produits intérieurs « interne » et « externe » :

$$\boxed{u \lrcorner\, \alpha = (\wedge^p\rho \cdot u) \lrcorner\, \alpha}.$$

d) le produit intérieur droit « interne » $\llcorner$, adjoint de la multiplication extérieure à gauche dans $\mathfrak{A}$.

e) l'opérateur étoile : $*$, de dualité, déterminé par l'identité : $\beta \wedge *\alpha = (\beta|\alpha)\omega$ quel que soit β (qui dépend donc de la métrique *et* de l'orientation).

Sur une p-forme $\alpha = \sum \alpha_\mathrm{I} dx^\mathrm{I}$, $(*\alpha)_{\mathrm{H}'} = \varepsilon(\mathrm{H})|\overline{g}|^{1/2} \sum g^{\mathrm{HI}}\alpha_\mathrm{I}$.

Soit en faisant apparaître les composantes contravariantes $\alpha^\mathrm{H} = \sum g^{\mathrm{HI}}\alpha_\mathrm{I}$ de α : $*\alpha = |\overline{g}|^{1/2} \sum \varepsilon(\mathrm{H})\alpha^\mathrm{H} dx^{\mathrm{H}'}$.

On a $*1 = \omega$ et $*\omega = (-1)^s$.

Les calculs effectués au § V.5 entraînent que sur les p-formes, on a : $* * -(-1)^{p(n-p)+s}$.

$$\boxed{*\alpha = (-1)^{p(n-p)}\alpha \lrcorner\, \omega} \text{ et } (*\alpha|\beta) = (\alpha|(-1)^s \overset{-1}{*} \beta).$$

Le *gradient* d'une fonction numérique différentiable f sur E est le champ de vecteurs contravariants $\operatorname{grad} f = \overset{-1}{\rho}\, df$, représentation contravariante de la 1-forme df. La *codifférentielle* δ est l'opérateur dual de la différentielle extérieure. Sur les p-formes : $\delta = (-1)^p \overset{-1}{*}\, d*$.

Elle diminue le degré d'une unité et $\delta\delta = 0$.

La *divergence* d'un champ de vecteurs X est le scalaire $\boxed{\mathrm{div}\, X = -\delta(\rho X)} = -\delta\xi$ où $\xi = \rho X$ est la 1-forme associée à X. Nous allons vérifier que cette définition est bien identique aux définitions antérieures (§ V.5 et VII.9) et ce faisant obtiendrons encore d'autres expressions de la divergence. En appliquant $*$ aux deux membres : $*(\mathrm{div}\, X) \cdot 1 = d * (\rho X)$ soit : $\mathrm{div}\, X \cdot \omega = d(-1)^{n-1}(\rho X) \lrcorner\, \omega = \mathrm{d}i(X)\omega$.

Dans un système de coordonnées différentiables quelconque, $\omega = |\overline{g}|^{1/2} dx^1 dx^2 \ldots dx^n$, et $i(X)\omega = |\overline{g}|^{1/2} \sum (-1)^j X^j dx^1 \ldots \widehat{dx^j} \ldots dx^n$; $\mathrm{d}i(X)\omega = |\overline{g}|^{1/2} \sum \frac{\partial}{\partial x^j}(|\overline{g}|^{1/2} X^j)\omega$, d'où l'expression, déjà obtenue par une autre voie au paragraphe précédent (dans E_n) :

$$\boxed{\mathrm{div}\, X = |\overline{g}|^{1/2} \sum \frac{\partial}{\partial x^j}(|\overline{g}|^{1/2} X^j)}\,.$$

En coordonnées rectilignes elle donne l'expression classique, et puisque l'actuelle définition comme la définition antérieure div X = Trace DX sont intrinsèques, elles coïncident. (on verra une troisième définition de la divergence : définition IX.3.D, et une généralisation : définition IX.4.C).

VII.11 – Harmoniques sphériques. Potentiel newtonien et multipoles

Le Laplacien Δ associé à la métrique d'un espace euclidien E_n, est invariant par le groupe d'isométries G de E_n. G opère sur les fonctions numériques $\mathfrak{C}^r$ de E_n par : $(g \cdot f)(x) = f(\overset{-1}{g}\, x)$. Δ commute avec les opérations de G ainsi qu'avec tout opérateur différentiel à coefficients constants sur E_n. Une fonction deux fois continûment différentiable est dite *harmonique* dans un ouvert U de E_n si $\Delta f = 0$ dans U. Il en résulte alors que f est nécessairement $\mathfrak{C}^\infty$ dans U. L'importance des fonctions harmoniques vient de ce que le potentiel newtonien d'une distribution de charges est, dans l'espace extérieur à ces charges, une fonction harmonique. Si $r = ||OP||$ est la fonction distance à l'origine dans E_n, de $r^2 = \sum (x^i)^2$ résultent : $\frac{\partial r}{\partial x^i} = \frac{x^i}{r}$ et $\Delta r^{-p} = p(p+2-n)r^{-p-2}$. Dès lors si $n \neq 2$ une fonction harmonique élémentaire dans $E_n - \{O\}$ est r^{-n+2}, soit $\frac{1}{r}$ dans E_3. C'est $\log r$ dans E_2.

On peut montrer que les seuls difféomorphismes d'un ouvert U de E_n sur un ouvert V qui transforment les fonctions harmoniques sur U en les fonctions harmoniques sur V sont les transformations conformes, qui pour $n \geqslant 3$ sont des produits d'isométries, d'homothéties et d'inversions. En particulier, l'inversion relative à la sphère unité de E_n qui s'écrit en coordonnées orthonormales : $x'^i = \frac{x^i}{r^2}$ transforme une fonction harmonique $U(x'^1, x'^2, \dots x'^n)$ sur $E_n - \{O\}$ en une autre $\boxed{\widetilde{U}(x^1, x^2, \dots x^n) = \frac{1}{r^{n-2}} U\left(\frac{x^1}{r^2}, \frac{x^2}{r^2}, \dots \frac{x^n}{r^2}\right)}$ que nous appellerons sa *duale* (transformation de Kelvin, 1847). Cette transformation est involutive : $\widetilde{\widetilde{U}} = U$. Par exemple, la fonction harmonique élémentaire $\frac{1}{r^{n-2}}$ est duale de la fonction constante égale à 1. Nous allons montrer que $\widetilde{U}$ est harmonique en considérant que l'inversion sert à définir de nouvelles coordonnées « curvilignes » $x^1, x^2, \dots x^n$ dans $E_n - \{O\}$ au lieu des coordonnées euclidiennes $x'^1, x'^2, \dots x'^n$. La métrique euclidienne ds'^2 et le laplacien Δ' s'écrivent : $ds'^2 = \frac{ds^2}{r^4}$ et

$$\Delta' = \overline{g}^{-1/2} \sum \frac{\partial}{\partial x^i}\left(\overline{g}^{-1/2} g^{ij} \frac{\partial}{\partial x^j}\right) = r^{2n} \sum \frac{\partial}{\partial x^i}\left(\frac{1}{r^{2n-4}} \cdot \frac{\partial}{\partial x^i}\right).$$

On peut écrire :

$$\begin{aligned} \frac{1}{r^{2n-4}} \frac{\partial U}{\partial x^i} &= \frac{1}{r^{n-2}} \cdot \frac{1}{r^{n-2}} \frac{\partial U}{\partial x^i} \\ &= \frac{1}{r^{n-2}} \left\{ \frac{\partial}{\partial x^i}\left(\frac{1}{r^{n-2}} U\right) - U \frac{\partial}{\partial x^i} \frac{1}{r^{n-2}} \right\}. \end{aligned}$$

Après simplification, on obtient l'expression du laplacien dans les nouvelles coordonnées :

$$\Delta' U = r^{n+2} \left\{ \sum \frac{\partial^2}{(\partial x^i)^2}\left(\frac{1}{r^{n-2}} U\right) - U \sum \frac{\partial^2}{(\partial x^i)^2} \frac{1}{r^{n-2}} \right\}.$$

Puisque $\dfrac{1}{r^{n-2}}$ est une fonction harmonique, on obtient finalement pour l'expression de Δ' : $\Delta'\mathrm{U} = r^{n+2}\sum \dfrac{\partial^2}{(\partial x^i)^2}\left(\dfrac{1}{r^{n-2}}\mathrm{U}\right)$ d'où il résulte que si U est une fonction harmonique, sa duale : $\widetilde{\mathrm{U}}(x^1, x^2, \dots x^n) = \dfrac{1}{r^{n-2}}\mathrm{U}\left(\dfrac{x^1}{r^2}, \dots \dfrac{x^n}{r^2}\right)$ l'est également.

Le groupe orthogonal $\mathrm{O}(n)$, donc aussi son algèbre de Lie $\mathfrak{O}(n)$, opèrent dans l'espace vectoriel P_l des fonctions polynomiales homogènes de degré l sur E_n. Cette représentation n'est pas irréductible. Par exemple si l est pair, les multiples scalaires de r^l forment une droite invariante de P_l. Le sous-espace $r^2\mathrm{P}_{l-2}$ et le sous-espace des *polynômes harmoniques de degré* l : $\mathrm{H}_l = \mathrm{Ker}\Delta$, sont appliqués, en eux-mêmes par les opérations de $\mathrm{O}(n)$, et décomposent P_l en la somme directe de deux $\mathrm{O}(n)$-modules : $\mathrm{P}_l = r^2\mathrm{P}_{l-2} \oplus \mathrm{H}_l$. La propriété : $r^2\mathrm{P}_{l-2} \cap \mathrm{H}_l = 0$ résulte du lemme VII.11.B ci-dessous.

Lemme VII.11.A. Si f est une fonction $\mathfrak{C}^2$ positivement homogène de degré l : $\Delta(r^2 f) = (2n + 4l)f + r^2\Delta f$.

Preuve : Calcul immédiat utilisant l'identité d'Euler : $\sum x_i \dfrac{\partial f}{\partial x^i} = lf$.

Lemme VII.11.B. Si $f \in \mathrm{P}_l$ et si λ est un nombre réel $\geqslant 0$, $\Delta(r^2 f) + \lambda f = 0$ entraîne $f = 0$. En particulier $\Delta(r^2 f) = 0$ entraîne $f = 0$. L'application de P_l dans lui-même $f \to \Delta(r^2 f)$ est un automorphisme linéaire qui applique en eux-mêmes H_l et $r^2\mathrm{P}_{l-2}$.

Preuve : d'après le lemme précédent :

$$\mathrm{O} = \Delta(r^2 f) + \lambda f = (2n + 4l + \lambda)f + r^2\Delta f.$$

Supposons le théorème démontré pour le degré $(l-2)$ et prenons le laplacien : $\Delta(r^2\Delta f) + (2n + 4l + \lambda)\Delta f = 0$. Δf est degré $(l-2)$ donc nul d'après l'hypothèse de récurrence. En reportant dans la première égalité il en résulte que $f = 0$.

Un calcul simple des dimensions des divers espaces P_{l-2}, P_l, et H_l (par exemple dim $H_l = (n+2l-2)\dfrac{(n+l-3)!}{(n-2)!}$) achève la preuve de la décomposition en somme directe.

Corollaire VII.11. L'espace vectoriel P_l des polynômes homogènes de degré l sur E_n se décompose en la somme directe de sous $O(n)$-modules : $P_l = H_l \oplus r^2 H_{l-2} \oplus r^4 H_{l-4} \oplus \dots r^{2[l/2]} H_{l-[l/2]}$. Il en résulte que la fonction obtenue par restriction d'un polynôme homogène à la sphère unité coïncide avec la restriction à cette sphère d'un polynôme harmonique (en général non homogène !).

Remarque : On peut démontrer que les représentations de $O(n)$ dans chacun des H_l sont irréductibles. Commes elles sont distinctes la décomposition précédente est donc unique.

La sphère unité S est une sous-variété de E_n. Elle hérite de la métrique induite qui détermine le laplacien Δ_s de S. Dans un système de coordonnées curvilignes (u^i) (§ VII.9) : $\Delta_S f = (\bar{g})^{-1/2} \sum \dfrac{\partial}{\partial u^i}\left(\bar{g}^{1/2} g^{ij} \dfrac{\partial f}{\partial u^j}\right)$.

Le système de coordonnées sphériques dans E_n où l'on sépare la coordonnée radiale r des coordonnées angulaires détermine sur S pour $r = 1$, des coordonnées curvilignes orthogonales. Dans une base orthonormale de E_n elles s'écrivent, avec $O \leqslant r$, $O \leqslant \theta_1 < 2\pi$, $O \leqslant \theta_i < \pi$ pour $i = 2, \dots n-1$:

$$
\begin{aligned}
x^1 &= r \sin \theta_{n-1} \sin \theta_{n-2} \dots \sin \theta_2 \sin \theta_1 \\
x^2 &= r \sin \theta_{n-1} \dots \sin \theta_2 \cos \theta_1 \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
x^n &= r \cos \theta_{n-1}
\end{aligned}
$$

et la métrique de E_n devient :

$$
ds^2 = dr^2 + r^2(d\theta_{n-1}^2 + \sin^2\theta_{n-1} d\theta_{n-2}^2 + \dots \sin^2\theta_{n-1} \dots \sin^2\theta_2 d\theta_1^2).
$$

Le laplacien de E_n en coordonnées sphériques s'écrit alors :

$$
\Delta = \frac{\partial^2}{\partial r^2} + \frac{n-1}{r} \cdot \frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_s = \frac{1}{r^{n-1}} \frac{\partial}{\partial r}\left(r^{n-1} \frac{\partial}{\partial r}\right) + \frac{1}{r^2}\Delta_s
$$

où Δ_s ne fait intervenir que les $(n-1)$ coordonnées angulaires. Δ_s est appelé le *Laplacien sphérique de* E_n-$\{\mathrm{O}\}$.

Cette formule permet de définir le laplacien de la sphère S de façon élémentaire. Si φ est une fonction deux fois continûment différentiable sur S, on peut lui associer une fonction positivement homogène de degré zéro h sur $\mathrm{E}_n - \{\mathrm{O}\}$ par $h(x) = \varphi\left(\frac{x}{||x||}\right)$. Cette fonction h est par définition indépendante de r et son laplacien est égal d'après la formule précédente à $\Delta h = \frac{1}{r^2}\Delta_s h$.

On peut donc prendre comme définition élémentaire du *Laplacien sur la sphère* S : $\Delta_{\mathrm{S}}\varphi = r^2 \Delta h$. Connaissant Δ_{S} on peut alors définir directement le laplacien sphérique Δ_s. Si f est deux fois continûment différentiable dans $\mathrm{E}_n - \{\mathrm{O}\}$, on peut écrire $f(x) = f(ru)$ avec $r = ||x||$ et $u = \frac{x}{||x||} \in \mathrm{S}$. En laissant r fixe, $f(ru)$ devient une fonction sur S. On obtient ainsi l'expression du laplacien sphérique de f :

$$\Delta_{\mathrm{S}} f(ru) = r^2 (\Delta f)_{r \text{ fixe}} = \Delta_s f.$$

Soit f une fonction $\mathfrak{C}^2$ positivement homogène de degré l sur E_n. La fonction $\mathrm{Y} = \frac{1}{r^l} f$ est positivement homogène de degré zéro. Elle est indépendante de la coordonnée radiale r et peut être identifiée à une fonction sur la sphère unité S. Un calcul direct, ou l'expression de Δ en coordonnées sphériques, donne alors pour le laplacien de $f = r^l \mathrm{Y}$:

$$\Delta f = l(n-2+l)\frac{1}{r^2} f + r^l \Delta_{\mathrm{Y}} = r^{l-2}[l(n-2+l)\mathrm{Y} + \Delta_{\mathrm{S}}\mathrm{Y}].$$

D'où le :

Théorème VII.11. Une fonction $\mathfrak{C}^2$ sur E_n positivement homogène de degré l, écrite $f = r^l \mathrm{Y}$, est harmonique *si et seulement si* la fonction Y sur la sphère unité S est une fonction propre du laplacien Δ_{S} pour la valeur propre $-l(n-2+l)$. Une telle fonction Y sur S est appelée un *harmonique sphérique* associé à f. Y est évidemment associé de la même façon à la fonction harmonique duale de f : $\widetilde{f} = \frac{1}{r^{n-2+l}}\mathrm{Y}$ positivement homogène de degré : $-(n-2+l)$.

Pour un l fixé, les harmoniques sphériques forment un espace vectoriel $\mathfrak{H}_l$ qui est naturellement muni d'une représentation du groupe orthogonal $\mathrm{O}(n)$ et de son algèbre de Lie $\mathfrak{O}(n)$.

Remarque : l'ensemble des valeurs propres du laplacien d'une variété riemannienne compacte est appelé le *spectre* de la variété. Par des considérations de théorie des groupes (les représentations de $\mathfrak{O}(n)$ dans les $\mathfrak{H}_l$ sont irréductibles) ou d'analyse, on peut montrer qu'il n'y a pas d'autres harmoniques sphériques que les traces sur S des polynômes harmoniques (exemple 1 ci-dessous). Le paramètre l est donc toujours entier. Nous vérifierons ce fait pour $n = 3$. Le spectre de S est donc formé des entiers $\{-l(l+n-2)\,;\ l = 0, 1, 2, \ldots\}$.

Exemples :

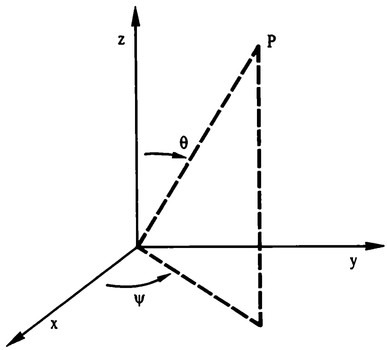
1) un polynôme harmonique homogène P de degré l détermine un harmonique sphérique $\mathrm{Y}_l = \dfrac{1}{r^l}\mathrm{P}$ dit de « *degré l* » et qui est aussi associé à $\widetilde{\mathrm{P}} = \dfrac{\mathrm{P}}{r^{n+2l-2}}$. L'application $\mathrm{P} \to \mathrm{Y}$ est un isomorphisme de $\mathrm{O}(n)$-modules de H_l sur $\mathfrak{H}_l$.

2) les harmoniques sphériques sur la $(n-1)$-sphère unité S de E_n généralisent les fonctions trigonométriques qui correspondent au cas $n = 2$, S étant alors le cercle S_1. Dans ce cas, en effet, les fonctions propres de $\Delta_{\mathrm{S}_1} = \dfrac{\partial^2}{\partial\theta^2}$ pour les valeurs propres $-l^2$, sont les solutions de $\dfrac{\partial^2\varphi}{\partial\theta^2} + l^2\varphi = 0$.

Le cas de très loin le plus important dans les applications est celui de l'espace euclidien ordinaire E_3. En nous conformant à l'usage, nous prenons pour coordonnées polaires dans E_3 : φ et θ tels $x = r\sin\theta\cos\varphi$, $u = r\sin\theta\sin\varphi$, $z = r\cos\theta$.

En coordonnées polaires, le laplacien de E_3 s'écrit : $\Delta = \dfrac{1}{r^2}\dfrac{\partial}{\partial r}\left(r^2\dfrac{\partial}{\partial r}\right) + \dfrac{1}{r^2}\Delta_s$ avec $\Delta_s = \dfrac{1}{\sin\theta}\dfrac{\partial}{\partial\theta}\left(\sin\theta\dfrac{\partial}{\partial\theta}\right) + \dfrac{1}{\sin^2\theta}\dfrac{\partial^2}{\partial\varphi^2}$.

Cet opérateur Δ_s dans E_3 a une interprétation fort élégante en termes de l'algèbre de Lie $\mathfrak{O}(3)$ du groupe orthogonal. Rappelons l'expression des opérateurs de rotation infinitésimale. Dans E_n



rapporté à une base orthonormale, une rotation $r_{jk}(\alpha)$ agit sur la fonction f par :

$$(r_{jk}(\alpha)\cdot f)(x) = f(r_{jk}(-\alpha)x)$$
$$= f(x^1,\dots x^j\cos\alpha + x^k\sin\alpha, x^{k+1},\dots -x^j\sin\alpha + x^k\cos\alpha,\dots n^n)$$

et $$\frac{d}{d\alpha}(r_{jk}(\alpha)\cdot f)_{\alpha=0} = -\left(x^j\frac{\partial}{\partial x^k} - x^k\frac{\partial}{\partial x^j}\right)f.$$

Les opérateurs de rotation infinitésimale autour des trois axes Ox, Oy, Oz de E_3 sont ainsi : $L_x = -\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$; $L_y = -\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$, $L_z = -\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$. r restant constant dans les rotations autour de O, ces trois opérateurs peuvent s'exprimer à l'aide de θ et φ :

$$L_x = \sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\,\cos\varphi\frac{\partial}{\partial\varphi};$$
$$L_y = -\cot\varphi\frac{\partial}{\partial\theta} + \cot\theta\,\sin\varphi\frac{\partial}{\partial\varphi};\ L_z = -\frac{\partial}{\partial\varphi}.$$

Ces opérateurs forment une base de l'algèbre de Lie $\mathfrak{O}(3)$ qui est complètement déterminée par leurs règles de commutation :

$$[L_x, L_y] = L_xL_y - L_yL_x = L_z;\ [L_yL_z] = L_x;\ [L_z, L_x] = L_y.$$

Il en résulte que la somme des carrés de ces opérateurs $L^2 = L_x^2 + L_y^2 + L_z^2$ commute avec chacun d'eux donc avec tous les opérateurs de $\mathfrak{O}(3)$, L^2 est *l'opérateur de Casimir* de l'algèbre de Lie $\mathfrak{O}(3)$ (au facteur 1/2 près). Un calcul direct montre que :

Proposition VII.11.A

L'opérateur de Casimir, appelé aussi carré du moment angulaire de E_3 : $L^2 = L_x^2 + L_y^2 + L_3^2$ est identique au laplacien sphérique Δ_s de E_3.

Corollaire : L'opérateur de Casimir de $\mathfrak{O}(3)$ ayant pour valeurs propres les nombres $l(l+1)$, l'entier $= 0, 1, \dots$ dans les représentations du groupe $O(3)$, ces nombres forment le spectre de la sphère S_2.

Les harmoniques sphériques sur la sphère unité S de E_3 sont les fonctions $Y(\theta, \varphi)$ *finies et de classe* $\mathfrak{C}^2$ *sur* S solutions de l'équation :

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\,\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2} = -l(l+1)Y$$

où le paramètre l est à déterminer. On peut chercher *des* solutions à variables séparées : $Y(\theta, \varphi) = T(\theta)F(\varphi)$ ou solutions axiales privilégiant l'axe Oz. On obtient :

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dT}{d\theta}\right), T(\theta)) + l(l+1)\sin^2\theta = -\frac{F''(\varphi)}{F(\varphi)} = a$$

où a est nécessairement une constante, F devant être une fonction univalente de période 2Π, $a = -m^2$, m entier, et $F(\varphi) = a_m \cos m\varphi + b_m \sin m\varphi$. On peut aussi prendre $F_m(\varphi) = C_m e^{im\varphi}$ et $F_{-m} = C_{-m}e^{-im\varphi}$ en prenant ultérieurement les combinaisons réelles de ces fonctions. T est solution de :

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\sin\theta\frac{dT}{d\theta} + \left\{l(l+1) - \frac{m^2}{\sin^2\theta}\right\}T = 0$$

équation qui prend une allure plus humaine avec le changement de variable défini par $u = \cos\theta$, d'où $\dfrac{d}{d\theta} = -\sin\theta\dfrac{d}{du}$. On obtient :

$$\frac{d}{du}\left\{(1-u^2)\frac{dT}{du}\right\} + \left\{l(l+1) - \frac{m^2}{1-u^2}\right\}T = 0.$$

Pour $m \neq 0$, c'est l'*équation de Legendre associée* L_m, pour $m = 0$, c'est l'*équation de Legendre classique* L_0. Elle a pour points singuliers $u = \pm 1$.

Pour $m = 0$, l'équation L_0 : $(1-u^2)\dfrac{d^2\mathrm{T}}{du^2} - 2u\dfrac{d\mathrm{T}}{du} + l(l+1)\mathrm{T} = 0$ ne peut avoir de solutions qui restent finies en $u = -1$ et en $u = 1$ *que* si le paramètre l est un entier $\geqslant 0$ (ainsi que le prouvent les développements en série aux voisinages de ces points). Si l'on remarque que la fonction $t = (u^2-1)^l$ satisfait à $(1-u^2)t' + 2lut = 0$, sa dérivée l-ème est une solution de l'équation obtenue en dérivant $(l+1)$ fois la relation précédente, qui n'est autre que L_0. l étant fixé, cette solution dont nous allons voir qu'à un coefficient près elle est l'unique solution régulière en $u = \pm 1$ est le *polynôme de Legendre de degré* l donné par la formule de *Rodrigues* : $\mathrm{P}_l(u) = \dfrac{1}{2^l \cdot l!}\dfrac{d^l}{du^l}(u^2-1)^l$.

Si Q est toute autre solution, linéairement indépendante de P_l, on a : $\mathrm{QP}'_l - \mathrm{P}_l\mathrm{Q}' = \dfrac{a}{1-u^2}$ où a est une constante non nulle. Il en résulte que Q ne peut être régulière en $u = \pm 1$. Les polynômes de Legendre sont orthogonaux sur $[-1, +1]$ et ont une quantité impressionnante de propriétés diverses.

Pour $u \neq 0$, on allège les singularités en $u = \pm 1$ par un changement de fonction : $\mathrm{T}(u) = (1-u^2)^{m/2}\mathrm{U}(u)$ et l'équation L_m devient :

$$(1-u^2)\frac{d^2\mathrm{R}}{du^2} - 2(m+1)u\frac{d\mathrm{R}}{du} + \{l(l+1) - m(m+1)\}\mathrm{R} = 0.$$

Or, l'équation L_0 différentiée m fois s'écrit :

$$(1-u^2)\frac{d^2}{du^2}\left(\frac{d^m\mathrm{T}}{du^m}\right) - 2(m+1)u\frac{d}{du}\left(\frac{d^m\mathrm{T}}{du^m}\right) + \{l(l+1) - m(m+1)\}\frac{d^m\mathrm{T}}{du^m} = 0.$$

La comparaison montre que la « fonction de Legendre associée » :

$$\mathrm{P}_l^m(u) = (1-u^2)^{m/2}\frac{d^m}{du^m}\mathrm{P}_l(u) = \frac{1}{2^l l!}(1-u^2)^{m/2}\frac{d^{m+l}}{du^{m+l}}(u^2-1)^l$$

pour $0 \leqslant m \leqslant l$ fixés, est la seule solution de l'équation de Legendre associée L_m qui reste finie pour $u = \pm 1$.

On obtient ainsi les $(2l+1)$ harmoniques sphériques formant une base de $\mathfrak{H}_l$:

$$\{\mathrm{P}_l(\cos\theta), \mathrm{P}_l^m(\cos\theta)\cos m\varphi, \mathrm{P}_l^m(\cos\theta)\sin m\varphi;\ m = 1, 2, \ldots l\}.$$

$\mathrm{P}_l(\cos\theta)$, indépendant de φ, est un harmonique sphérique « à symétrie axiale ». Afin d'exprimer de façon simple l'action des opérateurs de rotation dans les espaces vectoriels de fonctions continues sur S, on les complexifie, autrement dit, on considère les fonctions continues sur S à valeurs dans $\mathbb{C}$, et on y définit un produit scalaire hermitien par :

$$(f|h) = \int_{\mathrm{S}} \overline{f} h d\sigma = \int_{\mathrm{S}} \overline{f}(\theta, \varphi) h(\theta, \varphi) \sin\theta d\varphi d\theta$$

l'élément d'aire $d\sigma$ de S ayant pour expression en coordonnées polaires $d\sigma = \sin\theta d\varphi d\theta$.

Les rotations de SO(3) sont alors représentés par des opérateurs unitaires dans chacun des espaces $\mathfrak{H}_l \otimes \mathbb{C}$.

Les exponentielles $e^{im\varphi}$ prennent la place des $\cos m\varphi$ et $\sin m\varphi$, ce qui amène à faire varier le paramètre m de $-l$ à $+l$. La base précédemment trouvée, qui n'est qu'orthogonale, est remplacée par la base *orthonormale* suivante, dont les éléments sont appelés les *fonctions sphériques de Laplace*. Cette base est *canoniquement déterminée* par le système d'axes $\mathrm{O}xyz$:

$$\mathrm{Y}_{lm}(\theta, \varphi) = (-1)^m \left(\frac{2l+1}{4\Pi}\right)^{1/2} \left(\frac{(l-m)!}{(l+m!)}\right)^{1/2} e^{im\varphi}\mathrm{P}_l^m(\cos\theta) \text{ pour } \mathrm{O} \leqslant m \leqslant l$$

$$\mathrm{Y}_{lm}(\theta, \varphi) = \left(\frac{2l+1}{4\Pi}\right)^{1/2} \left(\frac{(l-|m|)!}{(l+|m|)!}\right)^{1/2} e^{im\varphi}\mathrm{P}_l^{|m|}(\cos\theta) \text{ pour } -l \leqslant m \leqslant 0.$$

On a : $\mathrm{Y}_{lo}(\theta, \varphi) = \left(\frac{2l+1}{4\Pi}\right)^{1/2} \mathrm{P}_l(\cos\theta)$; $(\mathrm{Y}_{lm}|\mathrm{Y}_{l'm'}) = \delta_{ll'}\delta_{mm'}$, et $\overline{\mathrm{Y}}_{lm} = (-1)^m \mathrm{Y}_{l,-m}$.

Exemples : $\mathrm{Y}_{1,\pm 1} = \pm\left(\frac{3}{8\Pi}\right)^{1/2} \sin\theta e^{\pm i\varphi}$,

$$\mathrm{Y}_{2,\pm 1} = \pm\left(\frac{15}{8\Pi}\right)^{1/2} \sin\theta\,\cos\theta e^{\pm i\varphi},$$

$$\mathrm{Y}_{2,\pm 2} = \frac{1}{4}\left(\frac{15}{2\Pi}\right)^{1/2} \sin^2\theta^{\pm 2i\varphi}.$$

Remarques : Les fonctions de Laplace comme les fonctions de Legendre font partie des « fonctions spéciales ». On donne de ces dernières deux définitions. La première appelle *fonction spéciale* toute fonction numérique sur un groupe G ou un espace homogène G/H (ici $\mathrm{S} = \mathrm{SO}(3)/\mathrm{SO}(2)$) qui apparaît comme coefficient $\mathrm{U}_{ij}(g)$ des matrices $\mathrm{U}(g)$ d'une représentation linéaire de dimension finie de G. La seconde, moins sérieuse, appelle fonction spéciale toute fonction qui a fait l'objet d'une publication !

L'activité mathématique, qui cherche en permanence à mieux « expliquer », trouve dans le domaine des fonctions spéciales un champ de manœuvres, riche et utile : on réussit, en effet , à « expliquer » les identités et formules classiques en montrant qu'elles ne font qu'exprimer la simple propriété : $\mathrm{U}(g' \cdot g) = \mathrm{U}(g')\mathrm{U}(g)$. En particulier, la démonstration élémentaire qui va suivre peut être avantageusement remplacée par des considérations plus élégantes de théorie des groupes.

Considérons deux systèmes d'axes ($\mathrm{O}xyz$ de coordonnées polaires θ, φ), ($\mathrm{O}x'y'z'$ de coordonnées polaires θ', φ'). Il leur correspond dans $\mathfrak{H}_l \otimes \mathbb{C}$ deux bases orthonormales de Laplace : les $\mathrm{Y}_{lm}(\theta,\varphi)$ et les $\mathrm{Y}_{lm'}(\theta',\varphi')$. Si R est la rotation qui fait passer de $\mathrm{O}x'y'z'$ à $\mathrm{O}xyz$, elle est représentée dans $\mathfrak{H}_l \otimes \mathbb{C}$ par une matrice D, et $\mathrm{Y}_{lm}(\theta,\varphi) = \sum \mathrm{D}_{mm'} \cdot \mathrm{Y}_{lm'}(\theta',\varphi')$.

$\mathrm{D} = |\mathrm{D}_{mm'}|$ est unitaire par construction. Son inverse est donc $\overset{-1}{\mathrm{D}} = {}^t\,\overline{\mathrm{D}}$, d'où : $\mathrm{Y}_{lm'}(\theta',\varphi') = \sum \overline{\mathrm{D}}_{mm'} \cdot \mathrm{Y}_{lm}(\theta,\varphi)$. Mais ces relations entre vecteurs de $\mathfrak{H}_l \otimes \mathbb{C}$ sont des relations entre fonctions sur S et nous allons considérer les *valeurs* de ces fonctions Y en deux points particuliers de S définis par les vecteurs unitaires $\overline{r}_1((\theta_1,\varphi_1)$ dans $\mathrm{O}xyz$, (θ'_1,φ'_1) dans $\mathrm{O}x'y'z'$) et $\overrightarrow{r}_2((\theta_2,\varphi_2)$ et $(\theta'_2,\varphi'_2))$:

$$\mathrm{Y}_{lm}(\theta_2,\varphi_2) = \sum \mathrm{D}_{mm'}, \mathrm{Y}_{lm'}(\theta'_2,\varphi'_2),$$

$$\mathrm{Y}_{lm'}(\theta'_1,\varphi'_1) = \sum \overline{\mathrm{D}}_{mm'}, \mathrm{Y}_{lm}(\theta_1,\varphi_1).$$

Supposons maintenant le système d'axes $Ox'y'z'$ choisi de telle sorte que $\overline{r}_2$ soit le vecteur unitaire de Oz'. On a dans ce cas : $\theta_2' = 0$, $\theta_1' = \gamma$ angle de $\overline{r}_1$ et $\overline{r}_2$. La deuxième égalité donne pour $m' = 0$:

$$\mathrm{Y}_{lo}(\gamma, \varphi_1') = \sqrt{\frac{2l+1}{4\Pi}}\mathrm{P}_l(\cos\gamma) = \sum \overline{\mathrm{D}}_{mo}\mathrm{Y}_{lm}(\theta_1, \varphi_1).$$

Mais pour $\theta_2' = 0$, les fonctions $\mathrm{Y}_{lm'}(\theta_2', \varphi_2')$ sont nulles si $m' \neq 0$ car $\mathrm{P}_{m,l}(\cos\theta)$ contient $\sin\theta$ en facteur pour $m \neq 0$. La première égalité s'écrit donc :

$$\mathrm{Y}_{lm}(\theta_2, \varphi_2) = \mathrm{D}_{mo}\mathrm{Y}_{lo}(\theta, \varphi') = \mathrm{D}_{mo}\sqrt{\frac{2l+1}{4\Pi}}.$$

En reportant la valeur trouvée de D_{mo} dans l'égalité précédente, on obtient le célèbre *théorème d'addition des harmoniques sphériques*, aux applications nombreuses, qui généralise la formule fondamentale de la trigonométrie sphérique :

$$\boxed{\mathrm{P}_l(\cos\gamma) = \frac{4\Pi}{2l+1}\sum_{m=-l}^{l} \overline{\mathrm{Y}}_{lm}(\theta_2, \varphi_2)\mathrm{Y}_{lm}(\theta_1, \varphi_1)}.$$

Nous allons montrer ci-dessous comment elle permet d'obtenir le développement multipolaire des potentiels newtoniens.

Considérons d'abord le potentiel exercé par une charge unité située en un point M_0 de coordonnées sphériques $(r_0, \theta_0, \varphi_0)$ et cartésiennes (ξ, η, ζ) sur un point P de coordonnées (r, θ, φ) et (x, y, z).

Si $\mathrm{R} = ||\mathrm{M}_0\mathrm{P}||$,

$$\begin{aligned}\mathrm{R}^2 = ||\mathrm{M}_0\mathrm{P}||^2 &= (r\sin\theta\cos\varphi - r_0\sin\theta_0\cos\varphi_0)^2 \\ &+ (r\sin\theta\sin\varphi - r_0\sin\theta_0\sin\varphi_0)^2 + (r\cos\theta - r_0\cos\theta_0)^2 \\ &= r^2 - 2rr_0\cos\gamma + r_0^2 = ||\vec{r} - \vec{r}_0||^2\end{aligned}$$

où γ est l'angle des vecteurs $(\vec{r}_0, \vec{r})$. Rappelons la formule de la trigonométrie sphérique :

$$\begin{aligned}\cos\gamma &= \frac{\vec{r}_0 \cdot \vec{r}}{r \cdot r_0} = \frac{x\xi + y\eta + z\zeta}{rr_0} \\ &= \cos\theta \, \cos\theta_0 + \sin\theta \, \sin\theta_0 \, \cos(\varphi - \varphi_0)\end{aligned}$$

$\dfrac{1}{\mathrm{R}} = \dfrac{1}{||\vec{r} - \vec{r}_0||} = \dfrac{1}{(r^2 - 2urr_0 + r_0^2)^{1/2}}$ où l'on peut mettre soit $\dfrac{1}{r}$ soit $\dfrac{1}{r_0}$ en facteur en vue du développement de $\dfrac{1}{\mathrm{R}}$ suivant les puissances de $\dfrac{1}{r}$ ou de $\dfrac{1}{r_0}$, qui s'obtient à l'aide du développement du binôme :

$$\frac{1}{(1-a)^{1/2}} = 1 + \frac{1}{2}a + \frac{1 \cdot 3}{2 \cdot 4}a^2 + \ldots \frac{1 \cdot 3 \ldots (2n-1)}{2 \cdot 4 \ldots 2n}a^n + \ldots$$

valable pour $|a| \leqslant 1$. En posant $a = 2ut - t^2$, on est assuré, puisque $|u| \leqslant 1$, que $||2ut - t^2| < 1$, dès que $1 - \sqrt{2} < t < 1$, donc pour $|t|$ assez petit. D'où :

$$\begin{aligned}\frac{1}{\sqrt{1 - 2ut + t^2}} &= 1 + \frac{1}{2}(2ut - t^2) + \ldots \\ &\quad + \ldots \frac{1 \cdot 3 \ldots (2n-1)}{2 \cdot 4 \ldots 2n}(2ut - t^2)^n + \ldots\end{aligned}$$

On peut chasser les parenthèses et réordonner suivant les puissances croissantes de t :

$$\frac{1}{\sqrt{1 - 2ut + t^2}} = \mathrm{P}_0(u) + \mathrm{P}_1(u)t + \mathrm{P}_2(u)t^2 + \ldots \mathrm{P}_n(u)t^n + \ldots$$

Le coefficient de t^n est un polynôme en u de degré n et de la parité de n, appelé *par définition le polynôme de Legendre de degré* n, et déjà considéré ci-dessus où l'on a donné son expression due à Rodrigues (1816).

$$\begin{aligned}&\mathrm{P}_0(u) = 1;\ \mathrm{P}_1(u) = u;\ \mathrm{P}_2(u) = \frac{1}{2}(3u^2 - 1); \\ &\mathrm{P}_3(u) = \frac{1}{2}(5u^3 - 3u);\ \mathrm{P}_4(u) = \frac{1}{8}(35u^4 - 30u^2 + 3);\ \ldots\end{aligned}$$

On démontre aisément, par récurrence par exemple, que les polynômes de Legendre s'expriment à l'aide des dérivées de la fonction $\frac{1}{r}$:

$$\frac{\partial^l}{\partial z^l}\left(\frac{1}{r}\right) = (-1)^l \frac{l!}{r^{l+1}} \mathrm{P}_l\left(\frac{z}{r}\right)$$

$r^n r_0^n \mathrm{P}_n(\cos\gamma) = r^n r_0^n \mathrm{P}_n\left(\frac{x\xi + y\eta + z\zeta}{r r_0}\right)$ est un polynôme homogène de degré n aussi bien par rapport aux variables (x, y, z) qu'aux variables (ξ, η, ζ) puisqu'il est symétrique par rapport à ces deux groupes de variables.

La fonction $\frac{1}{||\vec{r} - \vec{r}_0||}$ symétrique en $\vec{r}$ et $\vec{r}_0$ peut être développée suivant les puissances de r au voisinage de l'origine, sous la forme d'une série de polynômes harmoniques homogènes en (x, y, z) convergente pour $r < r_0$, dépendant des paramètres (ξ, η, ζ) :

$$\frac{1}{||\vec{r} - \vec{r}_0||} = \sum_{n=0}^{\infty} \frac{r^n}{r^{0}_{n+1}} \mathrm{P}_n(\cos\gamma) = \sum \mathrm{U}_n(x, y, z)$$

ou peut être développée « au voisinage de l'infini », suivant les puissances de $\frac{1}{r}$, pour $r > r_0$, en une série de fonctions harmoniques homogènes en (x, y, z) :

$$\frac{1}{||\vec{r} - \vec{r}_0||} = \sum_{n=0}^{\infty} \frac{r_0^n}{r^{n+1}} \mathrm{P}_n(\cos\gamma) = \sum \Phi_n(x, y, z).$$

Considérons maintenant un ensemble fini de charges ponctuelles positives ou négatives : $\{q_k \text{ en } \mathrm{M}_k\}$. Le potentiel qu'elles exercent en un point P s'écrit, en notant $\vec{r}_k = \overrightarrow{\mathrm{OM}}_k$, $r_k = ||\vec{r}_k||$, $\vec{r} = \overrightarrow{\mathrm{OP}}$, $r = ||\vec{r}||$:

$$\Phi(\mathrm{P}) = \sum_k \frac{q_k}{||\vec{r} - \vec{r}_k||} = \frac{1}{r} \sum_k \frac{q_k}{\left[1 - 2\frac{r_k}{r}\cos\gamma_k + \left(\frac{r_k}{r}\right)^2\right]^{1/2}}$$

et pour $r > \sup(r_k)$:

$$\Phi(\mathrm{P}) = \sum_1^\infty \Phi_j(\mathrm{P}) = \frac{1}{r}\sum_{p=0}^{\infty}\left[\sum_k q_k\left(\frac{r_k}{r}\right)^p \mathrm{P}_p(\cos\gamma_k)\right]$$
$$= \frac{1}{r}(\sum q_k) + \frac{1}{r^2}(\sum_k q_k\vec{r}_k)\cdot\frac{\vec{r}}{r} + \dots$$

Le vecteur $\sum q_k\vec{r}_k = \sum q_k\overrightarrow{\mathrm{OM}}_k$ est le moment du système de charges par rapport à 0. Si la somme $\sum q_k$ est nulle, ce vecteur est indépendant du choix de l'origine et est appelé le *moment dipolaire* $\overrightarrow{\mathrm{D}}$ du système de charges. Le développement du potentiel commence alors avec le terme $\Phi_2 = \dfrac{\overrightarrow{\mathrm{D}}\cdot\vec{r}}{r^3} = -\left(\overrightarrow{\mathrm{D}}\cdot\mathrm{grad}\dfrac{1}{r}\right)$.

Calculons le terme

$$\Phi_3(\vec{r}) = \frac{1}{r^3}\sum q_k r_k^2 \mathrm{P}_2(\cos\gamma_k) = \frac{1}{2r^5}\sum q_k(3(\vec{r}_k\cdot\vec{r})^2 - r_k^2 r^2).$$

Les dérivées secondes de $\dfrac{1}{r}$ forment la matrice de fonctions harmoniques $\mathrm{R}_{(2)}$:

$$\mathrm{R}_{(2)} = \frac{1}{r^5}\begin{vmatrix} 3x^2 - r^2 & 3xy & 3xz \\ 3xy & 3y^2 - r^2 & 3yz \\ 3xz & 3yz & 3z^2 - r^2 \end{vmatrix}$$

de trace nulle puisque $\dfrac{1}{r}$ est harmonique. $\Phi_3(r)$ est égal au $\dfrac{1}{6}$ du produit doublement contracté du tenseur $\mathrm{R}_{(2)}$ et du tenseur Q suivant, appelé le *moment quadrupolaire* des charges :

$$\mathrm{Q} = \sum_k q_k\begin{vmatrix} 3x_k^2 - r_k^2 & 3x_k y_k & 3x_k z_k \\ 3x_k y_k & 3y_k^2 - r_k^2 & 3y_k z_k \\ 3x_k z_k & 3y_k z_k & 3z_k^2 - r_k^2 \end{vmatrix}.$$

Lorsque le système se réduit à deux charges opposées $-q$ et q placées respectivement en des points M_1 et $\mathrm{M}_2 = \mathrm{M}_1 + h\vec{u}$ (où $||\vec{u}|| = 1$), leur potentiel en P : $qh\cdot\dfrac{1}{h}\left(\dfrac{1}{||\vec{r}-\vec{r}_1 - h\vec{u}||} - \dfrac{1}{||\vec{r}-\vec{r}_1||}\right)$

tend vers une limite lorsqu'on fait tendre h vers zéro et q vers $(+\infty)$ de telle sorte que qh reste fixe égal à m. Cette limite est appelée le *potentiel du dipole élémentaire* situé en M_1 de moment : $q(\vec{r}_1 + h\vec{u}) - q\vec{r}_1 = m\vec{u}$. L'expression ci-dessus montre qu'il est obtenu en dérivant relativement au vecteur $m\vec{u}$ et à la variable $\vec{r}_1$ la fonction $\dfrac{1}{||\vec{r}-\vec{r}_1||}$. C'est l'opposé de la dérivée par rapport à la variable $\vec{r}$. On a donc pour ce potentiel l'expression :

$$-\mathrm{D}^{(r)}_{m\vec{u}}\frac{1}{||\vec{r}-\vec{r}_1||} = -(m\vec{u}\cdot\nabla)^{(r)}\frac{1}{||\vec{r}-\vec{r}_1||}.$$

En particulier le potentiel d'un dipole de moment $\vec{m}$ situé à l'origine est $-\mathrm{D}_{\vec{m}}\left(\dfrac{1}{r}\right)$. Si $\vec{m}$ est le vecteur unitaire de l'axe $\mathrm{O}z$, c'est $-\nabla_z\dfrac{1}{r} = \dfrac{\mathrm{P}_1(\cos\theta)}{r^2}$.

On peut ainsi considérer que par l'application de p opérateurs de dérivation à l'origine : $(-1)^p\mathrm{D}_{\vec{m}_p}\mathrm{D}_{\vec{m}_{p-1}}\ldots\mathrm{D}_{\vec{m}_1}$ à la fonction $\dfrac{1}{r}$, on obtient une fonction harmonique représentant le potentiel d'un multipole élémentaire d'ordre p, placé à l'origine, de multimoment $(\vec{m}_1,\vec{m}_2,\ldots\vec{m}_p)$.

La disposition des moments $\vec{m}_1,\ldots\vec{m}_p$ détermine les propriétés de symétrie du multipole et de son potentiel. Les applications font souvent usage des opérateurs de dérivations complexes :$\nabla_\pm = \dfrac{\partial}{\partial x} \pm i\dfrac{\partial}{\partial y}$. On obtient ainsi par un calcul simple : $(-1)^{p+q}\nabla^p_\pm\nabla^q_z\dfrac{1}{r} = \dfrac{1}{r^{p+q+1}}q!e^{\pm ip\varphi}\mathrm{P}^p_{p+q}(\cos\theta)$ qui fournit un exemple de potentiel d'un multipole d'ordre $p+q$.

Revenons au développement de Legendre du potentiel $\dfrac{1}{||\vec{r}-\vec{r}_0||}$ exercé par une charge unité placée en $\mathrm{M}_0(\vec{r}_0;\theta_0;\varphi_0)$ en un point $\mathrm{P}(\vec{r};\theta;\varphi)$.

La formule d'addition des harmoniques sphériques permet de séparer dans ce développement la dépendance de M_0 de celle de P en écrivant pour $r > 0$:

$$\frac{1}{||\vec{r}-\vec{r}_0||} = \sum \frac{r_0^l}{r^{l+1}} \mathrm{P}_1(\cos\gamma)$$
$$= \sum_l \frac{4\Pi}{2l+1} \Big(\sum_{m=-l}^{+1} \overline{\mathrm{Y}}_{lm}(\theta_0,\varphi_0) r_0^l \mathrm{Y}_{lm}(\theta;\varphi) \cdot \frac{1}{r^{l+1}}.$$

Etant donnée une distribution de charges de densité ρ dans un domaine intérieur à une sphère de centre O et de rayon a, on peut développer la fonction potentiel $\Phi(\vec{r})$ pour $r > a$ en utilisant l'expression précédente.

$$\Phi(\vec{r}) = \int \frac{\rho(\vec{r}_0)}{||\vec{r}-\vec{r}_0||} d\omega_0 = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \int \mathrm{P}_n(\cos\gamma) r_0^n \rho(\vec{r}_0) d\omega_0.$$

Soit $\mathrm{Q}_{lm} = \int \overline{\mathrm{Y}}_{lm}(\theta,\varphi) r^l \rho(r,\theta,\varphi) r^2 \sin\theta \, \mathrm{d}r \, d\theta d\varphi$. C'est *le moment multipolaire* de la distribution de charges, de type (lm). Le développement du potentiel obtenu :

$$\Phi(\vec{r}) = \int \frac{\rho(\vec{r}_0) d\omega_0}{||\vec{r}-\vec{r}_0||} = \sum_{n=0}^{\infty} \frac{4\Pi}{l+1} \left(\sum_{m=-l}^{+l} \mathrm{Q}_{lm} \mathrm{Y}_{lm}(\theta,\varphi) \right) \frac{1}{r^{l+1}}$$

donne une expression du comportement angulaire des termes du développement de Taylor en fonctions harmoniques homogènes :

$$\Phi(\vec{r}) = \Phi_1(x,y,z) + \ldots \Phi_n(x,y,z) + \ldots$$

$$\Phi_1(\vec{r}) = \frac{1}{r} \int_\rho(\vec{r}_0) d\omega_0 = \frac{\mathrm{M}}{r} \text{ où M est la somme des charges}$$
$$\Phi_2(\vec{r}) = (\int_\rho(\vec{r}_0)\vec{r}_0 d\omega_0) \cdot \frac{\vec{r}}{r^3} = -\overrightarrow{\mathrm{D}} \cdot \mathrm{grad} \frac{1}{r} \text{ où le vecteur } \overrightarrow{\mathrm{D}} \text{ est}$$
le *potentiel dipolaire* de la distribution.

Plus généralement, $\Phi_n(\vec{r})$ est à un facteur numérique près le contracté sur tous les indices d'un tenseur symétrique $\mathrm{M}_{(n)}$ appelé *potentiel multipolaire d'ordre* n, avec le champ de tenseurs symétriques $\mathrm{E}_{(n)}$ que forment les dérivées d'ordre n de $\frac{1}{r}$ dont *toutes les traces sont nulles* ($\frac{1}{r}$ est harmonique) :

$$(\mathrm{E}_{(n)} i_1 i_2 \ldots i_n = \frac{\partial}{\partial x^{i_1}} \frac{\partial}{\partial x^{i_2}} \cdots \frac{\partial}{\partial x^{i_n}} \left(\frac{1}{r}\right).$$

Il en résulte que toute distribution de charges peut être *remplacée* par un ensemble de multipoles placées en un point O qui donne le *même potentiel* pour r assez grand.

VII.12 – Orientation des espaces pseudoeuclidiens. Relativité restreinte

Soit E un espace vectoriel réel de dimension n muni d'une forme quadratique q non-dégénérée. On distingue dans E les sous-espaces F strictement positifs pour lesquels $q(x) > 0$ quel que soit le vecteur x non nul de F, strictement négatifs : $q(x) < 0 \forall x \neq 0 \in \mathrm{F}$, isotropes, ou singuliers, si la restriction q_{F} de q à F est dégénérée, totalement isotropes, ou totalement singuliers si $q_{\mathrm{F}} \equiv 0$. On démontre immédiatement que les sous-espaces strictement positifs maximaux P ont tous la même dimension r, et que les sous-espaces strictement négatifs maximaux ont tous la même dimension s. (r, s) est l'indice de (E, q) avec $r + s = n$, et dans ce paragraphe, on suppose r et s tous deux non nuls : (E, q) est l'espace pseudoeuclidien $\mathrm{E}_{r,3}$. Soit $g(x, y) = (x|y) = \dfrac{1}{2}\{q(x+y) - q(x) - q(y)\}$ la forme bilinéaire symétrique associée à q. Elle est non-dégénérée et définit le produit scalaire des vecteurs de (E, q).

Si $x_1, x_2, \ldots x_p$ sont p vecteurs de E on appelle indicateur $\varepsilon(x_1, x_2, \ldots x_p)$ le signe du déterminant $\text{dét}|(x_i|x_j)|$ si celui-ci est non nul. Puisque $\text{dét}|(x_i|x_j)| = (x_1 \wedge x_2 \wedge \ldots x_p | x_1 \wedge x_2 \wedge \ldots x_p)$ ce signe ne dépend que du p-vecteur $x_1 \wedge x_2 \wedge \ldots x_p$. En particulier l'indicateur d'un vecteur régulier x est $\varepsilon(x) =$ signe $q(x)$. Lorsque le déterminant $\text{dét}|(x_i|x_j)|$ est nul, on pose $\varepsilon(x_1, x_2, \ldots x_p) = 0$. Si maintenant $x_1, x_2, \ldots x_p$ sont p vecteurs linéairement indépendants de E formant une base d'un sous-espace F, toute autre base de F est l'image de cette dernière par un automorphisme a de F et l'on a (§ II.3) :

$$\text{dét}|(ax_i|ax_j)| = (\text{dét}\, a)^2 \text{dét}|(x_i|x_j)|.$$

Ainsi le déterminant $\text{dét}|(x_i|x_j)|$ est non nul et de même signe pour toutes les bases de F si F est régulier ($q|_{\mathrm{F}}$ non-dégénérée). L'indicateur de F est alors $\varepsilon(\mathrm{F}) =$ signe $\text{dét}|(x_i|x_j)|$.

Le déterminant $\text{dét}|(x_i|x_j)|$ est nul si et seulement si le sous-espace F est singulier. On pose alors $\varepsilon(\mathrm{F}) = 0$.

$q(x)$ est le « carré scalaire de x » et la « longueur » de x est $||x|| = (\varepsilon(x)q(x))^{1/2}$.

De même $\text{dét}|(x_i|x_j)|$ est appelé le carré scalaire de $x_1, x_2, \dots x_p$ et le « volume » de ces p vecteurs est $v(x_1, x_2, \dots x_p) = ||x_1 \wedge x_2 \wedge \dots x_p|| = (\varepsilon(x_1, x_2, \dots x_p)\text{dét}|(x_i|x_j)|)^{1/2}$.

Une base $e = (e_1, e_2, \dots e_n)$ de (E, q) est dite orthonormale si $q(e_j) = +1$ ou -1, standard si $q(e_j) = 1$ pour $j = 1, 2, \dots r$ et $q(e_k) = -1$ pour $k = r+1, \dots n$. Il faut bien prendre garde à ce que, dans une base standard, les composantes covariantes x_j d'un vecteur x sont : $x_j = x^j$ pour $j = 1, 2, \dots r$, $x_j = -x^j$ pour $j = r+1, \dots n$ et le produit scalaire s'écrit :

$$(x|y) = \sum_{j=1}^{r} x^j y^j - \sum_{k=r+1}^{n} x^k y^k = \sum x^j y_j$$

Dans une base quelconque e de E, on note $g_{ij} = g(e_i, e_j) = (e_i|e_j)$, G la matrice $|g_{ij}|$ et $\overline{g} = \text{dét G}$. Le produit scalaire de deux vecteurs xy, s'écrit matriciellement dans e : $g(x, y) =^t \mathrm{XGY}$. Si maintenant X désigne la matrice $n \times n$ des composantes de n vecteurs $x_1, x_2, \dots x_n$ on a : $v(x_1, x_2, \dots x_p) = |\overline{g}|^{1/2}|\text{dét X}|$.

Le groupe d'automorphismes de $\mathrm{E}_{r,s}$ est le groupe orthogonal $\mathrm{O}(r, s)$. $\mathrm{O}(1, 3)$ est plus particulièrement appelé groupe de Lorentz.

Soit N_0 un sous-espace strictement négatif maximal, par exemple le sous-espace ayant pour base les s derniers vecteurs $e_{r+1}, \dots e_n$ d'une base standard de $\mathrm{E}_{r,s}$. Tout sous-espace strictement positif maximal P est un supplémentaire de N_0 puisque $\mathrm{P} \cap \mathrm{N}_0 = \{\mathrm{O}\}$ et $\dim \mathrm{P} + \dim \mathrm{N}_0 = n$. En particulier : $\mathrm{P}_0 = \mathrm{N}_0^{\perp}$. La projection Π de P sur P_0 parallèlement à N_0 projette P orthogonalement sur P_0 et est bijective. En effet, on a $x - \Pi x \in \mathrm{N}_0$, d'où $q(x - \Pi x) \leqslant 0$. Si $\Pi x = 0$, $q(x) \leqslant 0$ et $x = 0$. Le choix d'une orientation de P_0, par exemple définie par la base orthonormale $(e_1, e_2, \dots e_r)$ détermine sur chaque P au moyen de la base $(\overset{-1}{\Pi}\, e_1, \dots \overset{-1}{\Pi}\, e_r)$ une orientation canonique. De la même façon l'orientation de N_0 déterminée par la base $(e_{r+1}, \dots e_n)$ définit une orientation canonique des sous-espaces strictement négatifs maximaux. On a ainsi défini une « *orientation complète* » de $\mathrm{E}_{r,s}$: orientation cohérente de tous les sous-espaces strictement positifs ou négatifs maximaux.

Soient maintenant u un automorphisme quadratique de $\mathrm{E}_{r,s}$ ($u \in \mathrm{O}(r,s)$), U sa matrice dans la base standard e. Les vecteurs $(ue_1, ue_2, \ldots ue_r)$ forment une base orthonormale du sous-espace strictement positif maximal $\mathrm{P} = u(\mathrm{P}_0)$, image de P_0, les vecteurs $(ue_{r+1}, \ldots ue_n)$ une base orthonormale de $\mathrm{N} = u(\mathrm{N}_0)$. La sous-matrice U_r que forment les r premières lignes et les r premières colonnes de U est la matrice des composantes relativement à $(e_1, e_2, \ldots e_r)$ des vecteurs projections orthogonales sur $\mathrm{P}_0 : \Pi u(e_1), \ldots, \Pi u(e_r)$.

Puisque Π est bijective, ils sont linéairement indépendants. DétU_r est non nul, et son signe indique l'orientation relative de $(e_1, \ldots e_r)$ et $(\Pi u(e_1), \ldots \Pi u(e_r))$, donc la propriété pour u de conserver ou non l'orientation des sous-espaces positifs maximaux. Un raisonnement analogue vaut pour la sous-matrice U_s de U que forment les s dernières lignes et les s dernières colonnes. Le signe de détU'_s indique si u conserve ou non l'orientation des sous-espaces strictement négatifs maximaux. On en déduit aisément (cf DR) que lorsque r et s sont non nuls, $\mathrm{O}(r,s)$ se scinde en quatre morceaux connexes que l'on peut indexer par : $(++)$, sous-groupe connexe des *rotations propres* $\mathrm{R}(r,s) = \mathrm{O}_{++}(r,s)$ conservant les orientations des sous-espaces strictement positifs et négatifs maximaux, $(+-)$, $(-+)$ et $(--)$.

L'espace temps de la relativité restreinte, ou espace de Minkowski est l'espace affine $\widetilde{\mathrm{E}}_{1,3}$ associé à l'espace pseudoeuclidien $\mathrm{E}_{1,3}$. Choisissant un point O et un repère standard S d'origine O, la métrique s'exprime dans S par $q(x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$. On décide que, *relativement à* S, x^1, x^2, x^3 sont les coordonnées d'espace et que $x^0 = ct$, où c est la vitesse de la lumière et t le *temps mesuré dans* S : x^0 a bien la dimension d'une longueur. Il est clair qu'en ce qui concerne les phénomènes mécaniques terrestres, la coordonnée x^0 est, en valeur numérique, démesurément supérieure aux trois autres. Il n'en est plus de même dans les phénomènes électromagnétiques.

Dans $\widetilde{\mathrm{E}}_{1,3}$, les droites strictement positives sont les « droites de temps ». Leur orientation commune est l'orientation du temps. Les droites strictement négatives sont les « droites d'espace ». Une orientation cohérente des sous-espaces strictement négatifs maximaux, de dimension trois, est une orientation de l'espace.

Soient G la matrice diagonale $(1, -1, -1, -1)$, X et Y les matrices colonnes des composantes de x, y dans un repère standard de

$\mathrm{E}_{1,3}$. Le produit scalaire s'écrit $(x|y) =^t \mathrm{XGY}$. Une matrice L appartient au groupe de Lorentz $0(1,3)$ si et seulemnt si ${}^t\mathrm{LGL} = \mathrm{G}$, d'où $\overset{-1}{\mathrm{L}} = \mathrm{G}^t\mathrm{LG}$ et $\mathrm{dét L} = \pm 1$. Si l'on écrit :

$$\mathrm{L} = \left|\begin{array}{c|ccc} \mathrm{L}_{00} & \mathrm{L}_{10} & \mathrm{L}_{20} & \mathrm{L}_{30} \\ \hline \mathrm{L}_{01} & \mathrm{L}_{11} & \mathrm{L}_{21} & \mathrm{L}_{31} \\ \mathrm{L}_{02} & \mathrm{L}_{12} & \mathrm{L}_{22} & \mathrm{L}_{32} \\ \mathrm{L}_{03} & \mathrm{L}_{13} & \mathrm{L}_{23} & \mathrm{L}_{33} \end{array}\right|,$$

les vecteurs colonnes sont les images des vecteurs de base. D'où, en particulier $\mathrm{L}_{00}^2 - \mathrm{L}_{01}^2 - \mathrm{L}_{02}^2 - \mathrm{L}_{03}^2 = 1$. Il en résulte que $|\mathrm{L}_{00}| \geqslant 1$. Les quatre morceaux connexes de $0(1,3)$ peuvent être caractérisés par L_{00} et $\mathrm{dét\,L}$, soit en utilisant la notation des physiciens : une flèche pour l'orientation du temps, $\pm$ pour le déterminant :

$$\mathrm{L}_+^{\uparrow} : \mathrm{L}_{00} \geqslant 1,\ \mathrm{dét\,L} = +1;\ \mathrm{L}_-^{\uparrow} : \mathrm{L}_{00} \geqslant 1,\ \mathrm{dét\,L} = -1,$$
$$\mathrm{L}_+^{\downarrow} : \mathrm{L}_{00} \leqslant -1,\ \mathrm{dét\,L} = +1;\ \mathrm{L}_-^{\downarrow} \leqslant -1;\ \mathrm{dét\,L} = -1.$$

Une matrice $\mathcal{L}$ appartient à l'algèbre de Lie de $0(1,3)$ si et seulement si ${}^t\mathcal{L}\mathrm{G} + \mathrm{G}\mathcal{L} = 0$, ou ${}^t(\mathrm{G}\mathcal{L}) = -\mathrm{G}\mathcal{L}$ et

$$\mathcal{L} = \left|\begin{array}{c|ccc} 0 & \mathcal{L}_{01} & \mathcal{L}_{02} & \mathcal{L}_{03} \\ \hline \mathcal{L}_{01} & 0 & \mathcal{L}_{12} & -\mathcal{L}_{13} \\ \mathcal{L}_{02} & \mathcal{L}_{12} & 0 & -\mathcal{L}_{23} \\ \mathcal{L}_{03} & \mathcal{L}_{13} & \mathcal{L}_{23} & 0 \end{array}\right|.$$

Le groupe $0(1,3)$ est donc de dimension 6, son algèbre de Lie ayant six générateurs linéairement indépendants.

Remarque : Il est bien connu (DR) que $\mathrm{E}_{1,3}$ peut être représenté comme l'espace vectoriel réel $\mathfrak{H}$ des matrices hermitiennes d'ordre 2 par :

$$x \to \mathrm{X} = \begin{vmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{vmatrix} \quad \text{avec } q(x) = \mathrm{dét\,X}.$$

Les matrices complexes 2×2 de déterminant $+1$ forment le groupe $\mathrm{Sl}(2;\mathbb{C})$ et opèrent dans $\mathfrak{H}$. Si $\Lambda \in \mathrm{Sl}(2;\mathbb{C})$, $\mathrm{X} \in \mathfrak{H} \to \Lambda(\mathrm{X}) = \Lambda\mathrm{X}^t\overline{\Lambda}$ et $\mathrm{dét}\,\Lambda(\mathrm{X}) = \mathrm{dét\,X}$. On obtient ainsi un homomorphisme du groupe $\mathrm{Sl}(2;\mathbb{C})$, qui est connexe et simplement connexe sur le groupe $\mathrm{L}_+^\uparrow$ des rotations propres de Lorentz. Le noyau est formé de I et $-$I et $\mathrm{Sl}(2;\mathbb{C})$ est ainsi le revêtement universel de $\mathrm{L}_+^\uparrow$.

Un couple $(0, e_0)$ formé d'un point 0 de $\widetilde{\mathrm{E}}_{1,3}$ et d'un vecteur unitaire, $e_0 : q(e_0) = +1$, orienté dans le sens des temps positifs *détermine* une séparation *des coordonnées* d'espace et de temps, l'espace E_3 associé étant $e_0^\perp$. $(0, e_0)$ peut être appelé un observateur instantané de l'espace-temps.

Le groupe d'invariance des phénomènes physiques en relativité restreinte (imposé par les équations de Maxwell) est le *groupe de Poincaré* formé par l'extension du groupe des rotations propres de Lorentz $\mathrm{L}_+^\uparrow$ par les translations de l'espace affine $\widetilde{\mathrm{E}}_{1,3}$. Ni le temps, ni les longueurs, ni la simultanéité (égalité des coordonnées x^0) ne sont des invariants : ils dépendent du repère où on les évalue. Le *seul* invariant métrique est le carré de *l'intervalle de temps propre* entre deux points M_0, M_1 (ou évènements) de l'espace temps : $\tau^2 = \dfrac{1}{c^2} q(\mathrm{M}_1 - \mathrm{M}_0)$, qui, dans un repère standard, a pour expression :

$$\tau^2 = (t_1 - t_0)^2 - \frac{1}{c^2}\{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2\}$$

soit : $\Delta\tau^2 = \Delta t^2 - \dfrac{1}{c^2}\Delta r^2$.

Si dans le repère S, M_0 et M_1 occupent le même point « d'espace » au cours du temps, $\Delta\tau = \Delta t$. Il faut bien voir que dans l'espace-temps $\widetilde{\mathrm{E}}_{1,3}$, la coordonnée de temps n'est pas statique : le temps est constamment croissant. Un point M de l'espace-temps doit donc toujours être considéré comme « en mouvement ». Dans un repère standard S, il possède une vitesse (relative à S) v de composantes $\left(\dfrac{dx}{dt}, \dfrac{dy}{dt}, \dfrac{dz}{dt}\right)$ et la différentielle du temps propre le long de la trajectoire de M est $d\tau = \left(1 - \dfrac{v^2}{c^2}\right)^{1/2} dt$. La vitesse v ne peut donc jamais dépasser la vitesse c de la lumière dans le vide et le *facteur* $\overset{-1}{\gamma} = \left(1 - \dfrac{v^2}{c^2}\right)^{1/2}$ est < 1 si $v \neq 0$.

Or, on a vu que $d\tau$ est la différentielle dt' du temps le long de la trajectoire de M mesuré dans un repère S′ où M occupe le même point d'espace, c'est-à-dire dans un repère S′ qui suit le mouvement de M. On a donc $dt' = \overset{-1}{\gamma}\, dt < dt$ si $v \neq 0$. C'est le phénomène de la dilatation du temps : si S′ est en mouvement par rapport à S, le temps en un point d'espace fixe de S′ s'écoule plus lentement que le temps décrivant le mouvement de M vu de S.

Le calcul classique de la cinématique relativiste est le suivant. Soit S′ un repère de $\widetilde{\mathrm{E}}_{1,3}$ en mouvement rectiligne uniforme par rapport au repère S avec une vitesse v. Supposons la coïncidence des origines O et O′ de S et S′ comme évènement origine. Supposons également que les axes en soient choisis de telle sorte que $\mathrm{O}x^2, \mathrm{O}'x'^2$ d'une part, $\mathrm{O}x^3, \mathrm{O}'x'^3$ d'autre part, restent parallèles. La transformation du groupe de Poincaré qui fait passer de S à S′ est, puisque $\mathrm{O}' = \mathrm{O}$ lorsque toutes les coordonnées sont nulles, une transformation de Lorentz de l'espace quadratique S.

Elle s'écrit : $x'^2 = x^2$; $x'^3 = x^3$, et, pour les variables x^0 et x^1, est une rotation de Lorentz du plan des axes $(\mathrm{O}x^0, \mathrm{O}x^1)$. Or, avec la métrique induite $q(x) = (x^0)^2 - (x^1)^2$, ce plan est un plan hyperbolique (cf DR) et une transformation de Lorentz de ce plan est une rotation hyperbolique. Une telle rotation de l'angle $(-\varphi)$ autour de O s'écrit :

$$\left\{\begin{array}{l} x'^1 = x^1 \operatorname{ch}\varphi - x_0 \operatorname{sh}\varphi \\ x'^0 = -x^1 \operatorname{sh}\varphi + x^0 \operatorname{ch}\varphi \end{array}\right. .$$

La transformation réciproque est la rotation hyperbolique d'angle $(+\varphi)$:

$$\left\{\begin{array}{l} x'^1 = x'^1 \operatorname{ch}\varphi + x'^0 \operatorname{sh}\varphi \\ x'^0 = x'^1 \operatorname{sh}\varphi + x'^0 \operatorname{ch}\varphi \end{array}\right. .$$

Dans le repère S, la trajectoire de O′ a pour équation $x'^1 = 0$ soit

$$\left\{\begin{array}{l} x^1 = x = x'^0 \operatorname{sh}\varphi = ct' \operatorname{sh}\varphi \\ x^0 = ct = x'^0 \operatorname{ch}\varphi = ct' \operatorname{ch}\varphi \end{array}\right. .$$

d'où $t = t' \operatorname{ch}\varphi$ et $x = ct \operatorname{Th}\varphi$.

Sa vitesse est $v = \dfrac{dx}{dt} = c\mathrm{Th}\,\varphi$. En exprimant les fonctions hyperboliques de φ à l'aide de v, on obtient :

$$\mathrm{Th}\,\varphi = \frac{v}{c};\ \mathrm{ch}\,\varphi = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} = \gamma;\ \mathrm{sh}\,\varphi = \gamma\frac{v}{c}$$

ce qui donne finalement pour la transformation de Lorentz de S à S$'$ des coordonnées t, x, les expressions :

$$x' = (x - vt)\gamma;\ t' = \left(t - x\frac{v}{c^2}\right)\gamma.$$

Une conséquence est que, si $\mathrm{M}_1\mathrm{M}_2$ est un segment fixe de l'axe des x' dans S$'$ de longueur $l' = x'_2 - x'_1$, sa mesure l dans le repère S est liée à l' par :

$$l' = x'_2 - x'_1 = (x_2 - x_1)\gamma = l\gamma.$$

Ainsi $l = l'\gamma^{-1} < l'$. C'est la « contraction de Lorentz ».

Les dilatations du temps et contractions des longueurs ne doivent pas être considérés comme des phénomènes matériels. Elle reflètent simplement le changement de la mesure du temps et des longueurs qui résulte du mouvement du repère, et manifestent, par là même, qu'en relativité restreinte, ni le temps ni les longueurs ne sont des invariants absolus.

Remarquons que si l'on n'a pas la coïncidence des origines O et O$'$ des repères lorsque toutes les coordonnées sont nulles, la transformation de Poincaré qui fait passer de S à S$'$ s'obtient en ajoutant à la rotation de Lorentz une translation. Si maintenant S$''$ est un repère en mouvement rectiligne uniforme de vitesse u relativement à S$'$ dont les axes $\mathrm{O}''x''^2$ et $\mathrm{O}''x''^3$ restent parallèles à $\mathrm{O}'x'^2$, $\mathrm{O}'x'^3$ et $\mathrm{O}'x'^3$, $\mathrm{O}x^3$, la transformation de Poincaré faisant passer de S$'$ à S$''$ est somme d'une translation et d'une rotation hyperbolique du plan (x^0, x^1) caractérisés par un angle $(-\psi)$ avec $\dfrac{u}{c} = \mathrm{Th}\,\psi$. La rotation hyperbolique associée à la transformation de S à S$''$ a pour angle $-(\varphi + \psi)$. La vitesse V de translation de S$''$ par rapport à S est alors donnée par la formule d'addition des angles pour les tangentes hyperboliques :

$$\frac{\mathrm{V}}{c} = \mathrm{Th}\,(\varphi + \psi) = \frac{\mathrm{Th}\,\varphi + \mathrm{Th}\,\psi}{1 + \mathrm{Th}\,\varphi \cdot \mathrm{Th}\,\psi},$$

et l'on obtient la formule de composition des vitesses *parallèles* en relativité restreinte.

$$\mathrm{V} = \frac{u+v}{1+\dfrac{uv}{c^2}}.$$

En particulier, l'addition à la vitesse de la lumière $u = c$ d'une vitesse de translation quelconque v redonne toujours c :

$$\frac{c+v}{1+\dfrac{cv}{c^2}} = c \text{ quel que soit } v.$$

Autrement dit, si une particule est animée d'une vitesse c relativement à un repère, elle possède la même vitesse c dans n'importe quel autre repère en mouvement par rapport au premier.

La vitesse relative v d'un point M de coordonnées $(x^0 = ct, x, y, z)$ dans un repère standard S de l'espace-temps, a pour composantes dans S : $\dfrac{dx}{dt}, \dfrac{dy}{dt}, \dfrac{dz}{dt}$. Même en y ajoutant $c = \dfrac{dx^0}{dt}$, ces composantes *ne sont pas celles d'un vecteur* d'après ce qui précède.

Pour associer au point M de l'espace-temps une *vitesse relativiste* il faut faire intervenir le temps propre τ, qui est invariant, soit :

$$\mathrm{V} = \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau}\right) = (c, v)\frac{dt}{d\tau} = (c, v)\gamma = (c\gamma, v\gamma)$$

V est un vecteur tangent en M à sa trajectoire d'espace-temps de « longueur » fixe :

En effet : $q(\mathrm{V}) = ||\mathrm{V}||^2 = \dfrac{(dx^0)^2 - (dx^1)^2 - (dx^2)^2}{d\tau^2} = c^2.$

Il en résulte que : $\mathrm{O} = \dfrac{d}{d\tau}||\mathrm{V}||^2 = \dfrac{d}{d\tau}(\mathrm{V}|\mathrm{V}) = 2\left(\dfrac{d\mathrm{V}}{d\tau}\middle|\mathrm{V}\right).$

L'accélération relativiste $\dfrac{d\mathrm{V}}{d\tau}$ est orthogonale (pour la métrique de Minkowski à la vitesse relativiste V.

Il n'a été question jusqu'à présent que de cinématique. Une dynamique relativiste doit clarifier la notion de masse ainsi que son comportement par changement de repère. Une démarche classique consiste à les déduire de l'électrodynamique à partir de l'action de la force de Lorentz sur une masse ponctuelle chargée.

Mais de façon plus élémentaire, le choc élastique de deux particules permet, grâce à la conservation du moment, de définir le rapport de leurs masses à l'aide de vitesses. On peut ainsi partir de la notion de masse au repos m_0 d'une particule. Lorsque celle-ci est animée d'une vitesse v relativement au repère S son « moment relativiste » a pour expression dans S : $\mathrm{P} = m_0\mathrm{V} = (m_0\gamma c, m_0\gamma v)$; m_0 étant un scalaire, P est un vecteur et $||\mathrm{P}||^2 = m_0^2c^2$. La composante d'espace de P dans le repère S est le moment de la mécanique classique : $p = m_0\gamma v = mv$. Il en résulte de cette dernière expression que la mesure m de la masse d'une particule dans un repère S dépend de la vitesse v de cette particule relativement à S :

$$m = m_0\gamma = \frac{m_0}{\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}},$$

m tend vers l'infini lorsque v tend vers c.

La force relativiste agissant sur la particule est :

$$\begin{aligned} \mathrm{F} &= m_0\frac{d\mathrm{V}}{d\tau} = \frac{d\mathrm{P}}{d\tau} = \frac{d}{d\tau}(mc, p) = \frac{d}{dt}(mc, p)\frac{dt}{d\tau} \\ &= \left(\frac{dm}{dt}c, \frac{dp}{dt}\right)\gamma = \left(\frac{dm}{dt}c, f\right)\gamma \end{aligned}$$

où f est la force classique. En exprimant l'orthogonalité de F et de $\mathrm{V} = (c, v)\gamma$, on obtient :

$$\frac{dm}{dt}c^2 - f \cdot v = 0.$$

Or $f \cdot v$ représente le travail accompli par la force f pendant le mouvement. L'augmentation de l'énergie de la particule entre t_1 et t_2 est donc $\int_{t_1}^{t_2} f \cdot v = m_2c^2 - m_1c^2$.

Cela conduit à associer à une particule son énergie interne $\mathrm{E}_0 = m_0c^2$. Son énergie totale E lorsqu'elle est en mouvement est la somme de son énergie cinétique T et de son énergie interne, soit : $\mathrm{E} = mc^2 = \mathrm{T} + m_0c^2$. D'où :

$$\mathrm{T} = \frac{m_0c^2}{\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}} - m_0c^2 = \frac{1}{2}m_0v^2 + (\text{termes en } \frac{v^4}{c^2}).$$

Le vecteur moment relativiste $\mathrm{P} = (mc, mv)$ peut ainsi être écrit $\mathrm{P} = \left(\dfrac{\mathrm{E}}{c}, p\right)$ et être légitimement appelé vecteur de moment-énergie. C'est cette expression qui est la définition générale du moment d'une particule en relativité. En effet, dans le cas d'un photon, la masse est nulle mais non l'énergie, et son moment se réduit à $\mathrm{P} = \dfrac{\mathrm{E}}{c}$. La mécanique quantique prescrit à un photon de fréquence ν une énergie égale à $h\nu$, où h est la constante de Planck. Le moment d'un tel photon a donc pour valeur $\dfrac{h\nu}{c}$. (Rappelons les formules de dimensions, d'une vitesse $\mathrm{L}\,\mathrm{T}^{-1}$, d'un moment, $\mathrm{ML}\,\mathrm{T}^{-1}$, d'une force $\mathrm{ML}\,\mathrm{T}^{-2}$, d'une énergie $\mathrm{ML}^2\,\mathrm{T}^{-2}$).

La conservation du moment relativiste, parfaitement vérifiée par l'expérience, recouvre ainsi la conservation classique du moment cinétique et la conservation de l'énergie.

L'augmentation des contraintes imposées aux lois de la physique par un plus grand groupe de symétrie, le groupe de Poincaré, qui prend la place du groupe de Galilée, a un effet à la fois de simplification et de coordination. En particulier, l'ensemble des énergies présentes sous toutes les formes : champs, radiations, tensions, pressions, énergie cinétique, matière, etc..., à l'exclusion de la gravitation est exprimée par un tenseur symétrique du second ordre T. T est le « *stress-energy tensor* », désigné en français de diverses façons : *tenseur d'impulsion-énergie, de moment-énergie*, ou simplement d'*énergie*. C'est ce tenseur qui en relativité générale (§ IX.10) forme le second membre de l'équation d'Einstein, qui consiste à l'égaler à un tenseur géométrique de courbure G : $\mathrm{G} = \dfrac{1}{8\pi}\mathrm{T}$. L'essence de la théorie est que la distribution de l'énergie représentée par T détermine la métrique de l'espace par l'intermédiaire de sa courbure qui, elle, est l'expression du champ de gravitation. L'énergie gravitationnelle est donc fondamentalement séparée et distinguée des autres formes d'énergie. Il en résulte en particulier qu'on n'a pas en relativité générale de notion raisonnable de densité d'énergie gravitationnelle (en gravitation newtonnienne, celle-ci a pour valeur $\frac{1}{8\pi}|\mathrm{grad}\Phi|^2$ où Φ est le potentiel).

Exemples classiques de tenseurs d'énergie :

1) d'énergie purement cinétique. On considère un fluide de densité au repos ρ_0 où les interactions entre particules, en particulier la pression sont négligeables. Si u est la vitesse relativiste du fluide en un point de l'espace-temps, le tenseur d'énergie est $\mathrm{T} = \rho_0 u \otimes u$. Il a pour composantes dans un repère quelconque : $\mathrm{T}^{ij} = \rho_0 u^i u^j$. Dans un repère standard (ct, x, y, z) on a :

$$\mathrm{T}^{00} = \rho_0 \left(\frac{dx^0}{d\tau}\right)^2 = \rho_0 c^2 \left(\frac{dt}{d\tau}\right)^2 = \rho_0 c^2 \gamma^2 = \rho c^2,$$

$\rho = \rho_0 \gamma^2$ est la densité relativiste. Les volumes étant comme les longueurs, contractés dans le sens de la vitesse par un facteur $\overset{-1}{\gamma}$ et les masses augmentées par un facteur γ, la densité se trouve augmentée par un facteur γ^2 et $\mathrm{T}^{00} = \rho c^2$ est la densité d'énergie associée à la masse. Ce résultat, est général : pour tout tenseur d'énergie T, T^{00} est la densité de masse-énergie :

$$\mathrm{T}^{j0} = \mathrm{T}^{0j} = \rho_0 \frac{dx^0}{d\tau}\frac{dx^j}{d\tau} = \rho_0 c \gamma^2 v^j = c\rho v^j.$$

Les T^{0j}, $j = 1, 2, 3$ forment la densité de moment cinétique multipliée par c.

2) d'un fluide parfait dont la pression est notée π :

$$\mathrm{T} = (\rho + \pi) u \otimes u + \pi g \text{ (où } g \text{ est le tenseur métrique).}$$

3) nous examinerons dans le paragraphe suivant le tenseur d'énergie du champ électromagnétique.

La conservation de l'énergie s'exprime dans un repère standard pour un tenseur d'énergie T : $\mathrm{div}\,\mathrm{T}^0 = \sum \mathrm{T}^{0i}_{|i} = \sum \frac{\partial}{\partial x^i}\mathrm{T}^{0i} = 0$.

Dans le cas du tenseur 1) ci-dessus, cette équation est l'équation de continuité d'Euler de conservation de la masse : $\frac{\partial \rho}{\partial t} + \nabla \cdot \rho v = 0$.

VII.13 – Les équations de Maxwell

Dans un repère standard d'espace-temps, qui sépare donc la variable de temps des variables d'espace, les phénomènes électromagnétiques sont décrits : *a)* dans le vide, par deux champs de vecteurs : E, B, électrique et magnétique : *b)* en présence de matière par quatre champs de vecteurs $\mathrm{E}, \mathrm{B}, \mathrm{H}, \mathrm{D}$. Les équations de Maxwell expriment leurs relations mutuelles. Si ces équations sont parfaitement adaptées aux calculs, qui se font toujours avec des composantes, leur signification physique profonde ne devient manifeste que si on les écrit sous forme tensorielle, et elles apparaissent alors comme des relations entre les composantes d'un même tenseur. On peut mieux encore, les exprimer à l'aide de formes différentielles extérieures. C'est le but de ce paragraphe. Observons d'abord que les équations de Maxwell peuvent être écrites dans différents systèmes d'unités, ce qui fait apparaître ou disparaître des coefficients numériques. Nous utiliserons le système d'unités le plus courant, dit d'unités Gaussiennes. C'est le système CGS auquel on adjoint la condition que le coefficient numérique de la loi de Coulomb soit égal à un. Ce choix a l'inconvénient de *détruire* les dimensions des diverses grandeurs physiques en référence à un système d'unités indépendantes (voir A. Sommerfeld, Electrodynamics, Academic Press 1952, et « J.D. Jackson, Classical Electrodynamics, Wiley 1975 », l'« Appendix on Units and Dimensions »). Nous rappellerons d'abord rapidement l'origine et la formation des équations de Maxwell.

La loi élémentaire de Coulomb en $\dfrac{1}{r^2}$ d'attraction ou de répulsion, des charges électriques ponctuelles, a pour conséquence : *a)* la définition du champ électrique E exercé en un point par une distribution de charges de densité ρ, *b)* le théorème de Gauss : le flux du champ E à travers une surface fermée est égal à la charge totale contenue dans le volume limité par la surface multipliée par 4π.

L'expression différentielle du théorème de Gauss s'écrit : $\nabla \cdot \mathrm{E} = \operatorname{div} \mathrm{E} = 4\pi\rho$. C'est la première équation de Maxwell pour des charges situées *dans le vide*. La présence de matière complique la situation comme on va le voir. Alors que, dans un conducteur, des charges électriques peuvent se déplacer librement (électrons de conduction, ions dans un électrolyte), dans un isolant ou diélectrique, les charges électriques contenues dans ses atomes sont liées et ne peuvent effectuer que de très faibles déplacements sous l'action d'un champ électrique. Ces déplacements provoquent l'apparition d'un champ de dipoles électriques à l'intérieur de l'isolant. Il peut aussi se faire que les molécules d'un corps isolant soient des dipoles électriques naturels en raison de la répartition de leurs charges. Ces dipoles sont disposés de façon aléatoire en l'absence d'un champ extérieur et s'orientent par un phénomène analogue à l'aimantation sous l'action d'un champ électrique. L'apparition de dipoles électriques orientés de façon cohérente est appelée *polarisation*. Elles se décrit par un champ de vecteurs p : p est la densité de moment des dipoles. La relation entre p et le champ électrique E causant la polarisation est généralement une relation de proportionnalité : $p = \chi \mathrm{E}$ où χ est la susceptibilité électrique de l'isolant, mais χ peut dans certains cas être un opérateur linéaire.

Or, un dipole élémentaire de moment $\vec{p}$ situé en un point O détermine, en un point M, un potentiel $\varphi = \vec{p} \cdot \mathrm{grad}\dfrac{1}{r}$ où $r = |\mathrm{OM}|$.

Le champ de moments dipolaires $\vec{p}$ détermine ainsi un potentiel $\Phi = \int_{\mathrm{V}} \left(\vec{p} \cdot \mathrm{grad}\dfrac{1}{r}\right) dv$ (V : volume de l'isolant).

Cette intégrale se transforme, vue l'égalité :

$$\vec{p} \cdot \mathrm{grad}\frac{1}{r} = \mathrm{div}\frac{\vec{p}}{r} - \frac{1}{r}\mathrm{div}\vec{p}$$

et le théorème de Gauss, en :

$$\Phi = \int_{\mathrm{S}} \frac{\vec{p}\vec{n}}{r} d\sigma - \int_{\mathrm{V}} \frac{\mathrm{div}\vec{p}}{r} dv \text{ (S surface limitant V).}$$

Cette expression du potentiel signifie que la polarisation d'un diélectrique fait apparaître sur sa surface des charges ayant une densité superficielle égale à $\omega = \vec{p} \cdot \vec{n}$ et des charges volumiques ayant une densité $\rho_p = -\mathrm{div}\vec{p}$ dans V, prédiction mathématique confirmée par l'expérience.

A l'intérieur d'un isolant on a donc :

$$\mathrm{div}\,\mathrm{E} = 4\Pi(\rho + \rho_p) = 4\Pi\rho - 4\Pi\,\mathrm{div}\,\vec{p}.$$

Alors que E est lié à toutes les charges y compris celles qui sont induites par la polarisation, le vecteur $\mathrm{D} = \mathrm{E} + 4\Pi\vec{p}$ ne dépend que des vraies charges :

$$\boxed{\nabla \cdot \mathrm{D} = \mathrm{div}\,\mathrm{D} = 4\Pi\rho} \text{ (lois de Coulomb et de Gauss).}$$

D est le vecteur *déplacement électrique* et l'équation précédente est la *première équation de Maxwell* en présence de matière. Elle est aussi valable à la surface d'un diélectrique, la composante normale de D étant continue lorsqu'on traverse cette surface.

Lorsque le champ électrique E est variable, la polarisation l'est aussi et il en résulte un courant de polarisation. Un calcul simple montre que sa densité est $\frac{\partial p}{\partial t}$. Ce courant est généralement négligeable.

Des circuits portant des courants électriques agissent l'un sur l'autre. La force que peut exercer un circuit déterminé en un point M se manifeste par un champ de vecteurs B appelé *induction magnétique*. On a pour calculer B une expression différentielle (loi de Biot et Savart) qu'il suffit ensuite d'intégrer : un élément de courant $\mathrm{I}\vec{ds}$ en P produit une induction élémentaire $d\mathrm{B}$ en M avec $\vec{r} = \overrightarrow{\mathrm{PM}}$ donné par : $d\mathrm{B} = \frac{\mathrm{I}}{c}\frac{\vec{ds} \times \vec{r}}{r^3}$. C'est aussi une loi en $\frac{1}{r^2}$ mais le champ est ici orthogonal au plan $(\vec{ds}, \vec{r})$. (Rappelons la convention usuelle : le vecteur $\vec{r} = \overrightarrow{\mathrm{PM}}$ est *toujours* dirigé de la source en $\mathrm{P}(\mathrm{I}\vec{ds})$ *vers* le point d'observation M).

L'expression générale de l'induction magnétique déterminée en un point M_0 par une densité de courant J est alors :

$$\mathrm{B} = \frac{1}{c}\int \mathrm{J}(\mathrm{P}) \times \frac{\vec{r}}{r^3} dv_p \ (\vec{r} = \overrightarrow{\mathrm{PM}_0}).$$

Or le champ de vecteurs $\frac{\vec{r}}{r^3} = \left(\frac{x_0 - x}{r^3}, \frac{y_0 - y}{r^3}, \frac{z_0 - z}{r^3}\right)$ s'écrit $-\left(\frac{\partial}{\partial x_0}\left(\frac{1}{r}\right), \frac{\partial}{\partial y_0}\left(\frac{1}{r}\right), \frac{\partial}{\partial z_0}\left(\frac{1}{r}\right)\right) = -\mathrm{grad}_0\frac{1}{r}$ et l'intégrale se transforme en :

$$\mathrm{B} = \frac{1}{c}\int (\mathrm{grad}_0\frac{1}{r} \times \mathrm{J}(\mathrm{P}))dv = \frac{1}{c}\mathrm{rot}_0 \int \frac{\mathrm{J}(\mathrm{P})}{r}dv = \mathrm{rot}_0\mathrm{A}.$$

Le champ de vecteurs $\mathrm{A} = \frac{1}{c}\int \frac{\mathrm{J}(r)}{r}dv$, dont B est le rotationnel, est appelé le *potentiel vecteur d'induction magnétique*.

Puisque B est un rotationnel, on a $\boxed{\nabla \cdot \mathrm{B} = \mathrm{div}\,\mathrm{B} = 0}$. C'est la *deuxième équation de Maxwell*.

Sur un courant stable d'intensité I parcourant la courbe fermée Γ, la force a pour expression : $\mathrm{F} = \dfrac{\mathrm{I}}{c} \int \vec{ds} \times \overrightarrow{\mathrm{B}}$ soit aussi : $d\mathrm{F} = \dfrac{\mathrm{I}}{c} \vec{ds} \times \overrightarrow{\mathrm{B}}$.

Il en résulte que sur une particule de charge q animée d'une vitesse v l'action électromagnétique provenant d'un champ électrique E et d'une induction magnétique B est une force, dite *de Lorentz* égale à :

$$\mathrm{F} = q\left(\mathrm{E} + \frac{1}{c} v \times \mathrm{B}\right).$$

Si l'on déplace un circuit conducteur Γ où ne passe initialement aucun courant, dans un champ d'induction magnétique B, chaque électron mobile du circuit conducteur subit ainsi une force $\mathrm{F} = \dfrac{e}{c} v \times \mathrm{B}$ qui le met en mouvement et détermine un courant électrique. La force électromagnétique est $\dfrac{1}{c} \int_\Gamma (v \times \mathrm{B}) \cdot \vec{ds}$ intégrale qu'une transformation simple montre être égale à $-\dfrac{1}{c}\dfrac{\partial}{\partial t} \int_{\sum} \mathrm{B} \cdot \vec{n} d\sigma$ où $\int_{\sum} \mathrm{B} \cdot \vec{n} d\sigma$ est le flux du champ B à travers une surface $\sum$ de bord Γ. En interprétant la création d'un courant dans Γ par celle d'un champ électrique E qui l'induit on aboutit à la loi d'induction de Faraday : $\int_\Gamma \mathrm{E} \cdot \vec{ds} = -\dfrac{1}{c} \int_{\sum} \dfrac{\partial \mathrm{B}}{\partial t} \vec{n} d\sigma$ dont la forme différentielle déduite par le théorème de Stokes s'écrit :

$$\boxed{\nabla \times \mathrm{E} = \operatorname{rot} \mathrm{E} = -\frac{1}{c}\frac{\partial \mathrm{B}}{\partial t}} \text{ (loi de Faraday).}$$

C'est *la troisième équation de Maxwell.*

On peut aisément résoudre cette équation en E en y remplaçant B par le rotationnel du potentiel vecteur d'induction magnétique A, ce qui donne $\operatorname{rot}\left(\dfrac{1}{c}\dfrac{\partial \mathrm{A}}{\partial t}\right) = 0$, et $\mathrm{E} = -\dfrac{1}{c}\dfrac{\partial \mathrm{A}}{\partial t} - \operatorname{grad} \mathrm{Q}$, où Q est le potentiel résultant des charges électriques.

Nous arrivons maintenant à la plus difficile et aussi la plus ingénieuse des équations de Maxwell. Son point de départ, qui n'est valable que dans la situation statique et en l'absence de matière magnétisable est la *formule d'Ampère*. Une démonstration élégante de cette formule s'appuie sur des considérations géométriques. On sait que l'angle solide (orienté) $\delta\Omega$ d'où l'on voit d'un point M un élément de surface $\delta\sigma$ situé en P, orienté par le choix d'un vecteur unitaire normal $\vec{n}$, est la mesure algébrique de l'aire orientée découpée sur la sphère unité de centre M par le cône d'origine M s'appuyant sur $\delta\sigma$:

Si nous maintenons notre convention pour le vecteur $\vec{r} = \overrightarrow{\mathrm{PM}}$ dirigé de la source P vers le point d'observation M :

$$\boxed{\delta\Omega = -\frac{1}{r^3}(\vec{r}|\vec{n})\delta\sigma = -(\mathrm{grad}_{\mathrm{P}}\frac{1}{r}|\vec{n})\delta\sigma = (\mathrm{grad}_{\mathrm{M}}\frac{1}{r}|\vec{n})\delta\sigma}\,.$$

L'angle solide d'où l'on voit d'un point M une surface S orientée, *ne contenant pas* M, ayant pour bord une courbe Γ est :

$$\Omega = -\int_{\mathrm{S}}\frac{1}{r^3}(\vec{r}\cdot\vec{n})\delta\sigma = \int_{\mathrm{S}}(\mathrm{grad}_{\mathrm{M}}\frac{1}{r})\cdot\vec{n}\delta\sigma_{\mathrm{P}}.$$

Si S' est une autre surface orientée de bord Γ pouvant se déduire de S par déformation continue *sans rencontrer* M, elle donne la même valeur de Ω puisque div $\mathrm{grad}\frac{1}{r} = \Delta\frac{1}{r} = 0$ ($\frac{1}{r}$ est harmonique). Si au contraire elle traverse M dans sa déformation, Ω subit une discontinuité de $\pm 4\Pi$. S étant fixée, Ω est une fonction de M continue et différentiable dans tout l'espace moins S dont nous allons calculer la différentielle $d\Omega = \mathrm{grad}\,\Omega\cdot\vec{du}$. Or, déplacer M de $\vec{du}$ est équivalent pour Ω à déplacer S, donc Γ, de $(-\vec{du})$; Γ et $\Gamma - \vec{du}$ forment le bord d'un « ruban » R dont l'élément de surface orienté est $(-\vec{du})\times\vec{ds}$, où $\vec{ds}$ est l'élément d'arc orienté de Γ. D'où :

$$d\Omega = \int_{\mathrm{R}}\frac{1}{r^3}\vec{r}\cdot(\vec{du}\times\vec{ds}) = \vec{du}\cdot\int_{\Gamma}\frac{\vec{ds}\times\vec{r}}{r^3}$$

et $\mathrm{grad}\ \Omega = \int_{\Gamma}\frac{\vec{ds}\times\vec{r}}{r^3}$, qui est un champ de vecteurs défini dans tout l'espace moins Γ.

Si l'on a maintenant un courant électrique *stable* d'intensité I parcourant Γ dans le sens positif, la loi de Biot et Savart ci-dessus donne pour son induction magnétique en M :

$$\mathrm{B_M} = \frac{\mathrm{I}}{c}\int_\Gamma \frac{\vec{ds} \times \vec{r}}{r^3} = \frac{\mathrm{I}}{c}\mathrm{grad}\,\Omega.$$

On voit donc que B *est*, en dehors de Γ, *un champ de gradients* et dérive d'un *potentiel scalaire* $\Phi = \dfrac{-\mathrm{I}}{c}\Omega$ qui est une fonction *multivalente*. En effet si l'on calcule la circulation du vecteur B le long d'une courbe C : $\int_\mathrm{C} \overrightarrow{\mathrm{B}} \cdot \vec{dl} = \frac{\mathrm{I}}{c}\int_\mathrm{C} \mathrm{grad}\,\Omega \cdot \vec{dl} = \frac{\mathrm{I}}{c}\int_\mathrm{C} d\Omega$ et si la courbe C entoure une fois la courbe Γ dans le sens positif, on obtient :

$$\int_\mathrm{C} \overrightarrow{\mathrm{B}} \cdot \vec{dl} = \frac{\mathrm{I}}{c}4\Pi.$$

C'est la formule (intégrale) d'Ampère dont l'expression différentielle obtenue en remplaçant l'intégrale curviligne par une intégrale de surface étendue à une surface $\sum$ bordant la courbe C, est :

$$\boxed{\nabla \times \mathrm{B} = \mathrm{rot}\,\mathrm{B} = \frac{4\Pi \mathrm{J}}{c}} \quad \textit{(formule d'Ampère)}$$

où J est la densité de courant.

Cette formule n'est qu'une partie de la quatrième équation de Maxwell, et ceci pour deux raisons :

La première est le phénomène de magnétisation. Chaque atome est le siège de courants électriques engendrés par des mouvements orbitaux des électrons autour des noyaux. Or, la valeur en M du potentiel scalaire d'induction magnétique d'un courant d'intensité I dans un circuit Γ, bordant une surface S, peut s'écrire

$$\Phi_\mathrm{M} = -\frac{\mathrm{I}}{c}\Omega_\mathrm{M} = \frac{\mathrm{I}}{c}\int_\mathrm{S} \frac{\vec{n} \cdot \vec{r}}{r^3} d\sigma = \int_\mathrm{S} \left(\frac{\mathrm{I}}{c}\vec{n} \cdot \mathrm{grad_P}\frac{1}{r}\right) d\sigma.$$

Rappelons qu'un dipole élémentaire de moment $\vec{p}$ situé en P, produit un champ B qui peut être représenté aussi bien comme champ de gradients d'un potentiel φ : $\mathrm{B_M} = -\mathrm{grad}\varphi_\mathrm{M}$ avec $\varphi_\mathrm{M} = \dfrac{\vec{p} \cdot \vec{r}}{r^3} = \vec{p} \cdot \mathrm{grad_P}\dfrac{1}{r}(\vec{r} = \overrightarrow{\mathrm{PM}})$ que comme champ de rotationnels d'un potentiel vecteur A : $\mathrm{B_M} = \mathrm{rot}\,\mathrm{A_M}$ avec $\mathrm{A_M} = \dfrac{\vec{p} \times \vec{r}}{r^3} = \vec{p} \times \mathrm{grad_P}\dfrac{1}{r}$.

Le potentiel Φ_M apparaît donc comme le potentiel d'un champ de dipoles répartis sur la surface S dont la densité superficielle de moments est $\dfrac{\mathrm{I}\vec{n}}{c}$.

Un dipole élémentaire dont le potentiel et le champ sont le potentiel scalaire et le champ d'induction magnétique d'un courant infinitésimal d'intensité I autour de l'élément de surface $d\sigma$ est appelé un *dipole magnétique* (ou aimant élémentaire). Son moment $\vec{m} = \dfrac{\mathrm{I}}{c}\vec{n} \cdot d\sigma$ est son *moment magnétique*. On introduit ainsi, de façon purement mathématique une notion de charge magnétique, ces charges n'intervenant que par paires sous formes de dipoles équivalentes à des courants. Elles ne possèdent (jusqu'à nouvel ordre) aucune existence physique.

Outre les moments magnétiques provenant des courants orbitaux des électrons, les électrons eux-mêmes possèdent un moment magnétique qui ne peut pas s'interpréter physiquement comme celui d'un courant.

Si l'on applique à un corps un champ d'induction magnétique B, les couples qu'exerce B sur les moments magnétiques de ses atomes tendent à les aligner, malgré l'agitation thermique : c'est le phénomène de magnétisation, qui peut être éphémère ou permanente.

Supposons donc une densité J de courants électriques et une densité m de moments magnétiques. Ils créent ensemble un champ d'induction magnétique B que l'on peut représenter comme le rotationnel d'un potentiel vecteur A :

$$\mathrm{A} = \int \left(\frac{\mathrm{J(P)}}{cr} + \vec{m} \times \mathrm{grad}_{\mathrm{P}} \frac{1}{r} \right) dv_{\mathrm{P}}.$$

La formule classique : $\mathrm{rot}\,(f\vec{u}) = f\,\mathrm{rot}\,\vec{u} - \vec{u} \times \mathrm{grad}\, f$ permet de remplacer la partie de l'intégrale venant de la magnétisation par :

$$\int \frac{1}{r} \mathrm{rot}\, \vec{m} \cdot dv - \int \mathrm{rot} \frac{\vec{m}}{r} \cdot dv.$$

La formule classique : $\int_{\mathrm{V}} \mathrm{rot} \cdot \vec{u} dv = - \int_{\mathrm{S}} \vec{u} \times \vec{n} \cdot d\sigma$ (intégration par parties) permet de transformer la seconde intégrale et on obtient finalement :

$$\mathrm{A} = \int_{\mathrm{V}} \frac{1}{r} \left(\frac{\mathrm{J(P)}}{c} + \mathrm{rot}\, \vec{m} \right) dv + \int_{\mathrm{S}} \frac{\vec{m} \times \vec{n}}{r} d\sigma.$$

Cette formule générale montre qu'un champ d'induction magnétique quelconque B résulte d'un champ superficiel tangent à la surface des corps magnétiques et d'une densité de courant ayant pour valeur $\mathrm{J} + c\,\mathrm{rot}\,\vec{m}$. Lorsque $\mathrm{rot}\,\vec{m} \neq 0$ c'est-à-dire lorsque la magnétisation n'est pas homogène, elle apporte une contribution effective à la densité de courant déterminant l'induction magnétique, et la formule d'Ampère devient : $\mathrm{rot}\,\mathrm{B} = \dfrac{4\Pi \mathrm{J}}{c} + 4\Pi\,\mathrm{rot}\,\vec{m}$.

Combinée avec B, la magnétisation forme donc un nouveau champ de vecteurs $\mathrm{H} = \mathrm{B} - 4\Pi m$.

H est appelé le *champ magnétique*. Il possède les propriétés (courants et magnétisations *fixes*) :

$$\mathrm{rot}\,\mathrm{H} = \frac{4\Pi \mathrm{J}}{c};\ \mathrm{div}\,\mathrm{H} = 0$$

où J est la densité de courants *libres*.

La charge électrique étant conservée, on a dans le cas de courants et de magnétisations variables dans le temps, l'équation de continuité :

$$\mathrm{div}\ (\mathrm{J} + c\,\mathrm{rot}\,\vec{m}) + \frac{\partial \rho}{\partial t} = 0$$

soit $\mathrm{div}\,\mathrm{J} + \dfrac{\partial \rho}{\partial t} = \nabla \cdot \mathrm{J} + \dfrac{\partial \rho}{\partial t} = 0$.

L'équation d'Ampère reliant H et J ci-dessus n'est plus valable puisqu'elle entraîne $\dfrac{4\Pi}{c}\mathrm{div}\,\mathrm{J} = 0$ alors qu'on doit avoir : $\dfrac{4\Pi}{c}\mathrm{div}\,\mathrm{J} = -\dfrac{4\Pi}{c}\dfrac{\partial \rho}{\partial t}$. Le terme manquant s'obtient en faisant intervenir la loi de Coulomb (premère équation de Maxwell) : $\mathrm{div}\,\mathrm{D} = 4\Pi\rho$, d'où $\mathrm{div}\left(\dfrac{4\Pi}{c}\mathrm{J} + \dfrac{1}{c}\dfrac{\partial \mathrm{D}}{\partial t}\right) = 0$. C'est ce terme correctif le plus simple (on pourrait lui ajouter un rotationnel), $\dfrac{1}{c}\dfrac{\partial \mathrm{D}}{\partial t}$, qui a été introduit par Maxwell en 1865, et appelé par lui courant de déplacement.

La *quatrième équation de Maxwell* s'écrit :

$$\boxed{\nabla \times \mathrm{H} = \mathrm{rot}\,\mathrm{H} = \frac{4\Pi \mathrm{J}}{c} + \frac{1}{c}\frac{\partial \mathrm{D}}{\partial t}}.$$

Elle régit la propagation des ondes électromagnétiques dont l'existence fut effectivement prédite à partir de cette équation avant qu'elles fussent observées.

Les quatre équations de Maxwell en présence de matière peuvent être ainsi classées :

a) relations entre D *et* H : (I) $\operatorname{div} \mathrm{D} = 4\Pi\rho$ (IV) $\operatorname{rot} \mathrm{H} = \frac{4\Pi \mathrm{J}}{c} + \frac{1 \partial \mathrm{D}}{c \partial t}$.

Ces équations font intervenir les densités de charge électrique ρ et de courant électrique J, sources des champs électriques et magnétiques : ce sont les « *équations aux sources* ». Elles ne sont pas indépendantes : il existe une condition de compatibilité qu'on obtient en prenant la divergence des deux membres de (IV) ce qui donne $0 = 4\Pi \operatorname{div} \mathrm{J} + \operatorname{div} \frac{\partial \mathrm{D}}{\partial t}$ soit en tenant compte de (I) : $\operatorname{div} \mathrm{J} + \frac{\partial \rho}{\partial t} = 0$. On retrouve l'équation de continuité ce qui n'est guère surprenant puisqu'elle avait déterminé la forme de (IV).

b) Relations entre E *et* B : (II) $\operatorname{div} \mathrm{B} = 0$, (III) $\operatorname{rot} \mathrm{E} = -\frac{1}{c} \frac{\partial \mathrm{B}}{\partial t}$. Ces relations qui, on va le voir lient les composantes d'un même tenseur, sont les « *équations internes* ».

Elles non plus ne sont pas indépendantes puisqu'en prenant la divergence de (III) on obtient $\frac{\partial}{\partial t} \operatorname{div} \mathrm{B} = 0$.

Ces équations *doivent* être complétées par les relations entre D et E, J et E, H et B.

Ces relations, dites « constitutives » caractérisent la réponse du milieu matériel aux sollicitations électriques ou magnétiques. Certains corps ont un comportement exceptionnel présentant une réaction très forte et des phénomènes de rémanence ou hystérésis (substances ferromagnétiques pour l'aimantation : Fe, Ni, Co, ..., substances dites parfois « ferroélectriques » pour la polarisation : sel de Seignette). En général, pour des intensités et des fréquences raisonnables, la réponse est linéaire. Dans les milieux anisotropes : $\mathrm{D} = \varepsilon(\mathrm{E})$, soit $\mathrm{D}^i = \sum \varepsilon^i j \mathrm{E}^j$, et $\mathrm{B} = \mu(\mathrm{H})$ soit $\mathrm{B}^i = \sum \mu^i j \mathrm{H}^j$. ε et μ sont des opérateurs linéaires *symétriques* appelés respectivement perméabilité diélectrique et perméabilité magnétique. Dans des milieux isotropes, ils se réduisent à des scalaires. En l'absence de polarisation ou d'aimantation, $\varepsilon = \mu = 1$, soit : $\mathrm{D} = \mathrm{E}$ et $\mathrm{B} = \mathrm{H}$. C'est le cas du vide ou de l'air. De la même façon, on a $\mathrm{J} = \sigma(\mathrm{E})$, soit $J^i = \sum \sigma^i j E^j$, qui devient dans un milieu isotrope $J = \sigma E$ où le scalaire σ est la conductivité du corps (loi d'Ohm).

Nous allons maintenant adapter les équations de Maxwell aux principes relativistes. Outre une remarquable simplification, cette nouvelle écriture, intrinsèque, fait très vite la preuve qu'elle ne se réduit pas à un changement de formalisme car elle fait apparaître des relations *physiques* profondes, nullement évidentes *a priori*, entre diverses quantités.

Un premier point essentiel est l'invariance absolue de la charge électrique : la charge électrique est indépendante de la vitesse. Cela peut être justifié par divers raisonnements et vérifié par l'expérience (à 10^{-19} près). On a vu au paragraphe précédent que si ω est la mesure d'un volume au repos dans un repère S′, qui est, lui, animé d'une vitesse v par rapport au repère S, la mesure de ω dans S est $\overset{-1}{\gamma}\,\omega = \omega\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}$ (contraction de Lorentz). Si ρ_0 est la densité de charge « au repos » mesurée dans S, sa mesure ρ dans S est donc : $\rho = \rho_0\gamma > \rho_0$.

Considérons alors, dans un repère donné, l'équation de continuité : $\operatorname{div} J + \dfrac{\partial\rho}{\partial t} = 0$. Le vecteur densité de courant J s'exprime, en fonction de la vitesse v et de la densité ρ des charges, par $J = \rho v$. L'écriture relativiste de cette équation :

$$\operatorname{div} J + \frac{\partial\rho}{\partial t} = \frac{\partial}{\partial x^1}(\rho v x) + \frac{\partial}{\partial x^2}(\rho v y) + \frac{\partial}{\partial x^3}(\rho v z) + \frac{\partial}{\partial x^0}(c\rho) = 0$$

fait apparaître dans chaque repère standard des composantes qui s'écrivent : $(\rho c, J) = \rho(c, v) = \rho_0\gamma(c, v) = \rho_0 V$ où V est la vitesse relativiste. Ce sont donc bien les composantes dans chaque repère d'un même 4-vecteur $\mathfrak{J}$ de l'espace-temps appelé *densité de « charge-courant »*. L'équation de continuité prend la forme invariante relativiste : $\operatorname{div} \mathfrak{J} = 0$.

Donnons-en une conséquence physique : si l'on effectue le changement de repère : $S \Rightarrow S'$ déterminé par la rotation de Lorentz du plan (t, x) du paragraphe précédent VII.12, les composantes de $\mathfrak{J}$ dans S et S′ sont liées par :

$$J'_x = \gamma(J_x - \frac{v}{c}\rho);\ J'_y = J_y;\ J'_z = J_z;\ \rho' = \gamma\left(\rho - \frac{vJ_x}{c^2}\right).$$

Il en résulte que si un élément de conducteur où passe un courant de densité J a une densité de charge ρ nulle dans S, elle apparaît comme non nulle dans $S'(\rho \ni\neq 0)$. On peut suivre le même raisonnement avec la loi de Faraday (III) en partant du fait qu'elle doit avoir la même expression dans des repères en translation uniforme l'un par rapport à l'autre. Les rotations de Lorentz relatives aux plans $(t,x)(t,y)(t,z)$ montrent alors que les six composantes des champs de vecteurs E et B sont liées. Dans la rotation de Lorentz du plan (t,x) du § VII.12, on a par exemple :

$$\begin{aligned} E'_x &= E_x & B'_x &= B_x. \\ E'_y &= \gamma\left(E_y - \frac{v}{c}B_z\right) & B'_y &= \gamma\left(B_y + \frac{v}{c}E_z\right). \\ E'_z &= \gamma\left(E_z - \frac{v}{c}B_y\right) & B'_z &= \gamma\left(B_z + \frac{v}{c}E_y\right). \end{aligned}$$

On voit alors que, si dans un repère fixé, le champ électrique E et l'induction magnétique B apparaissent comme des entités distinctes, l'invariance relativiste montre qu'ils sont physiquement inséparables. On va même voir que les équations de Maxwell (II) et (III) sont deux aspects d'une même loi, la validité de l'une d'elles, $\operatorname{div} B = 0$ par exemple, dans un repère, entraînant automatiquement la loi de Faraday dans un autre. Le fait que l'on ait 6 quantités subissant des transformations linéaires par changement de repères suggère d'essayer de les considérer comme les 6 composantes d'un tenseur antisymétrique du second ordre, et de procéder de la même façon avec les champs de vecteurs D et H des équations (I) et (IV).

Considérons donc dans chaque repère standard les ensembles de composantes définis par les matrices :

$$|F^{ij}| = \begin{vmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{vmatrix},$$

$$|M^{ij}| = \begin{vmatrix} 0 & D_x & D_y & D_z \\ -D_x & 0 & H_z & -H_y \\ -D_y & -H_z & 0 & H_x \\ -D_z & H_y & -H_x & 0 \end{vmatrix}.$$

Il faut prouver que ce sont les composantes de tenseurs F (comme Faraday) et M (comme Maxwell) et pour cela montrer (chap.

II) qu'elles se transforment comme telles par un changement de repère. Il suffit de le vérifier pour la rotation de Lorentz du plan (t, x) du § précédent. Si A est la matrice de cette rotation, telle que $X' = AX$, on doit calculer $A|F^{ij}|^t A$ ce qui s'écrit :

$$\begin{vmatrix} \gamma & -\frac{v}{c}\nu & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{vmatrix} \begin{vmatrix} \gamma & -\frac{v}{c}\nu & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 0 & E_x & \gamma\left(E_y - \frac{v}{c}B_z\right) & \gamma\left(E_z + \frac{1}{c}B_y\right) \\ -E_x & 0 & \gamma\left(B_z - \frac{v}{c}E_y\right) & -\gamma\left(B_y + \frac{v}{c}E_z\right) \\ -\gamma\left(E_y - \frac{v}{c}B_z\right) & -\gamma\left(B_z - \frac{v}{c}E_y\right) & 0 & B_x \\ -\gamma\left(E_z + \frac{v}{c}B_y\right) & \gamma\left(B_y + \frac{v}{c}E_z\right) & -B_x & 0 \end{vmatrix}.$$

Ce sont exactement les formules de transformation obtenues ci-dessus. La vérification est identique pour $|M^{ij}|$. F et M sont appelés les *tenseurs électromagnétiques*.

Si $(0, e_0)$ est un « observateur instantané », il mesure par le contracté $e_0 \cdot F$, soit : $\sum(e_0)_i \cdot F^{ij}$ le champ électrique, qui est un vecteur de $e_0^\perp$ et par le contracté $e_0 \cdot ({}^*F)$ l'induction magnétique dans son repère.

On a également un « tenseur des charges » C déterminé par les champs des moments de polarisation p et de magnétisation m :

$$|C^{ij}| = \begin{vmatrix} 0 & p_x & p_y & p_z \\ -p_x & 0 & -m_z & m_y \\ -p_y & m_z & 0 & -m_x \\ -p_z & -m_y & m_x & 0 \end{vmatrix}.$$

Il est lié à F et M par la relation de Maxwell vue ci-dessus : $M = F + 4\Pi C$. La propriété pour C d'être un tenseur fournit immédiatement les formules de transformation de la polarisation et de la magnétisation lors d'un changement de repère. Une conséquence physique est qu'un corps au repos dans un repère S′ et qui y apparaît comme polarisé mais non magnétisé, apparaît dans un repère S en translation uniforme relativement à S′, comme possédant une magnétisation non nulle !

Relevons un point de détail, généralement laissé au lecteur, sur les passages entre composantes, covariantes, mixtes, et contravariantes *dans un repère standard* pour un tenseur du second-ordre de l'espace-temps. En prenant comme tenseur modèle le tenseur décomposable $t = x \otimes y$, les matrices de ses divers types de composantes sont :

$$|t^{ij}| = \begin{vmatrix} x^0y^0 & x^0y^1 & x^0y^2 & x^0y^3 \\ x^1y^0 & x^1y^1 & x^1y^2 & x^1y^3 \\ x^2y^0 & x^2y^1 & x^2y^2 & x^3y^3 \\ x^3y^0 & x^3y^1 & x^3y^2 & x^3y^3 \end{vmatrix},$$

$$|t_{ij}| = \begin{vmatrix} x_0y_0 & -x_0y_1 & -x_0y_2 & -x_0y_3 \\ -x^1y_0 & x_1y_1 & x_1y_2 & x_1y_3 \\ -x^2y_0 & x_2y_1 & x_2y_2 & x_2y_3 \\ -x^3y_0 & x_3y_1 & x_3y_2 & x_3y_3 \end{vmatrix},$$

$$|t^i_j| = \begin{vmatrix} x^0y_0 & -x^0y_1 & -x^0y_2 & -x^0y_3 \\ x^1x_0 & -x^2y_1 & -x^2y_2 & -x^2y_3 \\ x^2y_0 & -x^2y_1 & -x^2y_2 & -x^2y_3 \\ x^3y_0 & -x^3y_1 & -x^3y_2 & -x^3y_3 \end{vmatrix}.$$

Ainsi, le tenseur électromagnétique F s'écrit en composantes covariantes :

$$|\mathrm{F}_{ij}| = \begin{vmatrix} 0 & -\mathrm{E}_x & -\mathrm{E}_y & -\mathrm{E}_z \\ \mathrm{E}_x & 0 & \mathrm{B}_z & -\mathrm{B}_y \\ \mathrm{E}_y & -\mathrm{B}_z & 0 & \mathrm{B}_x \\ \mathrm{E}_z & \mathrm{B}_y & -\mathrm{B}_x & 0 \end{vmatrix}.$$

Cette remarque a son importance lorsqu'on associe à F et M les 2-formes différentielles extérieures qui leur correspondent naturellement, et que nous noterons également F et M par économie de notation :

$$\mathrm{F} = \sum_{i<j} \mathrm{F}_{ij} dx^i \wedge dx^j \qquad \mathrm{M} = \sum_{i<j} \mathrm{M}_{ij} dx^i \wedge dx^j.$$

ce qui donne pour F l'expression :

$$\mathrm{F} = \sum_{i<j} \mathrm{F}_{ij} dx^i \wedge dx^j = (\mathrm{E}_x dx + \mathrm{E}_y dy + \mathrm{E}_z dz) \wedge cdt$$
$$+(\mathrm{B}_x dy \wedge dz + \mathrm{B}_y dz \wedge dx + \mathrm{B}_z dx \wedge dy)$$

et pour la différentielle extérieure de F (en omettant les signes $\wedge$)

$$\begin{aligned} d\mathrm{F} = &\left(\frac{\partial \mathrm{B}_x}{\partial x} + \frac{\partial \mathrm{B}_y}{\partial y} + \frac{\partial \mathrm{B}_z}{\partial z}\right) dx\, dy\, dz \\ &- \left(\frac{\partial \mathrm{E}_y}{\partial x} - \frac{\partial \mathrm{E}_x}{\partial y} - \frac{1}{c}\frac{\partial \mathrm{B}_z}{\partial t}\right) cdtdxdy \\ &- \left(\frac{\partial \mathrm{E}_z}{\partial y} - \frac{\partial \mathrm{E}_y}{\partial z} - \frac{1}{c}\frac{\partial \mathrm{B}_x}{\partial t}\right) cdtdydz \\ &- \left(\frac{\partial \mathrm{E}_x}{\partial z} - \frac{\partial \mathrm{E}_z}{\partial x} - \frac{1}{c}\frac{\partial \mathrm{B}_y}{\partial t}\right) cdtdzdx. \end{aligned}$$

L'ensemble des deux équations de Maxwell (II) et (III), $\operatorname{div} \mathrm{B} = 0$ *et* $\operatorname{rot} \mathrm{E} = \dfrac{1\partial \mathrm{B}}{c\partial t}$ *est donc équivalent à* $d\mathrm{F} = 0$*, c'est-à-dire à la propriété pour* F *d'être fermée.*

Tensoriellement, l'équation $d\mathrm{F} = 0$ exprimant Maxwell (II) et (III) est un ensemble de quatre équations :

$$\boxed{\partial_i \mathrm{F}_{jk} + \partial_j \mathrm{F}_{ki} + \partial_k \mathrm{F}_{ij} = 0} \text{ pour } \forall (i,j,k) \subset (1,2,3,4).$$

A partir de la forme différentielle extérieure F, se construisent naturellement un scalaire $||\mathrm{F}||^2$ et un pseudoscalaire S par $\mathrm{F} \wedge \mathrm{F} = -\mathrm{S}\omega$ soient : $||\mathrm{B}||^2 - ||\mathrm{E}||^2$ et $\mathrm{E} \cdot \mathrm{B}$, qui sont donc invariants par les rotations de Lorentz :

$$||\mathrm{F}||^2 = \sum_{i<j} \mathrm{F}^{ij}\mathrm{F}_{ij} = ||\mathrm{B}||^2 - ||\mathrm{E}||^2$$

$$\mathrm{F} \wedge \mathrm{F} = (\mathrm{E}_x\mathrm{B}_x + \mathrm{E}_y\mathrm{B}_y + \mathrm{E}_z\mathrm{B}_z) dx\, dy\, dz\, cdt = -(\mathrm{E} \cdot \mathrm{B})\omega.$$

D'après le théorème de Poincaré (§ IV.14) l'équation $d\mathrm{F} = 0$ entraîne que dans un ouvert connexe et simplement connexe où elle est satisfaite, il existe une 1-forme Ω telle que $\mathrm{F} = d\Omega$; Ω n'est définie qu'à un « bord » ou gradient près c'est-à-dire à la différentielle près d'une fonction numérique ψ. Le 4-covecteur Ω est appelé le *potentiel ectromagnétique. L'addition à* Ω *d'une différentielle* $d\psi$ *s'appelle une transformation (ou changement) de jauge*. On a :

$$\Omega = \sum \Omega_i dx^i = \Phi cdt + \mathrm{A}_x dx + \mathrm{A}_y dy + \mathrm{A}_z dz$$

$$\mathrm{F}_{ij} = \frac{\partial \Omega_i}{\partial x^i} - \frac{\partial \Omega_i}{\partial x^j}$$

qui se décompose en : $\begin{cases} \mathrm{E} = \dfrac{1}{c}\dfrac{\partial \mathrm{A}}{\partial t} - \nabla\Phi \\ \mathrm{B} = \nabla \times \mathrm{A} = \mathrm{rot}\,\mathrm{A} \end{cases}$ où $\mathrm{A} = (\mathrm{A}_x, \mathrm{A}_y, \mathrm{A}_z)$ est le potentiel-vecteur magnétique et Φ le potentiel scalaire.

En électromagnétisme classique, seuls ont une signification physiques les F_{ij}. Cependant, le potentiel Ω se manifeste par de surprenants effets quantiques (effet Aharonov-Bohm).

Le contracté du vecteur densité de charge-courant $\mathfrak{J}$ (écrit en composantes covariantes) et du tenseur électromagnétique $\mathrm{F} = (\mathrm{F}^{ij})$ est un vecteur (contravariant) :

$$\left(\sum \mathfrak{J}_i \mathrm{F}^{ij}\right) = |\rho c - \mathrm{J}_x - \mathrm{J}_y - \mathrm{J}_z| \begin{vmatrix} 0 & \mathrm{E}_x & \mathrm{E}_y & \mathrm{E}_z \\ -\mathrm{E}_x & 0 & \mathrm{B}_z & -\mathrm{B}_y \\ -\mathrm{E}_y & -\mathrm{B}_z & 0 & \mathrm{B}_x \\ -\mathrm{E}_z & \mathrm{B}_y & -\mathrm{B}_x & 0 \end{vmatrix}$$

$$= \begin{vmatrix} \overrightarrow{E} \cdot \overrightarrow{\mathrm{J}} \\ \rho c \mathrm{E}_x + (\mathrm{J} \times \mathrm{B})_x \\ \rho c \mathrm{E}_y + (\mathrm{J} \times \mathrm{B})_y \\ \rho c \mathrm{E}_z + (\mathrm{J} \times \mathrm{B})_z \end{vmatrix}$$

qui, divisé par c est le vecteur *densité de force de Lorentz* f agissant sur les charges : $f = \dfrac{1}{c}\mathfrak{J} \cdot \mathrm{F} = \left(\dfrac{1}{c}\overrightarrow{\mathrm{E}} \cdot \overrightarrow{\mathrm{J}}\,; \rho\overrightarrow{\mathrm{E}} + \dfrac{1}{c}\overrightarrow{\mathrm{J}} \times \overrightarrow{\mathrm{B}}\right)$ $= -\dfrac{1}{c}\mathrm{F} \cdot \mathfrak{J}$.

Si $\vec{v}$ est le champ spatial des vitesses des charges électriques, $\overrightarrow{\mathrm{J}} = \rho\vec{v}$, et l'on retrouve dans la composante spatiale de :

$$f = \left(\frac{1}{c}\rho\overrightarrow{\mathrm{E}} \cdot \vec{v};\ \rho\left(\overrightarrow{\mathrm{E}} + \frac{1}{c}\vec{v} \times \overrightarrow{\mathrm{B}}\right)\right),$$

l'expression de la force de Lorentz déjà introduite.

La 0-composante est la densité de travail effectué par le champ E sur les charges, divisée par c. Puisque $\sum \mathfrak{J}_i \mathrm{F}^{ij} \mathfrak{J}_j$ est identiquement nul, et que $\mathfrak{J} = \rho \mathrm{V}$, les vecteurs d'espace-temps V et f sont orthogonaux.

Remarques : les équations du premier groupe étant homogènes en E et B (ρ et J n'interviennent pas) on peut utiliser $-\mathrm{F}$ au lieu de F sans changer l'équation $d\mathrm{F} = 0$.

Une autre possibilité (cf [SP]) est de remplacer F et M par $(-\mathrm{F})$ et $(-\mathrm{M})$ et de prendre pour codifférentielle $(-\delta)$ au lieu de δ.

Enfin, l'utilisation fréquente dans les textes anciens de la signature $(-+++)$ au lieu de $(+---)$ pour la métrique apporte des perturbations de signe dans les formules.

Les équations de Maxwell du second groupe (I) et (IV) expriment l'excitation des milieux matériels : électrique, par D, magnétique, par H, et sont mathématiquement duales des deux premières. La considération du tenseur M réunit ces deux équations en une seule. On peut écrire cette équation unique sous forme tensorielle : $\mathrm{Div}\,\mathrm{M} = \dfrac{-4\Pi}{c}\mathfrak{J}$, ou ce qui est équivalent, à l'aide des formes différentielles extérieures et de la codifférentielle : $\delta\mathrm{M} = \dfrac{4\Pi}{c}\mathfrak{J}$. Commençons par cette dernière. La 2-forme M s'écrit (en omettant les signes $\wedge$) : $\mathrm{M} = (\mathrm{D}_x dx + \mathrm{D}_y dy + \mathrm{D}_z dz)cdt + \mathrm{H}_x dydz + \mathrm{H}_y dzdx + \mathrm{H}_z dxdy$.

En se reportant à la définition de * et de δ, appliquées ici dans un repère orthonormal, on obtient par exemple (avec $cdt = dx^0$!) : $*(cdt\,dx) = -dy\,dz$; $*(dy\,dz) = cdt\,dx$; $^*(cdt\,dy\,dz) = dx$, $*(dx\,dy\,dz) = cdt$, etc... Appliquée à une 2-forme de $\mathrm{E}_{1,3}$: $\delta = (-1)^p \overset{-1}{*} d* = \overset{-1}{*} d*$; appliqué à une 3-forme : $\overset{-1}{*} = (-1)^{p(n-p)+s} * = *$, d'où $\delta = *d*$. Le calcul de $\delta\mathrm{M}$ se décompose en : $*\mathrm{M} = \mathrm{D}_x dydz + \mathrm{D}_y dzdx + \mathrm{D}_z dxdy + cdt(\mathrm{H}_x dx + \mathrm{H}_y dy \mathrm{H}_z dz)$.

$$\delta\mathrm{M} = (\nabla\cdot\mathrm{D})dxdydz + \left(\frac{1}{c}\frac{\partial\mathrm{D}_x}{\partial t} - \frac{\partial\mathrm{H}_z}{\partial y} + \frac{\partial\mathrm{H}_y}{\partial z}\right) cdt\,dydz + \dots$$

$$\delta\mathrm{M} = *d*\mathrm{M} = (\nabla\cdot\mathrm{D})cdt + \left(\frac{1}{c}\frac{\partial\mathrm{D}_x}{\partial t} - \frac{\partial\mathrm{H}_z}{\partial y} + \frac{\partial\mathrm{H}_y}{\partial z}\right) dx + \dots$$

Les composantes ainsi obtenues de la 1-forme $\delta\mathrm{M}$ sont *covariantes*. Le 4-vecteur $\mathfrak{J}$ densité de charge-courant introduit précédemment, de composantes contravariantes $(c\rho, \mathrm{J})$ a pour composantes covariantes $(c\rho, -\mathrm{J})$.

Vérifions maintenant que les équations de Maxwell (I) et (IV) s'écrivent simplement : $\boxed{\delta\mathrm{M} = \dfrac{4\Pi}{c}\mathfrak{J}}$ en prenant soin d'utiliser les composantes de même nature (ici covariantes) dans les deux membres. Les 0-composantes de δM et $\mathfrak{J}$ sont $(\delta\mathrm{M})_0 = \nabla \cdot \mathrm{D}$, $\mathfrak{J}_0 = c\rho$ d'où $\nabla \cdot \mathrm{D} = 4\Pi\rho$. Si l'on calcule les x-composantes covariantes de δM et de $\frac{4\Pi}{c}\mathfrak{J}$, on obtient : $\frac{1}{c}\frac{\partial \mathrm{D}_x}{\partial t} - \left(\frac{\partial \mathrm{H}_z}{\partial y} - \frac{\partial \mathrm{H}_y}{\partial z}\right) = -\frac{4\Pi}{c}\mathfrak{J}_x$ ce qui donne bien, avec les égalités analogues, Maxwell (IV).

Le fait que $\delta\delta = 0$ entraîne $\delta\mathfrak{J} = 0$ est l'équation de continuité ou de conservation de la charge que l'on retrouvera ci-dessous.

Voyons maintenant l'écriture tensorielle de l'équation unique remplaçant les équations de Maxwell (I) et (IV).

On peut utiliser le résultat précédent ou, beaucoup plus simplement remplacer dans (I) et (IV), D et H par les composantes contravariantes M^{ij} de M. Par exemple :

$$-\frac{1}{c}\frac{\partial \mathrm{D}_x}{\partial t} + \frac{\partial \mathrm{H}_z}{\partial y} - \frac{\partial \mathrm{H}_y}{\partial z} = \frac{\partial \mathrm{M}^{10}}{\partial x^0} + \frac{\partial \mathrm{M}^{12}}{\partial x^2} + \frac{\partial \mathrm{M}^{13}}{\partial x^3} = \frac{4\Pi}{c}\mathfrak{J}^1.$$

On obtient ainsi un ensemble de 4 équations tensorielles :

$$\boxed{\sum \frac{\partial \mathrm{M}^{\mu v}}{\partial x^v} = -\sum \frac{\partial \mathrm{M}^{v\mu}}{\partial x^v} = \frac{4\Pi}{c}\mathfrak{J}^1} \quad \mu = 0, 1, 2, 3.$$

En définissant Div M comme le vecteur de composantes $(\mathrm{div\ M})^j = \sum \frac{\partial \mathrm{M}^{ij}}{\partial x^i}$, on vérifie sur cet exemple une égalité, qui sera démontrée au § IX.12, entre l'opposée $(-\delta)$ de la codifférentielle et un opérateur de divergence qui sera défini sur les formes différentielles extérieures. Les 4 équations tensorielles précédentes se rassemblent alors en une équation vectorielle :

$$\boxed{\mathrm{div\ M} = -\frac{4\Pi}{c}\mathfrak{J}}.$$

Remarques :

a) au covecteur $\mathfrak{J} = (\rho c, -\mathrm{J})$ sont associées une 1-forme et la 3-forme duale : $j = *\mathfrak{J} = \rho c\, dx\, dy\, dz - \mathrm{J}_x c dt\, dy\, dz - \mathrm{J}_y c dt\, dz\, dx - \mathrm{J}_z c dt\, dx\, dy$, j permet de remplacer l'équation $\delta\mathrm{M} = \frac{4\Pi}{c}\mathfrak{J}$ par

$d(*\mathrm{M}) = \dfrac{4\Pi}{c}j$; et l'équation de continuité $\delta\mathfrak{J} = 0$ par $dj = 0$. On évite ainsi l'emploi de la codifférentielle.

b) on voit que c'est essentiellement le dual $*\mathrm{M}$ du tenseur M qui intervient, ce qui fait que, lorsque $\mathrm{M} = \mathrm{F}$, c'est avec le couple $(\mathrm{F}, *\mathrm{F})$ que s'écrivent les équations de Maxwell : $d\mathrm{F} = 0$, $d(*\mathrm{F}) = \dfrac{4\Pi}{c}j$.

Au tenseur M sont associés un scalaire $||\mathrm{M}||^2 = \sum_{i<j} \mathrm{M}^{ij}\mathrm{M}_{ij}$ $= ||\overrightarrow{\mathrm{H}}||^2 - ||\overrightarrow{\mathrm{D}}||^2$ et un pseudoscalaire obtenu par le carré extérieur $\mathrm{M} \wedge \mathrm{M}$, soit $\overrightarrow{\mathrm{H}} \cdot \overrightarrow{\mathrm{D}}$.

Avec les tenseurs F et M on peut former des scalaires et pseudoscalaires, qui sont invariants par les rotations de Lorentz : $(\mathrm{M}|\mathrm{F}) = \sum_{i<j} \mathrm{M}_{ij}\mathrm{F}^{ij} = \mathrm{B} \cdot \mathrm{H} - \mathrm{E} \cdot \mathrm{D}$, $(\mathrm{F}| * \mathrm{M}) = \mathrm{E} \cdot \mathrm{H} + \mathrm{B} \cdot \mathrm{D}$.

Le scalaire $\mathfrak{L} = \dfrac{-1}{8\Pi}(\mathrm{M}|\mathrm{F}) = \dfrac{1}{8\Pi}(\mathrm{E} \cdot \mathrm{D} - \mathrm{B} \cdot \mathrm{H})$ est appelé la *densité de Lagrangien* ou *l'action* du champ électromagnétique.

En électrodynamique non-relativiste, la densité d'énergie, somme des densités d'énergie magnétique et électrique a pour expression : $\mathrm{W} = \dfrac{1}{8\Pi}(\mathrm{B} \cdot \mathrm{H} + \mathrm{E} \cdot \mathrm{D})$, le malencontreux coefficient $\dfrac{1}{8\Pi}$, qui remonte à Maxwell lui-même, étant la conséquence du choix des unités gaussiennes. En supposant les opérateurs de perméabilité ε et μ constants on a, puisqu'ils sont symétriques $\dfrac{\partial \mathrm{B}}{\partial t} \cdot \mathrm{H} = \mu\left(\dfrac{\partial \mathrm{H}}{\partial t}\right) \cdot \mathrm{H} = \dfrac{\partial \mathrm{H}}{\partial t} \cdot \mu(\mathrm{H}) = \mathrm{B} \cdot \dfrac{\partial \mathrm{H}}{\partial t}$ et de même : $\dfrac{\partial \mathrm{E}}{\partial t} \cdot \mathrm{D} = \mathrm{E} \cdot \dfrac{\partial \mathrm{D}}{\partial t}$. D'où : $\dfrac{\partial \mathrm{W}}{\partial t} = \dfrac{1}{4\Pi}\left(\mathrm{E} \cdot \dfrac{\partial \mathrm{D}}{\partial t} + \mathrm{H} \cdot \dfrac{\partial \mathrm{B}}{\partial t}\right)$ ce qui conduit au théorème de l'énergie de la théorie de Maxwell classique, ou *théorème de Poynting* : on remplace les dérivées temporelles de B et D par leurs expressions tirées de Maxwell (III) et (IV) : $\dfrac{\partial \mathrm{W}}{\partial t} = \dfrac{c}{4\Pi}\{\mathrm{E} \cdot \mathrm{rot}\,\mathrm{H} - \mathrm{H} \cdot \mathrm{rot}\,\mathrm{E}\} - \mathrm{E} \cdot \mathrm{J}$ et on utilise l'identité : $\mathrm{H} \cdot \mathrm{rot}\,\mathrm{E} - \mathrm{E} \cdot \mathrm{rot}\,\mathrm{H} = \mathrm{div}(\mathrm{E} \times \mathrm{H})$, d'où :

$$-\frac{\partial \mathrm{W}}{\partial t} = \frac{c}{4\Pi}\mathrm{div}(\mathrm{E} \times \mathrm{H}) + \mathrm{E} \cdot \mathrm{J}.$$

En intégrant alors dans un volume V limité par une surface $\sum$ et en utilisant le théorème de Gauss :

$$\boxed{-\frac{\partial}{\partial t}\int_{\mathrm{V}} \mathrm{W} d\tau = \int_{\mathrm{V}} \mathrm{E}\cdot \mathrm{J} d\tau + \frac{c}{4\Pi}\int_{\sum} (\mathrm{E}\times \mathrm{H})\cdot \vec{n} d\sigma}.$$

Cette équation exprime que la perte d'énergie électromagnétique dans le volume V s'effectue de deux façons :

1) par le travail du vecteur électrique E sur les charges, travail de densité $\mathrm{E}\cdot\mathrm{J}$, qui est dissipé par effet Joule ou est transformé en énergie chimique.

2) par un flux d'énergie dissipé à travers la surface $\sum$ de V sous forme de radiation (voir ci-dessous). Sa densité volumique est la divergence du *champ de vecteurs de Poynting* S associé au champ électromagnétique : $\mathrm{S} = \frac{c}{4\Pi}\mathrm{E}\times\mathrm{H}$.

$\mathrm{W} = \frac{1}{8\Pi}(\mathrm{B}\cdot\mathrm{H} + \mathrm{E}\cdot\mathrm{D})$ n'est évidemment pas un invariant relativiste. Il n'est pas surprenant de trouver W comme composante T^{00} du tenseur d'énergie T associé au champ électromagnétique (cf § VII.12). La propriété pour T d'être un tenseur symétrique et la connaissance de sa composante T^{00} dans chaque repère standard suffisent à le déterminer et permettent de calculer ses composantes. Faraday, partant de l'impossibilité d'une action instantanée à distance, considérait que l'action des champs électrique et magnétique s'effectue au moyen de tensions et de pressions du milieu, à l'exemple des corps solides élastiques soumis à des contraintes. Maxwell donna une expression mathématique à ces tensions (tenseurs de Maxwell des tensions électrostatiques et magnétiques). Le tenseur relativiste d'énergie T rassemble toutes ces grandeurs en une seule, à savoir : la densité d'énergie W, le flux d'énergie (vecteur de Poynting), les tensions, et les densités de moment du champ. En l'absence d'aimantation et de polarisation du milieu ($\mathrm{E} = \mathrm{D}$, $\mathrm{B} = \mathrm{H}$), ce tenseur s'écrit :

$$\mathrm{T}^{ij} = \frac{1}{4\Pi}\big(\sum \mathrm{F}^{i}_{k}\mathrm{F}^{kj} + \frac{1}{2}g^{ij}\sum_{p<q}\mathrm{F}^{pq}\mathrm{F}_{pq}\big).$$

C'est une combinaison linéaire des deux seuls tenseurs symétriques du second ordre que l'on peut construire avec F. On a alors :

$$\mathrm{T}^{00} = \frac{1}{8\Pi}(\mathrm{E}^2 + \mathrm{B}^2) = \mathrm{W},$$
$$\mathrm{T}^{0j} = \mathrm{T}^{j0} = \frac{1}{c}\mathrm{S}^j = \frac{1}{4\Pi}(\mathrm{E} \times \mathrm{B})^j.$$

La partie purement spatiale de T s'écrit :

$$\frac{1}{4\Pi}\{\mathrm{E} \otimes \mathrm{E} + \mathrm{B} \otimes \mathrm{B} - \frac{1}{2}(\mathrm{E}^2 + \mathrm{B}^2)\mathrm{I}\}$$

et n'est autre que le tenseur des tensions de Maxwell.

En calculant à l'aide des équations de Maxwell la divergence du champ de tenseurs T

$$(\operatorname{div} \mathrm{T})^k = \sum_i \frac{\partial \mathrm{T}^{ik}}{\partial x^i} = f^k = \frac{1}{c}(\mathfrak{J} \cdot \mathrm{F})^k = \frac{1}{c}(\sum \mathfrak{J}_i \mathrm{F}^{ik})$$

on voit que *la divergence de* T *s'identifie au champ de forces de Lorentz* introduit plus haut. On peut d'ailleur partir de cette propriété pour déterminer T en remontant les calculs.

Remarque : si l'on inclut dans un tenseur d'énergie global toutes les énergies présentes, on obtient un tenseur de divergence nulle, cette nullité exprimant la conservation de l'énergie. Si on ne les inclut pas toutes, le tenseur obtenu a une divergence non nulle qui se manifeste comme une force agissant sur les énergies exclues.

Lorsque les champs électrique et magnétique E et H sont indépendants du temps, les équations de Maxwell sont les équations de l'électrostatique et de la magnétostatique. Lorsque les champs sont variables, il apparaît un phénomène d'émission de radiation.

Ce n'est cependant que lorsque la variation est rapide que le rayonnement émis transporte une fraction appréciable de l'énergie dissipée. A titre d'exemple, un champ électromagnétique variable satisfait aux équations : $\operatorname{div} \mathrm{E} = 0$; $\operatorname{rot} \mathrm{E} = -\frac{1}{c}\frac{\partial \mathrm{H}}{\partial t}$; $\operatorname{div} \mathrm{H} = 0$; $\operatorname{rot} \mathrm{H} = \frac{1}{c}\frac{\partial \mathrm{E}}{\partial t}$.

En utilisant l'identité : $\operatorname{rot} \operatorname{rot} \mathrm{X} = \operatorname{grad} \operatorname{div} \mathrm{X} - \Delta^2 \mathrm{X}$, on en déduit que E et H satisfont à l'équation des ondes : $\frac{1}{c^2}\frac{\partial^2 \mathrm{H}}{\partial t^2} -$

$\nabla^2 \mathrm{H} = 0$ et $\dfrac{1}{c^2}\dfrac{\partial^2 \mathrm{E}}{\partial t^2} - \nabla^2 \mathrm{E} = 0$. Ces équations s'écrivent aussi $\Box \mathrm{H} = 0$ et $\Box \mathrm{E} = 0$, où $\Box = \dfrac{1}{c^2}\dfrac{\partial^2}{\partial t^2} - \dfrac{\partial^2}{\partial x^2} - \dfrac{\partial^2}{\partial y^2} - \dfrac{\partial^2}{\partial z^2}$ est le Dalembertien, opérateur évidemment invariant par les opérations du groupe de Poincaré. Toutes les solutions (non constantes) de l'équation des ondes se propagent avec la vitesse c.

EXERCICES SUR LE CHAPITRE VII

EX. VII.1 :

Exprimer le contracté : $\sum\limits_{m} \varepsilon_{ijm}\varepsilon_{klm} = \sum \delta^{ijm}_{klm}$ sous la forme d'un déterminant par rapport aux δ^i_j et montrer qu'on peut en déduire les formules usuelles du produit vectoriel dans E_3 :

$$(x \times y) \times z = (x|z)y - (y|z)x$$
$$(x \times y | z \times u) = (x|z)(y|u) - (x|u)(y|z)$$
$$(x \times y) \times (z \times u) = (x|z \times u)y - (y|z \times u)x$$

EX. VII.2 :

Quelles sont les formes bilinéaires b sur l'espace euclidien E_n dont les composantes deux fois contravariantes b^{ij} sont celles de la forme bilinéaire inverse $\overset{-1}{b}$?

EX. VII.3 :

Soit t un tenseur du second ordre deux foix covariant symétrique de l'espace euclidien E_n. Si (t_{ij}) sont les composantes de t dans une base orthonormée de E_n, montrer que le polynôme caractéristique $\text{dét}|\lambda \mathrm{I} - t|$ a ses racines $\lambda_1, \lambda_2, \ldots \lambda_n$ réelles, et que ses coefficients ainsi que ses racines, fonctions des t_{ij}, sont des invariants. Calculer les coefficients dans le cas de E_3.

EX. VII.4 :

Soient $f = (f_1, f_2, \ldots f_n)$ une base quelconque de En, de volume v, $h = (h_1, h_2, \ldots h_n)$ la base supplémentaire. Montrer que f et h ont même orientation et que h_j s'exprime à l'aide du produit vectoriel par : $h_j = \dfrac{1}{v}(-1)^{n-j} f_1 \times f_2 \times \ldots f_{j-1} \times f_{j+1} \times \ldots f_n$.

EX. VII.5 :

Soit t un tenseur du second ordre de E_3. Dans chaque base on définit des composantes :

$$a_i = \frac{1}{2} \sum \varepsilon_{ijk} t^{jk}; \ b_{im} = \frac{1}{2} \sum \varepsilon_{ijk} \varepsilon_{mqr} t^{jq} t^{kr}$$

Etudier les propriétés tensorielles de a et de b. Cas où t est symétrique, antisymétrique, égal à $x \otimes y$.

EX. VII.6 :

Soient S un morceau de surface de E_3 défini par un paramétrage $(u, v) \to \mathrm{M}(u, v) \in \mathrm{E}_3$ de classe $\mathfrak{C}^3$, $\mathrm{E} = ||\mathrm{M}'_u||^2$, $\mathrm{F} = (\mathrm{M}'_u | \mathrm{M}'_v)$, $\mathrm{G} = ||\mathrm{M}'_v||^2$,

$$\mathrm{H} = ||\mathrm{M}'_u \wedge \mathrm{M}'_v||, \ \mathrm{N} = \frac{\mathrm{M}'_u \wedge \mathrm{M}'_v}{\mathrm{H}}$$

L, M, N les produits scalaires de N avec respectivement $\mathrm{M}''_{u^2}, \mathrm{M}''_{uv}$ et M''_{v^2}.

1) Montrer que si C est un arc de courbe orienté de S dont $\vec{t}$ et $\vec{n}$ sont les vecteurs unitaires de la tangente et de la normale principale et R le rayon de courbure, on a :

$$\frac{\vec{n} \cdot \overrightarrow{\mathrm{N}}}{\mathrm{R}} = \frac{\mathrm{L}\, du^2 + 2\mathrm{M}\, du\, dv + \mathrm{N}\, dv^2}{\mathrm{E}\, du^2 + 2\mathrm{F}\, du\, dv + \mathrm{G}\, dv^2} = \frac{1}{\mathrm{R_N}}$$

2) Montrer que si $\dfrac{1}{\mathrm{R}_1}$ et $\dfrac{1}{\mathrm{R}_2}$ désignent les extrêma de $\dfrac{1}{\mathrm{R_N}}$ en un point donné de S lorsque $\vec{t}$ tourne dans le plan tangent, on a : $\dfrac{1}{2}\left(\dfrac{1}{\mathrm{R}_1} + \dfrac{1}{\mathrm{R}_2}\right) = \dfrac{\mathrm{EN} - 2\mathrm{FM} + \mathrm{GL}}{2\mathrm{H}}$ et $\dfrac{1}{\mathrm{R}_1 \mathrm{R}_2} = \dfrac{\mathrm{LN} - \mathrm{M}^2}{\mathrm{H}}$.

3) Montrer que $\dfrac{\mathrm{LN}-\mathrm{M}^2}{\mathrm{H}}$ ne dépend *que* de $\mathrm{E}, \mathrm{F}, \mathrm{G}$ et de leurs dérivées premières et secondes (Théorème de Gauss).

EX. VII.7 :

Soient $\vec{u}, \vec{v}$ des vecteurs fixes de E_3, $\vec{r} = \overrightarrow{\mathrm{O}}\mathrm{M}$. Montrer : $\operatorname{div}[\vec{u}(\vec{v}|\vec{r})] = (\vec{u}|\vec{v})$; $\operatorname{div}(\vec{u} \times \vec{r}) = 0$; $\operatorname{div}[(\vec{r} \times \vec{u}) \times \vec{v}] = -2(\vec{u}|\vec{v})$; $\operatorname{rot}(\vec{u} \times \vec{r}) = 2\vec{u}$; $\operatorname{rot}[(\vec{u} \times \vec{r}) \times \vec{v}] = \vec{u} \times \vec{v}$.

EX. VII.8 :

Soit V un champ de vecteurs de E_3. Calculer $\Delta\overrightarrow{\mathrm{V}} = \nabla^2 \mathrm{V}$ en coordonnées sphériques.

EX. VII.9 :

Montrer que si f est une fonction harmonique dans E_3, le champ de vecteurs $f\vec{r}$ a un laplacien : $\Delta(f\vec{r}) = 0$.

EX. VII.10 :

Montrer en écrivant leur équation différentielle sous la forme $\dfrac{d}{dx}\left[(1-x^2)\dfrac{d\mathrm{P}_n}{dx}\right] = -n(n+1)\mathrm{P}_n(x)$ que les polynômes de Legendre sont orthogonaux sur $(-1, +1)$.

En déduire les relations de récurrence :

$$(n+1)\mathrm{P}_{n+1} - (2n+1) \times \mathrm{P}_n + n\mathrm{P}_{n-1} = 0$$
$$\mathrm{P}'_{n+1} - x\mathrm{P}'_n - (n+1)\mathrm{P}_n = 0$$

CHAPITRE VIII

Espaces symplectiques

Ce bref chapitre est destiné à montrer que la structure des équations d'Hamilton est « symplectique » c'est-à-dire déterminée par un tenseur antisymétrique d'ordre deux ω.

L'évolution d'un système dynamique « conservatif » n'est rien d'autre qu'une famille à un paramètre de difféomorphismes symplectiques, c'est-à-dire conservant le tenseur ω, classiquement appelés transformations canoniques. L'espace des « phases » d'un tel système est une « variété symplectique » dont la définition sera donnée au §IX.11 ainsi que le lien avec le calcul des variations.

Références :

V. GUILLEMIN et S. STERNBERG – Symplectic techniques in physics (Cambridge University Press, 1984)

N. WOODHOUSE – Geometric quantization (Oxford Mathematical Monographs), 1980.

VIII.1 – Espaces symplectiques

Définition VIII.1 :

On appelle *espace symplectique* sur le corps commutatif K, tout espace vectoriel E sur K, de dimension nécessairement paire $2n$, muni d'une *forme bilinéaire antisymétrique non-dégénérée* appelée *produit scalaire symplectique* et notée : $x, y \in \mathrm{E} \to [x \mid y]$. Le groupe que forment les automorphismes linéaires de E qui laissent invariante la forme $[|]$ est appelé le *groupe symplectique de* E et noté Sp(E) :

$$g \in \mathrm{Sp(E)} \Longleftrightarrow g \in \mathrm{Gl(E)} \text{ et } [gx \mid gy] = [x \mid y],\ \forall x, y.$$

En particulier, la forme sur K^{2n} :

$$[\mathrm{X} \mid \mathrm{Y}] = \sum_{j=1}^{n} (x^j y^{j+n} - x^{j+n} y^j) =^t \mathrm{XJY} = \sum \mathrm{J}_{kl} x^k y^l$$

$$\text{avec } \mathrm{J} = \begin{vmatrix} 0 & \mathrm{I}_n \\ -\mathrm{I}_n & 0 \end{vmatrix}$$

est appelée le *produit scalaire symplectique standard sur* K^{2n}, qui, muni de cette forme, est l'*espace symplectique standard* de dimenson $2n$ sur K.

Son groupe d'automorphismes, noté $\mathrm{Sp}(2n; \mathrm{K})$ est le groupe multiplicatif des matrices symplectiques M, caractérisées par : ${}^t\mathrm{MJM} = \mathrm{J}$.

Une base $(e_1, e_2, \ldots, e_{2n})$ d'un espace symplectique E est appelée une base symplectique lorsque le produit scalaire a dans cette base la forme standard. On note souvent $f_j = e_{n+j}$ pour $j = 1, 2, \ldots n$, d'où : $[e_i \mid f_j] = -[f_j \mid e_i] = \delta_{ij}$; $[e_i \mid e_j] = 0$, $[f_i \mid f_j] = 0$.

Les sous-espaces de base e et de base f sont des sous-espaces totalement isotropes maximaux (ou « lagrangiens ») de E.

On démontre immédiatement par récurrence que tout espace symplectique possède une base symplectique, qui détermine un isomorphisme de cet espace sur l'espace symplectique standard de même dimension : tous les espaces symplectiques de même dimension sur un même corps K sont isomorphes.

Exemples :

1) Le plan symplectique standard K^2, où

$$[\mathrm{X} \mid \mathrm{Y}] = x^1 y^2 - x^2 y^1 = \text{dét} \begin{vmatrix} x^1 & y^1 \\ x^2 & y^2 \end{vmatrix}$$

est la forme qui mesure l'« aire algébrique » de deux vecteurs relativement à la base.

Son groupe symplectique est $\mathrm{Sp}(2;\mathrm{K}) = \mathrm{Sl}(2;\mathrm{K})$.

2) Sur la somme directe d'un espace vectoriel V de dimension finie n et de son dual V^*, on définit un produit symplectique par :

$$[x \oplus \alpha \mid y \oplus \beta] = \langle x, \beta \rangle - \langle y, \alpha \rangle.$$

Une base $(e_1, e_2, \ldots, e_n)$ de V et la base duale $(f_1 = e^{*1}, f_2 = e^{*2}, \ldots f_n)$ de V^* forment, en identifiant V à $\mathrm{V} \oplus 0$ et V^* à $0 \oplus \mathrm{V}^*$, une base symplectique de $\mathrm{V} \oplus \mathrm{V}^*$. Tout espace symplectique est isomorphe à un tel espace $\mathrm{V} \oplus \mathrm{V}^*$.

3) Soient F un espace vectoriel complexe de dimension n, $\mathrm{E} = {}_{\mathbb{R}}\mathrm{F}$ l'espace vectoriel réel de dimension $2n$ obtenu par restriction des scalaires. Le scalaire i de F devient sur E un opérateur linéaire de carré $i^2 = -$Identité, et les structures F et $({}_{\mathbb{R}}\mathrm{F}; i)$ sont identiques.

Nous allons voir que la *partie imaginaire d'un produit scalaire pseudohermitien sur* F *détermine sur* ${}_{\mathbb{R}}\mathrm{F}$ *une structure symplectique invariante par* i.

Si f est une forme sesquilinéaire sur F : $f(x, y) \in \mathbb{C}$; $f(\lambda x, y) = \lambda f(x, y)$; $f(x, \mu y) = \overline{\mu} \cdot f(x, y)$ elle est invariante par i : $f(ix, iy) = f(x, y)$. On peut l'écrire : $f(x, y) = \mathrm{R}(x, y) + i\mathrm{I}(x, y)$ où R et I sont des formes bilinéaires réelles sur ${}_{\mathbb{R}}\mathrm{F}$ invariantes par i. Dès lors, un calcul immédiat montre que les trois propriétés suivantes sont équivalentes :

a) f est hermitienne : $f(x, y) = \overline{f(x, y)}$;

b) R est symétrique;

c) I est antisymétrique. Dans ce cas R ou I détermine f car alors : $\mathrm{I}(x, y) = \mathrm{R}(x, iy)$ et $\mathrm{R}(x, y) = \mathrm{I}(ix, y)$.

Si f est un produit scalaire pseudohermitien sur F c'est-à-dire une forme hermitienne non-dégénérée, R et I sont des formes bilinéaires réelles non-dégénérées, et déterminent sur ${}_{\mathbb{R}}\mathrm{F}$: R une structure pseudoeuclidienne et I une structure symplectique. Le cas usuel est celui où f est un produit scalaire *hermitien* : $f(x, x) > 0$, $\forall x \neq 0$. Si $e = (e_1, e_2, \ldots, e_n)$ est une base orthonormée de F, $(e_1, e_2, \ldots, e_n, ie_1 = e_{n+1}, ie_2 = e_{n+2}, \ldots, ie_n = e_{2n})$ est alors une base de ${}_{\mathbb{R}}\mathrm{F}$ qui est à la fois orthonormée pour la structure euclidienne définie par R et symplectique pour la structure symplectique définie par I.

Réciproquement, si E est un espace symplectique réel et $e = (e_1, \dots e_n, e_{n+1}, \dots e_{2n})$ une base symplectique de E, soit j_e l'opérateur linéaire de E ayant pour matrice dans e l'*opposée* $(-\mathrm{J})$ de la matrice J du produit $[||]$. On a $j_e^2 = -$Identité et $[x \mid y] = [j_e x \mid j_e y]$. L'opérateur j_e que nous noterons simplement j définit sur E une structure complexe pour laquelle $(e_1 e_2 \dots e_n)$ est une base complexe. La base réelle associée : $(e_1, e_2, \dots j e_1, j e_2 \dots)$ est la base symplectique de départ $(j e_k = -\mathrm{J} e_k = e_{n+k}\,;\, j e_{n+k} = -e_k)$. Pour un opérateur linéaire j de E, deux des trois conditions suivantes entraînent la troisième : *a)* $[jx \mid jy] = [x \mid y]$, $\forall x, y$ $(j \in \mathrm{Sp(E)})$; *b)* j est antisymétrique : $[jx \mid y] = -[x \mid jy]$; *c)* $j^2 = -\mathrm{Id}$. Un tel opérateur définit sur E une structure complexe dite *compatible* avec le produit symplectique. La forme $f(x, y) = [x \mid jy] - i[x \mid y]$ est alors un produit scalaire pseudohermitien sur $(\mathrm{E}; j)$. j est dit *positif* si la forme bilinéaire symétrique réelle $[x \mid jy]$ est strictement positive, ce qui est le cas pour l'opérateur j_e défini ci-dessus.

Un produit scalaire symplectique sur un espace vectoriel réel E peut donc toujours être considéré comme la partie imaginaire d'un produit scalaire hermitien relatif à une structure complexe sur E.

La donnée d'un produit scalaire symplectique sur un espace vectoriel E de dimension $2n$ sur K est équivalente à la donnée de la forme extérieure associée ω, de degré deux et de rang maximum n. Dans une base symplectique $(e_1, \dots e_n f_1 \dots f_n)$, la forme ω s'écrit à l'aide de la base duale : $\omega = \sum_j e^{*j} \wedge f^{*j}$. Les automorphismes symplectiques sont les automorphismes linéaires g tels que $(\wedge^2 \check{g})\omega = \omega$ (où $\check{g} = {}^t \overset{-1}{g}$, est le contragrédient de g). Les puissances extérieures de ω jusqu'à la n-ème sont des formes non nulles de degrés $4, \dots 2n$, invariantes par Sp(E). ω^n définit canoniquement une unité de volume de E. Afin d'éliminer les coefficients parasites, on prend pour l'unité de volume symplectique sur E : $v = (-1)^{n-1} \dfrac{\omega^n}{n!}$ $= e^{*1} \wedge \dots \wedge e^{*n} \wedge f^{*1} \wedge \dots \wedge f^{*n}$ pour toute base symplectique de E. Puisque $(\wedge^{2n} \check{g}) \cdot v = (\text{dét } \check{g}) \cdot v$, il en résulte que tout automorphisme symplectique est de déterminant $+1$. Si K est le corps réel, v définit une orientation canonique sur E. Un produit scalaire symplectique sur E détermine un isomorphisme de E sur son dual E^* en faisant correspondre à chaque vecteur X de E la forme ζ image de X par l'application ρ *associée à droite* de la forme bilinéaire :

$$\omega(\mathrm{Y},\mathrm{X}) = [\mathrm{Y} \mid \mathrm{X}] = \langle \mathrm{Y}, \rho(\mathrm{X})\rangle = \langle \mathrm{Y}, \zeta\rangle = \zeta(\mathrm{Y}).$$

Exprimée à l'aide de la forme ω : $\zeta = -i(\mathrm{X})\omega = -\mathrm{X} \lrcorner\, \omega$ (l'application γ associée à gauche est évidemment égale à $-\rho$).

Réciproquement, à chaque forme linéaire $\alpha \in E^*$ le produit symplectique fait correspondre un vecteur $\mathrm{A} = \overset{-1}{\rho}\alpha$ tel que $[. \mid \mathrm{A}] = \alpha$.

VIII.2 – Equations d'Hamilton

Nous supposerons dans tout ce qui suit que K *est le corps réel.* Au champ de vecteurs X sur l'espace symplectique réel E correspond la 1-forme différentielle $\zeta = -i(\mathrm{X})\omega$ et à la 1-forme différentielle α un champ de vecteurs A.

En particulier si α est la différentielle d'une fonction numérique $\mathfrak{C}^1$ sur E : $\alpha = df$, le champ de vecteurs G_f qui lui est associé est le *gradient symplectique de f* (cf. §IX.11). La valeur de la différentielle pour le vecteur Y : $df(\mathrm{Y}) = \mathrm{Y}\cdot f = [\mathrm{Y} \mid \mathrm{G}_f]$ est le produit symplectique de Y et de G_f : $df = -i(\mathrm{G}_f)\omega$.

Les champs de vecteurs sur E qui sont les gradients symplectiques des fonctions numériques $\mathfrak{C}^\infty$ de E sont appelés *champs hamiltoniens*.

Adoptons les notations de la mécanique pour les coordonnées relatives à une base symplectique de E : $(x^1, x^2, \dots x^{2n}) = (p_1 p_2 \dots p_n; q^1 \dots q^n)$ avec $\omega = \sum dp_j \wedge dq^j$ (coordonnées dites *normales*). On a alors : $df = \sum \dfrac{\partial f}{\partial q^j}\, dq^j + \sum \dfrac{\partial f}{\partial p_k}\, dp_k$ d'où : $\mathrm{G}_f = \sum \dfrac{\partial f}{\partial p_j}\,\dfrac{\partial}{\partial q^j} - \sum \dfrac{\partial f}{\partial q^k}\,\dfrac{\partial}{\partial p_k}$.

Si l'on note pour simplifier D_f le champ de vecteurs dont les composantes dans une base symplectique forment la matrice colonne des dérivées partielles de f

$$\mathrm{D}_f = \begin{vmatrix} \dfrac{\partial f}{\partial q^1} \\ \dfrac{\partial f}{\partial q^2} \\ \vdots \\ \dfrac{\partial f}{\partial p^1} \\ \dfrac{\partial f}{\partial p^2} \\ \vdots \end{vmatrix},$$

on a $\mathrm{G}_f = \mathrm{J} \cdot \mathrm{D}f$.

Nous noterons également $\mathrm{D}^2 f$ la matrice symétrique des dérivées secondes des f relativement aux coordonnées (q^j, p_k).

Les trajectoires du champ de vecteurs G_f de E sont les solutions du système différentiel formé des $2n$ équations :

$$\frac{dx}{du} = \mathrm{G}_f \text{ soit : } \frac{dq^j}{du} = \frac{\partial f}{\partial p_j}; \; \frac{dp_k}{du} = -\frac{\partial f}{\partial q^k}$$

dites *équations d'Hamilton.*

Définition VII.2 :

Soit $(\mathrm{E}; \omega)$ un espace symplectique réel. On appelle *transformation canonique* tout difféomorphisme d'un ouvert U sur un ouvert U′ qui respecte la structure symplectique, autrement dit, qui conserve la forme extérieure ω.

Exemple : Les automorphismes symplectiques sont les transformations canoniques linéaires.

Lemme VIII.2 : Un difféomorphisme Φ d'un ouvert U sur un ouvert U′ d'un espace symplectique réel $(\mathrm{E}; \omega)$ est une transformation canonique si et seulement si sa différentielle $\mathrm{D}\Phi_x$ est, en chaque point x de U, un automorphisme symplectique.

Preuve : En effet la conservation par Φ de la structure symplectique s'exprime par l'égalité $[\mathrm{D}\Phi_x \cdot \mathrm{X} \mid \mathrm{D}\Phi_x \cdot \mathrm{Y}] = [\mathrm{X} \mid \mathrm{Y}]$ quels que soient les vecteurs X, Y et $x \in \mathrm{U}$.

Théorème VIII.2 : Un système d'équations d'Hamilton est équivalent à une famille à un paramètre de transformations canoniques.

Preuve : La preuve classique consiste à montrer que ω reste constante le long d'une solution du système en vérifiant que $\dfrac{\partial\omega}{\partial u} = \sum\left(d\,\dfrac{\partial p_j}{\partial u}\wedge dq^j + dp_j\wedge d\,\dfrac{\partial q^j}{\partial u}\right)$ est identiquement nul, ce qui est immédiat.

Nous allons voir une autre démonstration, plus élaborée, mais dont le développement conduit de façon naturelle à diverses applications.

De façon générale, un système d'équations différentielles dans $\mathbb{R}^m$:

$$\frac{dx}{du} = \mathrm{X} \text{ soit : } \frac{dx^j}{du} = \mathrm{X}^j(x^1, x^2, \dots x^m) \text{ pour } j = 1, 2, \dots m$$

détermine par sa solution, au voisinage de x_0 : $x = x(u; x_0)$ soit $x^j = x^j(u; x_{1,0}, x_{1,0}, \dots x_{m,0})$ pour $j = 1, 2, \dots m$ un difféomorphisme local Φ pour chaque u assez petit : $x_0 \to x(u, x_0) = \Phi(u; x_0)$ se réduisant à l'identité pour $u = 0$. Φ représente la « translation » le long des courbes intégrales, trajectoires du champ de vecteurs X. Si DΦ désigne la différentielle de Φ (à u constant), en prenant les différentielles relatives à x_0 des deux membres de l'identité : $\dfrac{\partial}{\partial u}\,\Phi(u, x_0) = \mathrm{X}(\Phi(u, x_0))$, on obtient l'équation : $\dfrac{\partial}{\partial u}\,\mathrm{D}\Phi = \mathrm{DX}\cdot\mathrm{D}\Phi$ dans $\mathfrak{L}(\mathrm{E})$.

Nous allons l'interpréter dans le cas des équations d'Hamilton. On a vu au §VIII.1 que le groupe symplectique réel Sp$(2n, \mathbb{R})$ est le groupe multiplicatif des matrices M satisfaisant à ${}^t\mathrm{MJM} = \mathrm{J}$. Soit : $u \in [0, 1] \to \mathrm{M}(u)$ un chemin différentiable issu de l'élément neutre : $\mathrm{M(O)} = \mathrm{I}$ dans Sp$(2n; \mathbb{R})$. En différentiant pour $u = 0$ l'égalité : ${}^t\mathrm{M}(u)\mathrm{JM}(u) = \mathrm{J}$ on obtient : ${}^t\mathrm{M}'(\mathrm{O})\mathrm{J} + \mathrm{JM}'(0) = 0$. Les matrices H vérifiant ${}^t\mathrm{HJ} + \mathrm{JH} = 0$ sont les vecteurs tangents en l'élément neutre de Sp$(2n; \mathbb{R})$. L'espace vectoriel qu'elles forment est stable par le crochet $[\mathrm{H}, \mathrm{K}] = \mathrm{HK} - \mathrm{KH}$: c'est l'algèbre de Lie $\mathfrak{S}p(2n; \mathbb{R})$ du groupe Sp$(2n; \mathbb{R})$.

En différentiant l'égalité ${}^t\mathrm{M}(u)\mathrm{JM}(u) = \mathrm{J}$ pour u quelconque, on obtient : ${}^t\mathrm{M}'(u)\mathrm{JM}(u) +^t \mathrm{M}(u)\mathrm{JM}'(u) = 0$ qui s'écrit

${}^t\{\mathrm{M}'(u)\overset{-1}{\mathrm{M}}(u)\}\mathrm{J}+\mathrm{J}\{\mathrm{M}'(u)\overset{-1}{\mathrm{M}}(u)\} = 0$ soit $\mathrm{M}'(u)\overset{-1}{\mathrm{M}}(u) \in \mathfrak{S}p(2n;\mathbb{R})$, quel que soit u.

Réciproquement si $u \to \mathrm{M}(u)$ est un chemin différentiable dans $\mathrm{Gl(E)}$ tel que $\mathrm{M}(0) = \mathrm{I}$ et $\mathrm{M}'(u)\overset{-1}{\mathrm{M}}(u) \in \mathfrak{S}p(2n;\mathbb{R})$ quel que soit u, en remontant les calculs précédents, on obtient : $\dfrac{d}{du}\{{}^t\mathrm{M}(u)\mathrm{JM}(u)\} = 0$, d'où ${}^t\mathrm{M}(u)\mathrm{JM}(u) = \mathrm{J}$ quel que soit u, ce qui prouve que le chemin est contenu dans le groupe $\mathfrak{S}p(2n,\mathbb{R})$.

Or, si l'on écrit une matrice H de $\mathfrak{S}p(2n;\mathbb{R})$ en blocs $n \times n$: $\mathrm{H} = \begin{vmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{vmatrix}$, l'égalité ${}^t\mathrm{HJ} + \mathrm{JH} = 0$ est équivalente aux trois égalités : ${}^t\mathrm{B} = \mathrm{B}$, ${}^t\mathrm{C} = \mathrm{C}$, $\mathrm{D} = -{}^t\mathrm{A}$

On constate alors que la différentielle DG_f du gradient symplectique de la fonction f : $\mathrm{DG}_f = \mathrm{D}(\mathrm{J}\, df) = \mathrm{JD}^2 f$ est une matrice de l'algèbre de Lie $\mathfrak{S}p(2n;\mathbb{R})$. En effet :

$$ {}^t\mathrm{DG}_f \cdot \mathrm{J} + \mathrm{J} \cdot \mathrm{DG}_f = \mathrm{D}^2 f\, {}^t\mathrm{J} \cdot \mathrm{J} + \mathrm{J} \cdot \mathrm{JD}^2 f = 0 $$

ce que l'on vérifie directement :

$$ \mathrm{DG}_f = \begin{vmatrix} \left\| \dfrac{\partial^2 f}{\partial p^j \partial q^k} \right\| & \left\| \dfrac{\partial^2 f}{\partial p^j \partial p^k} \right\| \\ -\left\| \dfrac{\partial^2 f}{\partial q^j \partial q^k} \right\| & -\left\| \dfrac{\partial^2 f}{\partial q^k \partial p^j} \right\| \end{vmatrix}. $$

L'équation satisfaite par le difféomorphisme local d'un système d'Hamilton : $\dfrac{\partial}{\partial u}\mathrm{D}\Phi = \mathrm{DG}_f \cdot \mathrm{D}\Phi$ soit $\mathrm{D}\Phi' \cdot \mathrm{D}\Phi,^{-1} = \mathrm{DG}_f$ exprime donc que $u \to \mathrm{D}\Phi_u$ est un chemin dans le groupe $\mathrm{Sp}(2n;\mathbb{R})$. D'après le lemme ci-dessus, Φ_u est une transformation canonique pour chaque u, c.q.f.d.

Cas particulier : La matrice DG_f étant formée des dérivées secondes de la fonction f est une matrice constante si et seulement si f est une fonction polynomiale des (q^j, p_k) de degré $\leqslant 2$. Les termes constants et de degré un de f disparaissent dans DG_f. Considérons alors l'espace vectoriel Q des polynômes *homogènes* quadratiques en les variables (q^j, p_k). Les composantes du gradient symplectique G_f d'un tel polynôme f sont linéaires homogènes, DG_f est une matrice constante A et le vecteur G_f a pour valeur au point x (de coordonnées q^j, p_k) : $G_f = DG_f \cdot x = Ax$.

Les équations d'Hamilton associées au polynôme f : $\dfrac{dx}{du} = G_f = A \cdot x$ ont pour solution le sous-groupe à un paramètre $x = \exp Au$ du groupe symplectique $\mathfrak{S}p(2n; \mathbb{R})$. L'isomorphisme linéaire de l'espace vectoriel Q sur l'algèbre de Lie $\mathfrak{S}p(2n; \mathbb{R})$ sera complété au paragraphe suivant.

VIII.3 – Crochets et algèbres de Poisson

Si $\mathfrak{F}$ est l'algèbre associative des fonctions numériques $\mathfrak{C}^\infty$ sur l'espace symplectique E, les considérations précédentes permettent de définir sur $\mathfrak{F}$ une multiplication interne, dite *crochet de Poisson*, par :
si $f, g \in \mathfrak{F}$: $\{f, g\} = [G_f \mid G_g] = G_f \cdot g = -G_g \cdot f$ soit en coordonnées normales :

$$\{f, g\} = \sum \Big(\frac{\partial f}{\partial p_i}\,\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\,\frac{\partial g}{\partial p_i}\Big).$$

On définit ainsi sur $\mathfrak{F}$ une structure d'algèbre de Poisson conformément à la définition suivante :

Définition VIII.3.A :

Une *algèbre de Poisson* $\mathfrak{A}$ est un espace vectoriel muni de *deux* structures d'algèbres :

1) une structure d'algèbre associative commutative unitaire.

2) une structure d'algèbre de Lie, dont le crochet est généralement noté par des accolades $\{\quad\}$, compatible avec la première dans le sens suivant : les translations à gauche de la structure d'algèbre de Lie : $h \to \{a, h\} = \operatorname{ad} a \cdot h$ sont des dérivations de la structure d'algèbre associative :

$$\{a, bc\} = \{a, b\}c + b\{a, c\}$$

(d'où il résulte que $\{a, 1\} = 0\ \forall a$).

Nous allons montrer que le crochet de Poisson sur $\mathfrak{F}$ est bien celui d'une algèbre de Lie et que l'application $f \in \mathfrak{F} \to \mathrm{G}_f$ est un homomorphisme de l'algèbre de Lie de $\mathfrak{F}$ dans l'algèbre de Lie des champs de vecteurs. En effet, la définition du crochet de deux champs de vecteurs $[\mathrm{X}, \mathrm{Y}]$ et la formule suivante :

$$\begin{aligned} d\omega(\mathrm{X},\mathrm{Y},\mathrm{Z}) = {} & \mathrm{X}\cdot\omega(\mathrm{Y},\mathrm{Z}) + \mathrm{Y}\cdot\omega(\mathrm{Z},\mathrm{X}) + \mathrm{Z}\cdot\omega(\mathrm{X},\mathrm{Y}) \\ & -\omega([\mathrm{X},\mathrm{Y}],\mathrm{Z}) - \omega([\mathrm{Y},\mathrm{Z}],\mathrm{X}) - \omega([\mathrm{Z},\mathrm{X}],\mathrm{Y}) \end{aligned}$$

vus au §IV.14 permettent d'écrire, pour $f, g, h \in \mathfrak{F}$:

$$\begin{aligned} [\mathrm{G}_f, \mathrm{G}_g]\cdot h &= \mathrm{G}_f\cdot(\mathrm{G}_g\cdot h) - \mathrm{G}_g\cdot(\mathrm{G}_f\cdot h) \\ &= \mathrm{G}_f\cdot\{g,h\} - \mathrm{G}_g\cdot\{f,h\} \\ &= \{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} + \mathrm{G}_{\{f,g\}}\cdot h \end{aligned}$$

tandis que :

$$0 = d\omega(\mathrm{G}_f, \mathrm{G}_f, \mathrm{G}_h) = \{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\} \quad \text{c.q.f.d.}$$

Exemples :

1) Le crochet de Poisson permet de considérer une fonction numérique $\mathfrak{C}^\infty$: f, comme opérateur différentiel opérant sur l'algèbre des fonctions différentiables par : $\mathrm{G}_f\cdot g = \{f,g\}$ (première étape de l'opération de « quantification »). On peut ainsi écrire l'équation d'évolution d'un système mécanique décrit par n paramètres $q^1, q^2, \ldots q^n$, leurs « moments » associés $p_1, p_2, \ldots p_n$ et son énergie totale ou hamiltonien H sous la forme :

$$\frac{dq^j}{du} = \frac{\partial \mathrm{H}}{\partial p_j} = \{\mathrm{H}, q^j\} \qquad \frac{dp_j}{du} = -\frac{\partial \mathrm{H}}{\partial q^j} = \{\mathrm{H}, p_j\}$$

L'équation d'évolution d'une fonction $\mathrm{F}(q^1, q^2, \ldots p_1, p_2)$ est analogue : $\dfrac{\mathrm{DF}}{du} = \{\mathrm{H}, \mathrm{F}\}$ et H est ainsi interprété comme un opérateur.

2) L'expression du crochet de Poisson en coordonnés normales montre que sur l'espace symplectique réel $\mathbb{R}^{2n}(\omega = \sum_i dp_i \wedge dq^i)$ les polynômes quadratiques homogènes forment une sous-algèbre de Lie $\mathfrak{G}$ de l'algèbre de Poisson $\mathfrak{P}(\mathbb{R}^{2n}, \omega)$. Les gradients symplectique de ces polynômes sont des champs de vecteurs linéaires sur

$\mathbb{R}^{2n}$, hamiltoniens, donc laissant la forme ω Lie-invariante, ce qui détermine un isomorphisme de $\mathfrak{G}$ sur $\mathfrak{S}p(2n;\mathbb{R})$.

En particulier dans l'algèbre de Poisson de $\mathbb{R}^2$ muni des coordonnées (p,q) et de la forme $\omega = dp\wedge dq$ les polynômes quadratiques forment une sous-algèbre de Lie isomorphe de $\mathfrak{S}l(2;\mathbb{R}) = \mathfrak{S}p(2;\mathbb{R})$

Si $a = \frac{1}{2}\, q^2$, $b = -\frac{1}{2}\, p^2$, $c = -qp$

$$\mathrm{G}_a = -q\,\frac{\partial}{\partial p},\ \mathrm{G}_b = -\frac{\partial}{\partial p},\ \mathrm{G}_c = p\,\frac{\partial}{\partial p} - q\,\frac{\partial}{\partial q}.$$

L'hamiltonien de l'oscillateur harmonique : $\mathrm{H} = \frac{1}{2}\,(p^2+q^2)$ est représenté par la rotation hyperbolique infinitésimale : $p\,\frac{\partial}{\partial q} - q\,\frac{\partial}{\partial p}$

3) Une autre sous-algèbre de Lie de $\mathfrak{P}(\mathbb{R}^{2n},\sigma)$ est formée par les fonctions affines sur $\mathbb{R}^{2n}$.

Le crochet de $f = a + \sum \lambda^i q_i + \sum \mu^j p_j$ et $f' = a' + \sum \lambda'^i q_i + \sum \mu'^j p_j$ est la constante : $\{f, f'\} = \sum(\lambda'^i \mu^i - \lambda^i \mu'^i)$

Cette algèbre de Lie peut donc encore être décrite par l'espace vectoriel $\mathrm{L} = \mathbb{R} \oplus \mathbb{R}^{2n}$, somme directe de l'espace symplectique $(\mathbb{R}^{2n},\sigma)$ et des constantes muni du crochet : $[a+x, b+y] = \sigma(x,y) + 0$. C'est une *algèbre d'Heisenberg* d'ordre n.

VIII. 4 – Algèbre et groupe d'Heisenberg associés à un espace symplectique

Soit $\mathfrak{A}$ une algèbre associative unitaire (sur un corps K de caractéristique nulle !). Un élément a de $\mathfrak{A}$ est nilpotent s'il existe un entier m tel que $a^m = 0$. Un élément u de $\mathfrak{A}$ est unipotent si $u - 1 = a$ est nilpotent. Une série entière $\mathrm{S}(x)$ devient un polynôme en x lorsque x est nilpotent. D'où les propriétés suivantes :

1) tout élément unipotent $u = 1 + a$ est inversible et $\overset{-1}{u} = (1+a)^{-1} = 1 - a + a^2 + \ldots + (-1)^m a^m$ si $a^{m+1} = 0$.

2) l'exponentielle d'un élément nilpotent a est un élément unipotent :

$$\exp a = 1 + a\,\frac{a^2}{2} + \ldots + \frac{a^m}{m!} \text{ si } a^{m+1} = 0.$$

3) le logarithme d'un élément unipotent $u = 1+a$ est un élément nilpotent :

$$\operatorname{Log} u = \operatorname{Log}(1+a) = \sum_{n=1} (-1)^{n-1} \frac{a^n}{n}.$$

4) les identités entre séries entières donnent des identités entre polynômes lorsque l'argument est nilpotent :

– si a est nilpotent : $\operatorname{Log}(\exp a) = a$

– si u est unipotent : $\exp(\operatorname{Log} u) = u$

L'exponentielle et son application réciproque, le logarithme, établissent une bijection entre l'ensemble des éléments nilpotents et l'ensemble des éléments unipotents de $\mathfrak{A}$.

- si a et b sont nilpotents et $ab = ba$, $(a+b)$ est nilpotent, et $\exp(a+b) = \exp a \exp b$.
- si u et v sont unipotents et $uv = vu$, uv est unipotent et $\log(uv) = \log u + \log v$.

Exemples : Les matrices $n \times n$ surtriangulaires ($a_{ij} = 0$ si $i \geqslant j$) sont des éléments nilpotents de l'algèbre $\mathrm{K}(n)$ des matrices $n \times n$ sur K, et forment une algèbre de Lie $\mathfrak{T}(n;\mathrm{K})$ nilpotente. Les matrices triangulaires supérieures dont la diagonale est formée de 1 ($a_{ii} = 1, \forall i$, et $a_{ij} = 0$ si $i > j$) sont des éléments unipotents de $\mathrm{K}(n)$ et forment avec le produit, un groupe nilpotent $\mathrm{T}(n;\mathrm{K}) \subset \mathrm{Gl}(n;\mathrm{K})$. L'exponentielle et son application réciproque le logarithme établissent une bijection entre l'algèbre de Lie $\mathfrak{T}(n;\mathrm{K})$ et le groupe $\mathrm{T}(n;\mathrm{K})$ par : $a \in \mathfrak{T}(n;\mathrm{K}) \to \exp a \in \mathrm{T}(n;\mathrm{K})$.

Des propriétés analogues valent pour les matrices triangulaires inférieures.

Définition VIII.4 :

On appelle *algèbre d'Heisenberg d'ordre* n sur la corps K, l'algèbre de Lie $\mathfrak{h}_n$ nilpotente de dimension $(2n+1)$ obtenue à partir d'un espace symplectique E de dimension $2n$ sur K, de forme symplectique b, en prenant $\mathfrak{h}_n = \mathrm{E} \oplus \mathrm{K}$ et en y définissant le crochet par :

$$[x + \alpha \cdot 1, y + \beta \cdot 1] = 0 + b(x,y)1.$$

L'algèbre d'Heisenberg de dimension infinie $\mathfrak{h}_\infty$ est définie par une base $(e_i, e'_j : i, j = 1, 2, \ldots \infty)$ et la même table de multiplication : $[e_i, e'_j] = \delta_{ij}$.

On a une représentation linéaire *irréductible de dimension infinie* de l'algèbre de Lie nilpotente $\mathfrak{h}_n$ comme algèbre de Lie d'opérateurs linéaires sur l'espace vectoriel des polynômes de n variables :

$\mathrm{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$, en faisant correspondre à l'unité de K et aux vecteurs d'une base symplectique de E :

$$e_i \to \frac{\partial}{\partial \mathrm{X}_i},\ e'_j \to \text{ multiplication par } \mathrm{X}_j.$$

On a de la même façon une représentation linéaire irréductible de l'algèbre $\mathfrak{h}_\infty$ par les opérateurs analogues opérant dans l'espace vectoriel des polynômes à une infinité de variables $\mathrm{K}[\mathrm{X}_1, \mathrm{X}_2, \ldots]$

Dans la représentation linéaire précédente de $\mathfrak{h}_n$ par des matrices nilpotentes de $\mathrm{K}(n+2)$, l'exponentiation dans $\mathrm{K}(n+2)$ s'écrit :

$$\exp\{x + \alpha \cdot 1\} = \exp a = \mathrm{I} + a + \frac{a^2}{2} \begin{vmatrix} 1 & x^1 & x^n & \alpha + \dfrac{(u \mid u')}{2} \\ 0 & 1 & 0 & x'^1 \\ 0 & \ldots & 0 & x'^n \\ 0 & \ldots & 0 & 1 \end{vmatrix},$$

$$\exp\{x + \alpha \cdot 1\} \cdot \exp\{y + \beta \cdot 1\} = \exp\Big\{x + y + \Big(\alpha + \beta + \frac{b(x,y)}{2}\Big) \cdot 1\Big\}.$$

Les matrices de la forme $\{\exp x + \alpha \cdot 1\}$ forment ainsi un groupe H_n, sous-groupe de $\mathrm{T}(n+2; \mathrm{K})$, appelé *groupe d'Heisenberg associé* *à l'espace symplectique* (E, Φ). On peut évidemment définir directement H_n, sans faire appel à la représentation linéaire, comme l'ensemble $\mathrm{E} \oplus \mathrm{K}$ muni de la multiplication:

$$(x + \alpha \cdot 1) \cdot (y + \beta \cdot 1) = \Big\{(x + y) + \Big(\alpha + \beta + \frac{b(x, y)}{2}\Big) \cdot 1\Big\}.$$

Pour exprimer le fait que H_n est obtenu par exponentiation de $\mathfrak{h}_n$ on dit que $\mathfrak{h}_n$ est l'algèbre de Lie du groupe H_n.

Le groupe symplectique $\mathrm{Sp}(\mathrm{E})$, d'invariance de la forme b, opère sur $\mathfrak{h}_n = \mathrm{E} \oplus \mathrm{K}$ en respectant le crochet, donc comme groupe d'automorphismes de l'algèbre de Lie $\mathfrak{h}_n$, ainsi que de H_n.

VIII.5 – Théorème de Darboux

On verra au §IX.11 la définition des variétés symplectiques. Considérer un espace symplectique réel (E, ω_0) comme une variété symplectique revient à identifier à E l'espace $\mathrm{T}(\mathrm{M})$ des vecteurs tangents en chacun de ses points M et considérer de cette façon que ω_0 définit un produit scalaire symplectique sur $\mathrm{T}(\mathrm{M})$ en tout point M de E. Cela devient particulièrement simple si l'on considère le produit scalaire symplectique comme une forme différentielle extérieure sur E à coefficients constants (invariante par translations).

On a vu au §VIII.1 que si $(\mathrm{E}_1, [|]_1)$ et $(\mathrm{E}_2, [|]_2)$ sont deux espaces symplectiques de même dimension sur un même corps K, il existe toujours un isomorphisme linéaire Φ de E_1 sur E_2 tel que $[\Phi x \mid \Phi y] = [x \mid y]_1$, $\forall x, y \in \mathrm{E}_1$. Le théorème suivant généralise largement cette propriété pour $\mathrm{K} = \mathbb{R}$. On en verra la portée au §IX.11.

Théorème de Darboux VIII.5 : Soit, sur un ouvert U d'un espace vectoriel réel E de dimension $2n$, une forme différentielle extérieure à coefficients ω, *fermée* $(d\omega = 0)$ de rang maximum $2n$ en tout point de U. Soit ω_1 la forme différentielle extérieure à coefficients constants sur E égale à ω en un point x_1 de U. On peut alors trouver un difféomorphisme local Φ d'un voisinage ouvert $\mathrm{V} \subset \mathrm{U}$ de x_1 dans U, laissant x_1 fixe, transformant ω_1 en ω : $\Phi^* \omega_1 = \omega$. Si $x^1, x^2, \dots, x^n, y^1, y^2, \dots y^n$ sont des coordonnées sur E donnant à

ω_1 la forme standard $\omega_1 = \sum dx^1 \wedge dy^1$, les $(x^1 \circ \Phi, \dots, y^n \circ \Phi)$ forment un système de coordonnées curvilignes au voisinage de x_1 dans lequel la forme ω s'écrit $\omega = \sum d(x^j \circ \Phi) \wedge d(y^1 \circ \Phi)$

Preuve : On construit une famille à un paramètre Φ_t, $t \in [0,1]$, de difféomorphismes d'un voisinage V de x_1 dans U tels que $\Phi_t(x_1) = x_1$, $\forall t$, $\Phi_0 =$ identité et $\Phi_1 = \Phi$.

La forme différentielle extérieure

$$\omega_t = \omega + t(\omega_1 - \omega) = \sum_{i<j} \{a_{ij}(x) + t(a_{ij}(x_0) - a_{ij}(x))\} dx^i \wedge dx^j$$

est de rang maximum pour $0 \leqslant t \leqslant 1$ en $x = x_1$ donc dans un voisinage ouvert de x_1 que l'on choisir « étoilé ».

Il existe dans V une 1-forme α définie à la différentielle df d'une fonction numérique près, ce qui permet de prendre $\alpha(x_1) = 0$; telle que $\omega - \omega = d\alpha$. Mais ω_t étant de rang maximum il correspond à la forme α un champ de vecteur X_t nul en x_1 quel que soit t, tel que $i(\mathrm{X}_t)\omega_t = -\alpha$.

En réduisant éventuellement le voisinage ouvert V, le champ X_t définit un difféomorphisme local Φ_t pour $0 \leqslant t \leqslant 1$, tel que $\Phi_0 =$ Identité, $\Phi_0(x_1) = x_1$, $\forall t$, par : $\dfrac{\partial}{\partial t}\, \Phi_t(x) = \mathrm{X}_t(\Phi_t(x))$

Un calcul simple montre alors que :

$$\begin{aligned}
\frac{\partial}{\partial t}\,(\Phi_t^* \omega_t) &= \Phi_t^*\Big(\mathrm{L}_{\mathrm{X}}\omega_t + \frac{\partial}{\partial t}\,\omega_t\Big) \\
&= \Phi_t^*(di(\mathrm{X}_t)\omega_t + i(\mathrm{X}_t)d\omega_t + \omega_1 - \omega) \\
&= \Phi_t^*(-d\alpha + \omega_1 - \omega) \\
&= 0
\end{aligned}$$

Il en résulte que $\Phi_t^*\omega_t$ est indépendant de t, d'où :

$$\Phi_0^*\omega_0 = \Phi_0^*\omega = \omega = \Phi_1^*\omega_1 = \Phi^*\omega_1 \text{ c.q.f.d.}$$

CHAPITRE IX

Analyse tensorielle sur les variétés

C'est au départ, la relativité générale qui a imposé à l'analyse tensorielle de se concevoir dans des espaces généraux où les « liaisons rectilignes » sont absentes : l'espace est « courbe ».

Dans le même ordre d'idées, il s'est avéré nécessaire, dans l'élaboration de modèles des phénomènes physiques, de n'imposer à l'espace où ils se déroulent que le minimum de structures « *a priori* », afin qu'aucun préjugé, ou vice de forme, ne s'introduise dans leur analyse.

L'effort de construction de l'analyse tensorielle et de la géométrie dans le cadre des variétés différentielles a été largement compensé par une meilleure connaissance des structures et des algorithmes, et a permis de préciser en particulier le rôle exact de la métrique.

IX.1 / Définition des variétés différentielles, vecteurs et covecteurs tangents.
IX.2 / Champs de tenseurs. Formes différentielles extérieures. Volume.
IX.3 / Dérivations. Dérivations de Lie. Groupes et algèbres de Lie.
IX.4 / Dérivation covariante et transport parallèle. Connexions.
IX.5 / Connexions sur les fibrés vectoriels avec groupe structural.
IX.6 / Formes de courbure. Formes extérieures tensorielles.
IX.7 / Une application à la physique : les théories de jauge.
IX.8 / Tenseurs de courbure.
IX.9 / Géodésiques. Coordonnées normales. Déviations.
IX.10 / Intégration des formes extérieures. Formule de Stokes.
IX.11 / Calcul des variations sur une variété. Applications.

Vocabulaire :

- *différentiabilité :* voir le §II.2.
- de *classe* $\mathfrak{C}_r$ $(r \geqslant 1)$ signifie que l'on est assuré de l'existence et de la continuité des dérivées des fonctions et applications en jeu jusqu'à l'ordre r inclus.

– *difféomorphisme de classe* $\mathfrak{C}_r$: application bijective continûment différentiable jusqu'à l'ordre r ainsi que l'application inverse.

IX.1 – Définition des variétés différentielles. Vecteurs et covecteurs tangents.

Une variété différentielle, de classe $\mathfrak{C}_r$, $r \geqslant 1$, de dimension n, est, comme son nom l'indique, un espace sur lequel on a convenablement défini ce qu'est une fonction différentiable, de classe $\mathfrak{C}_r$ et qui, muni de cette « structure différentielle », est localement « difféomorphisme » à l'espace vectoriel réel $\mathbb{R}^n$.

Par analogie avec les représentations locales des sphères terrestre et céleste, ces difféomorphismes locaux, qui servent à la définition, sont appelés des *cartes* de la variété, et un ensemble de cartes, un *atlas*.

La définition ci-après va donc définir la « structure différentielle » à l'aide de cartes :

Définition IX.1.A :

Une variété différentielle $\mathfrak{V}$ de dimension n, de classe $\mathfrak{C}_r$, $r \geqslant 1$ est un *espace topologique séparé* muni de la structure définie par la donnée d'un *atlas* $\mathfrak{C}_r$ *sur* $\mathfrak{V}$, qui consiste en :

1) la donnée d'un recouvrement de $\mathfrak{V}$ par des ouverts U_j, $(j \in \mathrm{I})$;

2) pour chaque U_i, la donnée d'une application φ_j, bijective, continue ainsi que $\overset{-1}{\varphi}_j$, appelée *carte*, de U_j sur un ouvert de $\mathbb{R}^n$.

On impose à l'ensemble de ces cartes la condition suivante : (C) pour chaque couple $(i, j) \in \mathrm{I} \times \mathrm{I}$ tel que $\mathrm{U}_i \cap \mathrm{U}_j$ soit non vide, l'application bijective $\varphi_j \circ \overset{-1}{\varphi}_i$ de l'ouvert $\varphi_i(\mathrm{U}_i \cap \mathrm{U}_j)$ sur l'ouvert $\varphi_j(\mathrm{U}_i \cap \mathrm{U}_j)$, ou *changement de cartes* est un difféomorphisme de classe $\mathfrak{C}_r$.

Une fonction f sur $\mathfrak{V}$, à valeurs dans un espace vectoriel réel fixe E, est alors dite différentiable de classe $\mathfrak{C}_s$, $s \leqslant r$, en un point x de $\mathfrak{V}$ si pour un ouvert U_j contenant x, la fonction $f \circ \overset{-1}{\varphi}_j$ est s fois continûment différentiable au point $\varphi_j(x)$ de $\mathbb{R}^n$, et cette propriété ne dépend pas, vu la condition (C) de la carte φ_j choisie. f est dite de classe $\mathfrak{C}_s$ sur $\mathfrak{V}$ si elle l'est en tout point.

Si $(\mathrm{U}'_k\,;\varphi'_k\,;k\in \mathrm{J})$ est un atlas $\mathfrak{C}_r$ sur $\mathfrak{V}$, il est dit équivalent au premier s'il est compatible avec lui c'est-à-dire si la réunion $(\mathrm{U}_j,\mathrm{U}'_k\,;\varphi_j,\varphi'_k\,;j\in \mathrm{I},\,k\in \mathrm{J})$ est encore un atlas $\mathfrak{C}_r$ de $\mathfrak{V}$. Il donne alors la même définition des fonctions de classes $\mathfrak{C}_r$ et l'on convient de dire qu'il définit sur $\mathfrak{V}$ la même structure de variété différentielle.

Si l'on désigne par ζ^1, ζ^2, ..., ζ^n les fonctions coordonnées standard dans $\mathbb{R}^n$, les n fonctions composées $x^1=\zeta^1\circ\varphi_j,\ \dots x^n=\zeta^n\circ\varphi_j$ sont des fonctions à valeurs réelles, de classe $\mathfrak{C}_r$ sur l'ouvert U_j appelées coordonnées locales sur U_j définies par la carte φ_j.

Nous allons maintenant montrer, dans ce qui suit, comment un atlas $\mathfrak{C}_r$ permet de transporter à l'espace $\mathfrak{V}$ *toute les structures des ouverts de* $\mathbb{R}^n$ *qui sont conservées par les difféomorphismes de classe* $\mathfrak{C}_r$.

A toute fonction f de $\mathfrak{V}$ à valeurs dans un ensemble ou un espace quelconque E correspond localement, dans un ouvert U de $\mathfrak{V}$ possédant une carte φ, une fonction F des coordonnées locales $x^1,x^2,\dots x^n$ d'un point x de U définies par φ telle que : $f(x)=\mathrm{F}(x^1,x^2,\dots x^n)$. En effet, si $x\in \mathrm{U}$, $\mathrm{F}=f\circ\overset{-1}{\varphi}$ est une fonction sur l'ouvert $\varphi(\mathrm{U})$ de $\mathbb{R}^n$ c'est-à-dire une fonction de n variables réelles : $\zeta^1,\zeta^2,\dots\zeta^n$ à valeurs dans E, et

$$\begin{aligned}f(x)&=(f\circ\overset{-1}{\varphi})(\varphi(x))=\mathrm{F}(\varphi(x))=\mathrm{F}[\zeta^1(\varphi(x)),\dots\zeta^n(\varphi(x))]\\&=\mathrm{F}(x^1,\dots x^n).\end{aligned}$$

D'après la définition ci-dessus, dans le cas où E est un espace vectoriel réel, dire que f est différentiable de classe $\mathfrak{C}_s$ en point revient à dire que la fonction F de n variables réelles correspondante est de classe $\mathfrak{C}_s$ au point $\zeta=\varphi(x)$.

Remarquons que nous avons clairement défini la notion de différentiabilité de toutes classes sur les variétés mais que nous ne somme pas encore capables de dire ce qu'est la différentielle d'une fonction ! Le seul cas à notre portée est celui de la différentielle nulle.

En effet, si f est une fonction sur $\mathfrak{V}$ à valeurs dans un espace vectoriel réel E, différentiable en un point x, elle est dite *plate* en x, ou de différentielle nulle, si pour une carte φ en x, la fonction

$f \circ \overset{-1}{\varphi}$ a une différentielle $\mathrm{D}(f \circ \overset{-1}{\varphi})_{\varphi(x)}$ nulle au point $\varphi(x)$ de $\mathbb{R}^n$. Cette définition ne dépend évidemment pas de la carte choisie car, si ψ est une autre carte, on a, en vertu de la transitivité des différentielles :

$$\mathrm{D}(f \circ \overset{-1}{\psi})_{\psi(x)} = \mathrm{D}(f \circ \overset{-1}{\varphi})_{\varphi(x)} \circ \mathrm{D}(\varphi \circ \overset{-1}{\psi})_{\psi(x)}.$$

Dans le même ordre d'idées, si l'on peut immédiatement définir ce qu'est un chemin différentiable de classe $\mathfrak{C}_s$ d'une variété $\mathfrak{V}$ de classe $\mathfrak{C}_r$, $r \geqslant s$: application continue c d'un intervalle $[a, b]$ de $\mathbb{R}$ dans $\mathfrak{V}$ telle que pour toute carte φ, $\varphi(c(t))$ soit de classe $\mathfrak{C}_s$, rien, ne permet jusqu'à présent de donner un sens à la notion de vecteur-vitesse ou dérivée $\dfrac{dc}{dt}$ en un point. La définition suivante, qui reprend exactement l'interprétation des vecteurs issus d'un point de $\mathbb{R}^n$ comme opérateurs de dérivation (Définition II.2.), va résoudre ces problèmes.

Définition IX.1.B

Un *vecteur tangent* X en un point x d'une variété différentielle $\mathfrak{V}$ est un opérateur de dérivation des fonctions différentiables à valeurs réelles définies au voisinage de x. De façon précise, à une telle fonction f, X fait correspondre un nombre, noté $\mathrm{X} \cdot f$, et appelé *dérivée de f par rapport à* X, cette opération devant satisfaire aux deux propriétés qui lui servent d'axiomes :

1) $\mathrm{X} \cdot f = 0$ si f est plate en x, d'où il résulte que X est nul sur les constantes, et que $\mathrm{X} \cdot f = \mathrm{X} \cdot g$ lorsque f et g sont égales dans un voisinage arbitrairement petit de x.

2) linéarité : $\mathrm{X}(\lambda f + \mu g) = \lambda \mathrm{X} \cdot f + \mu \mathrm{X} \cdot g$, $\forall \lambda, \mu \in \mathbb{R}$

L'ensemble des vecteurs tangents en un point x de $\mathfrak{V}$ est un espace vectoriel appelé *l'espace tangent en x* et est noté $\mathrm{T}(\mathfrak{V})_x$, ou simplement T_x ou $\mathfrak{V}_x$.

Remarque : Si f et g sont deux fonctions numériques différentiables en x_0 et nulles en x_0, leur produit fg est différentiable et plat en x_0. En effet, si φ est une carte, $fg \circ \overset{-1}{\varphi} = (f \circ \overset{-1}{\varphi}) \cdot (g \circ \overset{-1}{\varphi})$ et en différentiant ces fonctions numériques définies sur un ouvert de $\mathbb{R}^n$:

$$d(fg \circ \overset{-1}{\varphi})_{\varphi(x_0)} = f(x_0) \cdot d(g \circ \overset{-1}{\varphi})_{\varphi(x_0)} + g(x_0) d(f \circ \overset{-1}{\varphi})_{\varphi(x_0)} = 0.$$

Si X est un vecteur tangent en x_0 et si f et g sont maintenant deux fonctions numériques différentiables en x_0 quelconques :

$$\begin{aligned} \mathrm{X}\cdot(fg) &= \mathrm{X}\cdot\{(f-f(x_0)+f(x_0))\cdot(g-g(x_0)+g(x_0))\} \\ &= f(x_0)\mathrm{X}\cdot g + g(x_0)\mathrm{X}\cdot f. \end{aligned}$$

Les axiomes définissant un vecteur tangent entraînent donc le comportement caractéristique des dérivations vis-à-vis du produit.

Les raisonnements suivants reprennent ceux du §II.2.

Soient φ une carte d'un voisinage U de $x_0 \in \mathfrak{V}$, $x^1, x^2, \dots x^n$ les coordonnées locales correspondantes, F la fonction différentiable de n variables correspondant à une fonction différentiable f à valeurs réelles définie sur U : $f(x) = \mathrm{F}(x^1, x^2, \dots x^n)$. Par définition, f est plate en x_0 si la différentielle $(\mathrm{DF})_{x_0} = \sum \left(\frac{\partial \mathrm{F}}{\partial x^j}\right)_{x_0} dx^j$ est nulle, soit $\left(\frac{\partial \mathrm{F}}{\partial x^j}\right)_{x_0} = 0\ \forall j$. Pour chaque $j = 1, 2, \dots, n$ on définit un vecteur tangent $\left(\frac{\partial}{\partial x^j}\right)_{x_0}$ par $\left(\frac{\partial}{\partial x^j}\right)_{x_0} \cdot f = \left(\frac{\partial \mathrm{F}}{\partial x^j}\right)_{x_0}$, qu'on peut noter simplement $\left(\frac{\partial f}{\partial x^j}\right)_{x_0}$. Puisque la fonction :

$f(x) - \sum \left(\frac{\partial \mathrm{F}}{\partial x^j}\right)_{x_0} (x^j - x_0^j) = \mathrm{F}(x^1, x^2, \dots x^n) - \sum \left(\frac{\partial \mathrm{F}}{\partial x^j}\right)_{x_0} (x^j - x_0^j)$ est plate en x_0, on a, si X est un vecteur tangent quelconque en x_0 :

$$\mathrm{X}\cdot f = \sum (\mathrm{X}\cdot x^j)\left(\frac{\partial}{\partial x^j}\right)_{x_0} \cdot f$$

ce qui prouve que les n dérivations par rapport à des coordonnées locales sur un ouvert U de $\mathfrak{V}$ forment en chaque point de U une base de l'espace tangent en ce point. Les espaces tangents de $\mathfrak{V}$ ont donc pour dimension celle de $\mathfrak{V}$.

Exemple : Vecteur-vitesse en un point $x_0 = c(t_0)$ d'un chemin différentiable $t \to c(t)$ de $\mathfrak{V}$. Si f est une fonction numérique différentiable en x_0, on définit le vecteur tangent $\left(\frac{dc}{dt}\right)_{x_0}$ comme l'opérateur de dérivation :

$$\left(\frac{dc}{dt}\right)_{x_0} \cdot f = \left(\frac{df(c(t))}{dt}\right)_{t_0} = \lim_{h\to 0} \frac{f(c(t_0+h)) - f(c(t_0))}{h}.$$

Si φ est une carte d'un voisinage U de $x_0 = c(t_0)$ cela s'écrit :

$$= \lim_{h\to 0} \frac{(f \circ \overset{-1}{\varphi})(\varphi(c(t_0+h))) - (f \circ \overset{-1}{\varphi})(\varphi(c(t_0)))}{h}$$

$$= \mathrm{D}(f \circ \overset{-1}{\varphi})_{\varphi(x_0)}\Big(\frac{d\varphi(c(t))}{dt} t_0\Big) \text{ (calculé dans } \mathbb{R}^n)$$

d'où résulte bien que $\left(\dfrac{dc}{dt}\right)_{x_0} \cdot f$ est nul lorsque f est plate en x_0.

On a ainsi attaché à chaque point d'une variété différentielle de dimension n un espace vectoriel de même dimension, l'espace tangent qui bien que de définition abstraite, va jouer le même rôle que les tangentes aux courbes, plans tangents aux surfaces, plongées dans $\mathbb{R}^n$, et qui sont, eux, définis géométriquement par l'espace vectoriel ambiant.

Le dual de l'espace tangent T_x est *l'espace cotangent* T_x^*. C'est donc, par définition, l'espace vectoriel des formes linéaires sur T_x. Toute fonction numérique différentiable en x définit un élément de T_x^*, appelé sa différentielle df_x : c'est la forme linéaire sur T_x qui à chaque vecteur tangent $\mathrm{X} \in \mathrm{T}_x$ fait correspondre la dérivée de f par rapport à X : $\mathrm{X} \cdot f$, qui peut donc aussi s'écrire $df(\mathrm{X})$. L'espace cotangent T_x^* peut donc être considéré comme l'espace vectoriel des différentielles des fonctions numériques en x. Les différentielles des fonctions coordonnées $(x^1, x^2, \dots x^n)$ soient : $(dx^1, dx^2, \dots dx^n)$ forment une base de T_x^* duale de la base $\left(\dfrac{\partial}{\partial x^1}, \dfrac{\partial}{\partial x^2}, \dots \dfrac{\partial}{\partial x^n}\right)$.

La réunion des espaces tangents T_x aux points de la variété $\mathfrak{V}$ est l'espace tangent $\mathrm{T}(\mathfrak{V})$, la réunion des T_x^*, l'espace cotangent $\mathrm{T}^*(\mathfrak{V})$. Il ne faut naturellement pas considérer $\mathrm{T}(\mathfrak{V})$ ou $\mathrm{T}^*(\mathfrak{V})$ comme des réunions d'espaces vectoriels sans lien entre eux. Ce sont, de façon naturelle des variétés de classe $\mathfrak{C}_{r-1}$. Si U est un ouvert de $\mathfrak{V}$ possédant une carte φ dans $\mathbb{R}^n$, définissant des coordonnées locales $x^1, x^2, \dots x^n$, la carte φ se prolonge en une carte Φ de T(U) dans $\mathbb{R}^{2n}$ qui au point (x, X) de T(U) avec $\mathrm{X} \in \mathrm{T}_x$, fait correspondre les $2n$ coordonnées $(x^1, x^2, \dots x^n, \mathrm{X}\cdot x^1, \dots \mathrm{X}\cdot x^n)$ et au point $(x; \alpha)$ de $\mathrm{T}^*(\mathrm{U})$, avec $\alpha \in \mathrm{T}_x^*$ les $2n$ coordonnées $\left(x^1, x^2, \dots x^n \Big(\dfrac{\partial}{\partial x^1}\Big), \dots \Big(\dfrac{\partial}{\partial x^n}\Big)\right)$.

Plus généralement, à tout symbole de variance (v) (cf. chapitre III) correspond en chaque point x de $\mathfrak{V}$, un espace tensoriel

$\otimes^{(v)}\mathrm{T}_x$. La réunion de ces espaces est *l'espace tensoriel tangent* $\otimes^{(v)}\mathrm{T}(\mathfrak{V})$ à la variété $\mathfrak{V}$. Il possède une structure de variété de classe $\mathfrak{C}_{r-1}$ définie par les prolongements des cartes de $\mathfrak{V}$, à valeurs dans $\mathbb{R}^n \times (\otimes^{(v)}\mathbb{R}^n)$. On définit de la même façon les espaces de pseudotenseurs tangents de variance tensorielle (v), de variance scalaire m, les cartes de $\mathfrak{V}$ dans $\mathbb{R}^n$ se prolongent en des cartes dans les espaces de pseudotenseurs de même variance (v), m, sur $\mathbb{R}^n$.

On peut maintenant étendre aux variétés différentielles les notions d'applications différentiables d'applications tangente et cotangente, et de différentielle.

Une application Φ d'un ouvert U d'une variété $\mathfrak{V}$ dans une variété $\mathfrak{V}'$ est dite différentiable de classe $\mathfrak{C}_s$ en un point x, si pour des cartes φ de $\mathfrak{V}$ au voisinage de x et φ' de $\mathfrak{V}'$ au voisinage de $\Phi(x)$, $\varphi' \circ \Phi \circ \varphi^{-1}$ est de classe $\mathfrak{C}_s$ en $\varphi(x)$. L'*application tangente* Φ_* de T(U) dans $\mathrm{T}(\mathfrak{V}')$ est définie par $\Phi_*(\mathrm{X}) \cdot f = \mathrm{X}(f \circ \Phi)$. L'application duale, ou *application cotangente* $\Phi^* =^t \Phi_*$ applique le covecteur $\alpha \in \mathrm{T}^*(\mathfrak{V}')_{\Phi\mathrm{X}}$ sur le covecteur $\Phi^*(\alpha) = \alpha \circ \Phi_*$.

Si α est la différentielle d'une fonction numérique sur $\mathfrak{V}$, $\Phi^*(df) = df \circ \Phi_* = d(f \circ \Phi)$. Si Ψ est une application différentiable de $\mathfrak{V}'$ dans $\mathfrak{V}''$, on a $(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*$ et $(\Psi \circ \Phi)^* = \Phi^* \circ \Psi^*$.

Le terme de *différentielle* s'applique spécialement aux fonctions sur $\mathfrak{V}$ à valeurs dans un espace vectoriel *réel fixe* L.

La différentielle df_x d'une telle fonction f en un point x de $\mathfrak{V}$ est alors l'application tangente $(f_*)_x$ considérée comme application linéaire de T_x dans L, soit $df_x \in \mathfrak{L}(\mathrm{T}_x; \mathrm{L})$. Si X est un vecteur tangent en $x \in \mathfrak{V}$ la dérivée de f par rapport à X est alors $\mathrm{X} \cdot f = df(\mathrm{X}) = (f_*)_x(\mathrm{X})$. Les vecteurs tangents, initialement définis comme dérivations des fonctions différentiables numériques, sont donc plus généralement des dérivations des fonctions différentiables à valeurs dans un espace vectoriel réel *fixe* L.

Exemple : si φ est une carte d'un ouvert U de $\mathfrak{V}$ dans $\mathbb{R}^n$, $\zeta^2, \zeta^2, \ldots \zeta^n$ les coordonnées de $\mathbb{R}^n$ et $x^j = \zeta^j \circ \varphi$ on a : $dx^j = d\zeta^1 \circ \varphi_*$ soit aussi :

$$dx^j = \varphi^*(d\zeta^j).$$

La structure de ces espaces : tangents, cotangents, tensoriels, sur une variété $\mathfrak{V}$ est décrite par le terme *d'espace fibré vectoriel* pour la raison suivante : un espace de tenseurs, $\otimes^{(\mathfrak{v})}\mathrm{T}(\mathfrak{V})$, sur $\mathfrak{V}$, par exemple est muni de l'application naturelle π qui à chaque élément t de $\otimes^{(\mathfrak{v})}\mathrm{T}(\mathfrak{V})_x$ associe le point x de $\mathfrak{V}$ au-dessus duquel il est défini : π est de classe $\mathfrak{C}_{r-1}$ et l'image réciproque $\overset{-1}{\pi}\, x$, ou fibre de x est un espace vectoriel, $\otimes^{(\mathfrak{v})}\mathrm{T}_x$ dans notre cas.

De plus, chaque carte de $\otimes^{(\mathfrak{v})}\mathrm{T}(\mathfrak{V})$ associée à une carte φ d'un ouvert U de $\mathfrak{V}$ applique, par une bijection de classe $\mathfrak{C}_{r-1}$, $\otimes^{(\mathfrak{v})}\mathrm{T}(\mathrm{U})$ sur le produit $\varphi(\mathrm{U}) \times (\otimes^{(\mathfrak{v})}\mathbb{R}^n)$, dont la restriction à chaque fibre $\otimes^{(\mathfrak{v})}\mathrm{T}_x$ est un isomorphisme sur l'espace tensoriel $\varphi(x) \times (\otimes^{(\mathfrak{v})}\mathbb{R}^n)$.

On est donc conduit à la définition générale suivante :

Définition IX.1.C :

Un *espace fibré vectoriel* E de classe $\mathfrak{C}_r$, de dimension p, au-dessus d'une variété différentielle $\mathfrak{V}$, est une variété différentielle, munie d'une application Π sur $\mathfrak{V}$, dite projection, et de la structure suivante :

1) une structure d'espace vectoriel réel sur chaque fibre $\mathrm{E}_x = \overset{-1}{\prod}\, x, \forall x \in \mathfrak{V}$.

2) un atlas de cartes déterminé par la donnée d'un recouvrement de $\mathfrak{V}$ par des ouverts $\mathrm{U}_\alpha, \alpha \in \mathrm{I}$ et pour chaque U_α, la donnée d'un difféomorphisme Ψ_α de classe $\mathfrak{C}_r$ de $\overset{-1}{\prod}\, \mathrm{U}_\alpha$ sur $\mathrm{U}_\alpha \times \mathbb{R}^p$ appliquant le fibre $\mathrm{E}_x = \overset{-1}{\prod}\, x$ sur $x \times \mathbb{R}^p$ par un isomorphisme d'espaces vectoriels. $\mathfrak{V}$ est appelée la *base* de E.

Une application u de $\mathfrak{V}$ dans E telle que pour chaque x de $\mathfrak{V}$, $u(x)$ appartienne à la fibre E_x au-dessus de x est appelée une *section de* E, ou encore un *champ de vecteurs de* E, *sur* $\mathfrak{V}$. L'ensemble des sections de E sur $\mathfrak{V}$, de classe $\mathfrak{C}_r$ forme un espace vectoriel que l'on note traditionnellement $\Gamma\mathrm{E}$. Si E et F sont deux espaces fibrés vectoriels $\mathfrak{C}_r$ au-dessus de $\mathfrak{V}$, une application Φ de E dans F, de classe $\mathfrak{C}_r$, telle que, pour chaque x de $\mathfrak{V}$, la restriction de Φ à la fibre E_x soit une application linéaire Φ_x de E_x dans F_x est appelée un *morphisme* de E dans F. Toutes les opérations de l'algèbre linéaire et multilinéaire peuvent être transportées aux espaces fibrés vectoriels.

On a vu au chapitre IV l'importance des tenseurs restreints et de leur groupe structural. On est ainsi à considérer la structure plus riche suivante :

Définition IX.1.D :

Un espace fibré vectoriel E de classe $\mathfrak{C}_r$ de dimension p *de groupe structural* G, sous-groupe du groupe linéaire $\mathrm{Gl}(p)$, est un espace fibré vectoriel au sens de la définition précédente, dont les changements de cartes s'effectuent dans le groupe G.

De façon précise, si $\mathrm{U}_\alpha \cap \mathrm{U}_\beta \neq \varnothing$, on a $\Psi_\alpha \overline{\Psi}_\beta^{1} = g_{\alpha\beta}$ où :
$x \in \mathrm{U}_\alpha \cap \mathrm{U}_\beta \to g_{\alpha\beta}(x) \in \mathrm{G}$ est une fonction de classe $\mathfrak{C}_r$.

Si $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots \varepsilon_p)$ désigne la base standard de $\mathbb{R}^p$, $\varepsilon_\alpha = \overline{\Psi}_\alpha^{1} \varepsilon$ et $\varepsilon_\beta = \overline{\Psi}_\beta^{-1} \varepsilon$ sont des champs de repères de E au-dessus de U_α et U_β respectivement. On a $\Psi_\alpha \overline{\Psi}_\beta^{1} \varepsilon = g_{\alpha\beta}\varepsilon$ que l'on peut écrire matriciellement : $g_{\alpha\beta}\varepsilon = \varepsilon \cdot \mathrm{G}_{\alpha\beta}$ ou $\mathrm{G}_{\alpha\beta}$ est la matrice $p \times p$ de $g_{\alpha\beta}$ dans la base ε. On a donc :

$$\varepsilon_\beta = \overset{-1}{\Psi}{}_\beta \varepsilon = \overset{-1}{\Psi}{}_\alpha \varepsilon \mathrm{G}_{\alpha\beta} = \varepsilon_\alpha \mathrm{G}_{\alpha\beta}.$$

Exemple : L'espace tangent d'une variété riemannienne $\mathfrak{V}$ de dimension n (voir ci-dessous Définition IX.2.C) ainsi que les espaces de tenseurs restreints sur $\mathfrak{V}$ sont des espaces fibrés vectoriels de groupe structural $\mathrm{O}(n)(\mathrm{SO}(n))$ si $\mathfrak{V}$ est orientée).

IX.2 – Champs de tenseurs. Formes différentielles extérieures. Volume

Définition IX.2.A :

Soit A une sous-variété de la variété différentielle $\mathfrak{V}$ (A peut être un morceau de courbe, ou de surface, ou un ouvert de $\mathfrak{V}$, etc...)

Un *champ de tenseurs* t de variance (v), de classe $\mathfrak{C}_s$ $(s \leqslant r)$ sur A est une fonction qui à chaque x de A fait correspondre un élément t_x de la fibre $\otimes^{(\mathfrak{v})}\mathrm{T}_x$ au-dessus de x, et qui, en tant qu'application de A dans la variété $\otimes^{(\mathfrak{v})}\mathrm{T}(\mathfrak{V})$ est de classe $\mathfrak{C}_s$, autrement dit une section de classe $\mathfrak{C}_s$ du fibré vectoriel $\otimes^{(\mathfrak{v})}\mathrm{T}(\mathfrak{V})$. Cette dernière condition peut s'exprimer simplement de la façon suivante :

Localement, un champ de tenseurs p fois contravariants et q fois covariants possède, relativement à des coordonnées locales

$(x^1, x^2, \ldots x^n)$ des composantes $t^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$ dans la base $\dfrac{\partial}{\partial x^{i_1}} \otimes \ldots \dfrac{\partial}{\partial x^{i_p}} \otimes dx^{j_1} \otimes \ldots dx^{j_q}$. t est de classe $\mathfrak{C}_s$ si ces composantes sont des fonctions numériques de classe $\mathfrak{C}_s$.

Par un changement de coordonnées $(x^1, \ldots x^n) \to (x'^1, \ldots x'^n)$, elles se transforment suivant les formules classiques :

$$\boxed{t'^{i_1 \ldots i_p}_{j_1 \ldots j_q} = \sum t^{k \ldots k_p}_{m_1 \ldots m_q} \frac{\partial x'^{i_1}}{\partial x^{k_1}} \cdots \frac{\partial x'^{i_p}}{\partial x^{k_p}} \cdot \frac{\partial x^{m_1}}{\partial x'^{j_1}} \cdots \frac{\partial x^{m_q}}{\partial x'^{j_q}}}.$$

Si t est un champ de tenseurs et f une fonction numérique tous deux de classe $\mathfrak{C}_s$ sur A, le produit ft est le champ tenseurs de classe $\mathfrak{C}_s$ défini en chaque point x par $(ft)(x) = f(x)t_x$.

Les opérations tensorielles : multiplications tensorielles, contractions, permutations d'arguments se transportent immédiatement aux champs de tenseurs sur A en s'effectuant en chaque point.

La contraction d'un champ de vecteurs X et d'un champ de covecteurs α par exemple est la fonction numérique sur A : $x \to \langle \mathrm{X}_x, \alpha_x \rangle = \left(\sum \mathrm{X}^j \alpha_j \right)_x$.

Ce qui caractérise les tenseurs et les opérations tensorielles sur un espace vectoriel fixe E est la multilinéarité. On peut se demander quelle forme prend cette propriété lorsque l'on considère des champs de tenseurs sur une variété $\mathfrak{V}$. Elle s'exprime alors de la façon suivante : si $\mathfrak{F}$ est l'algèbre des fonctions numériques de classe $\mathfrak{C}_s$ sur $\mathfrak{V}$, *les champs de tenseurs d'une variance* (v) *donnée, de même classe* $\mathfrak{C}_s$ *forment un module sur* $\mathfrak{F}$*, et les opérations tensorielles entre champs de tenseurs de classe* $\mathfrak{C}_s$ *sont* $\mathfrak{F}$*-multilinéaires.*

Les ensembles des champs de tenseur de diverses variances, ou simplement des champs de tenseurs contravariants, ou de champs de tenseurs covariants, d'une même classe $\mathfrak{C}_s$ sur $\mathfrak{V}$ sont, avec la multiplication tensorielle, des $\mathfrak{F}$-algèbres.

Nous allons maintenant définir le *crochet* de deux champs de vecteurs tangents d'une variété différentielle.

Soient Φ une application différentiable de $\mathfrak{V}$ dans $\mathfrak{W}$, Φ_* l'application tangente de $\mathrm{T}(\mathfrak{V})$ dans $\mathrm{T}(\mathfrak{W})$. Si X est un champ de vecteurs tangents sur $\mathfrak{V}$, $\Phi_*(\mathrm{X})$ *n'est pas* en général un champ de vecteurs tangents sur $\mathfrak{W}$, et ceci pour deux raisons :

1) $\Phi_*(\mathrm{X})$ en un point y de $\mathfrak{W}$ peut avoir autant de valeurs différentes qu'il y a de points x de $\mathfrak{V}$ tels que $\Phi(x) = y$ puisque $(\Phi_*)_x$ applique $\mathrm{T}(\mathfrak{V})_x$ dans $\mathrm{T}(\mathfrak{W})_{y=\Phi(x)}$.

2) l'image de Φ n'est pas en général la variété $\mathfrak{W}$ tout entière.

Dans le meilleur des cas, un champ X sur $\mathfrak{V}$ et un champ X' sur $\mathfrak{W}$ sont dits *compatibles pour l'application* Φ si quel que soit $x \in \mathfrak{V}$, $\Phi_*(\mathrm{X}_x) = \mathrm{X}'_{\Phi(x)}$. Dans ce cas quels que soient la fonction numérique différentiable h sur $\mathfrak{W}$ et le point x de $\mathfrak{V}$, on a : $(\mathrm{X}' \cdot h)_{\Phi(x)} = \Phi_*(\mathrm{X}_x) \cdot h = \mathrm{X}_x \cdot (h \circ \Phi)$ ce qui exprime *l'identité entre fonctions sur* $\mathfrak{V}$: $(\mathrm{X}' \cdot h) \circ \Phi = \mathrm{X} \cdot (h \circ \Phi)$

C'est évidemment toujours le cas lorsque Φ est un difféomorphisme et $\mathrm{X}' = \Phi_*(\mathrm{X})$.

Soient maintenant X et Y deux champs de vecteurs sur $\mathfrak{V}$, X' et Y' deux champs de vecteurs sur $\mathfrak{W}$, Φ une application de $\mathfrak{V}$ dans $\mathfrak{W}$ telle que X et X' soient compatibles et Y et Y' soient compatibles pour Φ.

Si tout est de classe $\mathfrak{C}_s$, $s \geqslant 2$, on a alors :

$$(\mathrm{Y}' \cdot (\mathrm{X}' \cdot h)) \circ \Phi = \mathrm{Y} \cdot ((\mathrm{X}' \cdot h) \circ \Phi) = \mathrm{Y} \cdot (\mathrm{X} \cdot (h \circ \Phi))$$

d'où l'égalité :

$$\{\mathrm{X}' \cdot (\mathrm{Y}' \cdot h) - \mathrm{Y}' \cdot (\mathrm{X}' \cdot h)\} \circ \Phi = \{\mathrm{X}(\mathrm{Y} \cdot (h \circ \Phi)) - \mathrm{Y} \cdot (\mathrm{X} \cdot (h \circ \Phi))\}.$$

Définition IX.2.B :

Si X et Y sont deux champs de vecteurs tangents de classe $\mathfrak{C}_s$, $s \geqslant 2$, sur la variété différentielle $\mathfrak{V}$, leur crochet, noté $[\mathrm{X}; \mathrm{Y}]$ est un champ de vecteurs tangents de classe $\mathfrak{C}_{s-1}$. Il est défini par l'application qui à chaque fonction numérique de classe $\mathfrak{C}_1$ en un point x fait correspondre le nombre : $([\mathrm{X}, \mathrm{Y}] \cdot f)_x = \mathrm{X}_x \cdot (\mathrm{Y} \cdot f) - \mathrm{Y}_x(\mathrm{X} \cdot f)$.

Cette définition demande un certain nombre d'explications et de justifications. Tout d'abord, cette définition permet d'écrire l'égalité qui vient d'être démontrée *entre champs compatibles* :

$$\boxed{([\mathrm{X}', \mathrm{Y}'] \cdot h) \circ \Phi = [\mathrm{X}, \mathrm{Y}] \cdot (h \circ \Phi)}.$$

On va l'appliquer en prenant pour Φ une carte φ d'un ouvert U de $\mathfrak{V}$ sur un ouvert $\mathrm{U}' = \varphi(\mathrm{U})$ de $\mathbb{R}^n$. Les champs de vecteurs compatibles avec X et Y sont alors leurs images : $X' = \varphi_*(X)$ et $Y' = \varphi_*(Y)$. Ce sont des champs de vecteurs sur U′ et on a vu au §IV.14 que leur crochet peut s'écrire $[X', Y'] = DY'(X) - DX'(Y)$ et est ainsi défini pour X, Y de classe $\mathfrak{C}_1$ seulement.

Pour définir l'action de ce crochet sur une fonction numérique différentiable h, il suffit aussi que h soit de classe $\mathfrak{C}_1$, puisque cette action peut s'écrire : $[X', Y'] \cdot h = dh(DY'(X') - DX'(Y'))$.

Il résulte de cela :

a) que le crochet $[X, Y]$ sur $\mathfrak{V}$, opère en chaque point x sur les fonctions numériques f de classe $\mathfrak{C}_1$ en ce point, puisqu'une telle fonction f s'écrit : $f = (f \circ \overset{-1}{\varphi}) \circ \varphi$,

b) que $[X, Y]_x \cdot f$ est nul si f est plate. $[X, Y]$ est donc bien un champ de vecteurs tangents sur V.

Une autre conséquence de l'égalité $([X', Y'] \circ \Phi = [X, Y] \circ (h \circ \Phi)$ est la proposition suivante :

Proposition IX.2 :

Soit Φ une application différentiable d'une variété $\mathfrak{V}$ dans une variété $\mathfrak{W}$. Si (X, X') et (Y, Y') sont deux couples de champs de vecteurs compatibles pour Φ : X, Y sur $\mathfrak{V}$ et X', Y' sur $\mathfrak{W}$, de classe $\mathfrak{C}_s$, $s \geqslant 2$, leurs crochets : $[X, Y]$ sur $\mathfrak{V}$ et $[X', Y']$ sur $\mathfrak{V}'$ sont également compatibles pour Φ.

Preuve : L'égalité ci-dessus exprime en effet qu'en chaque point x de $\Phi_*([X, Y]_x) = [X', Y']_{\Phi(x)}$.

Des exemples de champs de tenseurs sont donnés par la définition suivante :

Définition IX.2.C :

Sur une variété différentielle $\mathfrak{V}$ de classe $\mathfrak{C}_r$, $r \geqslant 3$, une *structure riemannienne*, resp. *pseudoriemannienne*, resp. *presque symplectique* est la donnée d'un champ de tenseurs g deux fois covariants, de classe $\mathfrak{C}_{r-1}$, tel que pour chaque x de $\mathfrak{V}$, $g_x \in T_x^* \otimes T_x^*$ soit, sur l'espace tangent, T_x, un produit scalaire euclidien, resp. pseudoeuclidien resp. symplectique.

Sur un voisinage possédant une carte, g s'écrit avec les coordonnées locales : $g = \sum g_{ij} dx^i \otimes dx^j$. Lorsqu'il s'agit d'une structure riemannienne ou pseudoriemannienne, g est symétrique non dégénéré et on l'écrit généralement de façon concise sous la forme

$\sum g_{ij}dx^i dx^j$. Or, on verra ci-dessous qu'une forme différentielle extérieure de degré deux sur $\mathfrak{V}$ s'écrit de façon abrégé $\sum\limits_{i<j} a_{ij}dx^i dx^j$ mais cette fois $dx^i dx^j$ signifie $dx^i \wedge dx^j$!

Ces deux simplifications d'écriture sont *parfaitement incompatibles*, mais néanmoins en usage permanent. Seul le contexte permet de leur attribuer leur sens.

Sur une variété riemannienne, pseudoriemannienne ou presque symplectique, la métrique g détermine des isomorphismes canonique entre champs de tenseurs de même ordre mais de variances distinctes. On a en effet un isomorphisme métrique ρ du fibré tangent $\mathrm{T}(\mathfrak{V})$ sur le fibré cotangent $\mathrm{T}^*(\mathfrak{V})$ qui fait correspondre au vecteur X_x la forme $\zeta_x = \rho\mathrm{X}_x = g(.,\mathrm{X}_x) = (. \mid \mathrm{X}_x)$, au champ de vecteurs X la forme $\zeta = \rho\mathrm{X}$ telle que $(\mathrm{Y} \mid \mathrm{X}) = \langle \mathrm{Y}, \zeta \rangle$ tandis que $\overset{-1}{\rho}$ change la forme α en le champ de vecteurs A tel que $(\mathrm{Y} \mid \mathrm{A}) = \langle \mathrm{Y}, \alpha \rangle$. En termes de composantes dans des coordonnées locales, ρ se traduit par la contraction par le tenseur métrique g, et $\overset{-1}{\rho}$ par la contraction par le tenseur inverse $\overset{-1}{g}$:

$$\zeta_i = \sum g_{ij}\mathrm{X}^j;\ \mathrm{X}^i = \sum g^{ij}\zeta_j;\ \mathrm{A}^i = \sum g^{ij}\alpha_j.$$

Les puissances tensorielles, extérieures ou symétriques de ρ ou du tenseur g assurent le passage d'une variance à l'autre. On a donc sur une variété métrique des champs restreints dont on peut à volonté choisir la variance.

On a déjà vu au §. IV qu'il y avait deux façons d'exprimer une forme différentielle extérieure sur un espace vectoriel réel de E. Cette propriété est générale. Elle est partagée par tous les champs de tenseurs qui peuvent ainsi être considérés, soit comme des sections d'un espace fibré, soit comme des fonctions de champs de vecteurs ou de covecteurs. L'identité entre ces deux expressions des champs de tenseurs résulte du théorème général suivant :

Théorème IX.2 : théorème de tensorialité :

Soient $\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p$, $\mathrm{F}, (p+1)$ espaces fibrés vectoriels sur une variété différentielle $\mathfrak{V}$ tous de classe $\mathfrak{C}_r$ $(r \geqslant 1)$, $\mathfrak{F}$ l'algèbre des fonctions numériques de classe $\mathfrak{C}_r$ sur $\mathfrak{V}$.

Un morphisme multilinéaire φ :

$$\mathrm{E}_1 \times \mathrm{E}_2 \times \ldots \mathrm{E}_p \to \mathrm{F}.$$

autrement dit un morphisme linéaire $\widetilde{\varphi}$:

$$\mathrm{E}_1 \otimes \mathrm{E}_2 \otimes \ldots \mathrm{E}_p \to \mathrm{F}$$

détermine une application $\mathfrak{F}$-multilinéaire Φ entre sections des espaces fibres vectoriels :

$$\Gamma\mathrm{E}_1 \times \Gamma\mathrm{E}_2 \times \ldots \Gamma\mathrm{E}_p \to \Gamma\mathrm{F}$$

(autrement dit, une application $\mathfrak{F}$-linéaire $\widetilde{\Phi} : \Gamma\mathrm{E}_1 \otimes_{\mathfrak{F}} \Gamma\mathrm{E}_2 \otimes_{\mathfrak{F}} \ldots \Gamma\mathrm{E}_p \to \Gamma\mathrm{F}$)

Réciproquement, une application $\mathfrak{F}$-multilinéaire Φ :

$$\Gamma\mathrm{E}_1 \times \Gamma\mathrm{E}_2 \times \ldots \Gamma\mathrm{E}_p \to \Gamma\mathrm{F}$$

provient toujours d'un morphisme multilinéaire φ entre espaces fibrés vectoriels qu'elle détermine.

Preuve :

a) φ détermine Φ : si $\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_p$ sont des sections respectivement de $\mathrm{E}_1, \mathrm{E}_2, \ldots \mathrm{E}_p$ l'application $x \in \mathfrak{V} \to \mathrm{Y}_x = \varphi(\mathrm{X}_{1x}, \mathrm{X}_{2x}, \ldots, \mathrm{X}_{px}) \in \mathrm{F}_x$ est une section de F de classe $\mathfrak{C}_r$: $\mathrm{Y} = \Phi(\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_p)$.

b) Réciproquement, partant de Φ, il faut reconstituer φ. Supposons d'abord que $(\mathrm{X}_j)_{x_0} = 0$. Soient un voisinage ouvert assez petit de x_0 pour que E_j possède une carte au-dessus de U_0 : $(\mathrm{E}_j)_{\mathrm{U}_0} \to \mathrm{U}_0 \times \mathbb{R}^p$, $e_1, e_2, \ldots e_p$ les sections de E_j au-dessus de U_0 images réciproques de la base standard de $\mathbb{R}^p$. Si maintenant V et U sont deux voisinages ouverts de x_0 tels que $\overline{\mathrm{V}} \subset \mathrm{U}$, $\overline{\mathrm{U}} \subset \mathrm{U}_0$, $\overline{\mathrm{V}}$ et $\overline{\mathrm{U}}$ compacts, il existe une fonction numérique $\mathfrak{C}_r$, α à valeurs dans $[0, 1]$, égale à 1 sur $\overline{\mathrm{V}}$ et nulle sur $\mathfrak{V} - \mathrm{U}$.

Si $\beta = 1 - \alpha$, on a $\alpha^2 + \alpha\beta + \beta = 1$, $\alpha\mathrm{X}_j = \sum_{k=1}^{p} a_k e_k$ avec $a_k(x_0) = 0$, $\forall k$, et puisque Φ est supposée $\mathfrak{F}$-linéaire :

$$\Phi(\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_p) = \alpha(x) \sum_k a_k(x)\Phi(\mathrm{X}_1, \ldots \alpha(x)e_k, \ldots \mathrm{X}_p)$$
$$+\beta(x)(1 + \alpha(x))\Phi(\mathrm{X}_1, \ldots \mathrm{X}_p)_x$$

Pour $x = x_0$, le second membre est nul, ce qui prouve qu'il suffit que X_j soit nul en un point x_0 pour que la section $\Phi(\mathrm{X}_1, \ldots \mathrm{X}_2)$ de F soit nulle également en x_0.

Si maintenant deux sections X_j et X'_j de E ont la même valeur en un point x_0, il en résulte que

$$\Phi(X_1, \ldots X_j, \ldots X_p)_{x_0} = \Phi(X_1, \ldots X'_j, \ldots X_p)_{x_0}$$

Ceci étant valable quel que soit j, Φ définit bien en chaque point x de $\mathfrak{V}$ une application φ_x de $(E_1 \times E_2 \times \ldots E_p)_x$ dans F_x et l'on vérifie sur une carte locale qu'elle est bien multilinéaire et de classe $\mathfrak{C}_r$ donc un morphisme de $E_1 \times E_2 \times \ldots E_p$ dans F.

Une application immédiate du théorème précédent est le :

Corollaire IX.2 : Critère de tensorialité. Soient $\mathfrak{V}$ une variété différentielle, $\mathfrak{F}$ l'algèbre des fonctions numériques de classe $\mathfrak{C}_r$ sur $\mathfrak{V}$. Soit T une application $\mathfrak{F}$-multilinéaire faisant correspondre à tout ensemble de p champs de vecteurs $X_1, X_2, \ldots X_p$ et q champs de covecteurs $\omega_1, \omega_2, \ldots \omega_q$ un champ de tenseurs $T(X_1, \ldots X_p; \omega_1, \ldots \omega_q)$ de variance $(v) = (a, b)$ tous de classe $\mathfrak{C}_r$. T définit alors canoniquement un champ de tenseurs t de variance $(a+q, b+p)$, de classe $\mathfrak{C}_r$ sur $\mathfrak{V}$.

Définition IX.2.D :

Sur une variété différentielle $\mathfrak{V}$ de dimension n, de classe $\mathfrak{C}_r$, $r \geqslant 2$, une *forme différentielle extérieure de degré* p $(0 \leqslant p \leqslant n)$ est une fonction ω qui fait correspondre à chaque point x de $\mathfrak{V}$ une forme extérieure de degré p sur l'espace tangent T_x, soit : $\omega_x \in \wedge^p T^*_x$. Sur un ouvert U de $\mathfrak{V}$ possédant une carte, les différentielles $dx^1, dx^2, \ldots dx^n$ des fonctions coordonnées forment une base de chaque T^*_x, $\forall x \in U$, et ω s'écrit donc sur U : $\omega_x = \sum\limits_{I} a_I(x) dx^{i_1} \wedge dx^{i_2} \wedge \ldots dx^{i_p} = \sum a_I(x) dx^I$, la somme étant étendue aux multiindices $I = \{1 \leqslant i_1 < i_2 < \ldots i_p \leqslant n\}$ qui parcourent les $\binom{n}{p}$ suites strictement croissantes de p éléments de $(1, 2, \ldots n)$. ω est alors dite de classe $\mathfrak{C}_s$, $s \leqslant r-1$, si les fonctions coordonnées a_I sont des fonctions de classe $\mathfrak{C}_s$.

Cette dernière propriété ne dépend pas du choix de la carte comme le prouve la formule du changement de composantes de ω lors du passage d'un système de coordonnées $(x^1, x^2, \ldots x^n)$ à un autre $(y^1, y^2, \dots y^n)$. Si $\omega = \sum a_{\mathrm{I}} dx^{\mathrm{I}} = \sum b_{\mathrm{J}} dy^{\mathrm{J}}$:

$$\begin{aligned}\omega &= \sum_{\mathrm{I}} a_{\mathrm{I}}(x) \Big(\sum \frac{\partial x^{i_1}}{\partial y^{k_1}}\, dy^{k_1} \Big) \wedge \dots \Big(\sum \frac{\partial x^{i_p}}{\partial y^{k_p}}\, dy^{k_p} \Big) \\ &= \sum_{\mathrm{J}} \Big(\sum_{\mathrm{I}} a_{\mathrm{I}}(x) \frac{\mathrm{D}(x^{i_1}, \dots x^{i_p})}{\mathrm{D}(y^{j_1}, \dots y^{j_p})} \Big)\, dy^{j_1} \wedge \dots dy^{j_p} \\ &= \sum_{\mathrm{J}} b_{\mathrm{J}}(x) dy^{\mathrm{J}}\end{aligned}$$

avec $\displaystyle b_{\mathrm{J}}(x) = \sum_{\mathrm{I}} (a_{\mathrm{I}}(x)\ \frac{\mathrm{D}(x^{i_1}, \dots x^{i_p})}{\mathrm{D}(y^{j_1}, \dots y^{p_p})}$

où
$$\frac{\mathrm{D}(x^{i_1}, \dots x^{i_p})}{\mathrm{D}(y^{j_1}, \dots y^{j_p})} = \underset{\substack{i\in \mathrm{I} \\ j \in \mathrm{J}}}{\text{dét}} \left| \frac{\partial x^i}{\partial y^j} \right|.$$

Les formes différentielles extérieures d'une même classe $\mathfrak{C}_s$ sur $\mathfrak{V}$ forment une *algèbre graduée anticommutative* sur l'algèbre $\mathfrak{F}$ des fonctions numériques de classe $\mathfrak{C}_s$, qui sont les formes de degré zéro : $\alpha \wedge \omega = (-1)^{d^0\alpha \cdot d^0\omega} \omega \wedge \alpha$.

Si $\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_p$ sont p champs de vecteurs sur $\mathfrak{V}$, et $\omega = \sum a_{\mathrm{I}} dx^{\mathrm{I}}$, leur contraction est la fonction numérique :

$$\omega(\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_p) = \sum a_{\mathrm{I}}\ \text{dét} |\langle dx^{i_k}; \mathrm{X}_j \rangle|$$

ω est une fonction $\mathfrak{F}$-p-linéaire alternée des champs de vecteurs sur $\mathfrak{V}$.

Si Φ est une application $\mathfrak{C}_r$ de $\mathfrak{W}$ dans $\mathfrak{V}$; l'image réciproque par Φ, dans $\mathfrak{W}$, d'une forme différentielle extérieure ω de degré p sur $\mathfrak{V}$ est définie en un point y de $\mathfrak{W}$ par : $(\Phi^*\omega)_y = (\wedge^p \Phi_y^*)(\omega_{\Phi y})$. On a : $\Phi^*(\alpha \wedge \omega) = (\Phi^*\alpha) \wedge (\Phi^*\omega)$.

Suivant le principe : « devient structure sur $\mathfrak{V}$ tout ce qui est conservé par changement de cartes (ou de coordonnées) », la *différentielle extérieure* d est intrinsèquement définie sur toute variété différentielle. En effet, on a défini au §IV.14 la différentielle extérieure des formes différentielles sur un espace vectoriel réel $\mathbb{R}^n$, et démontré que cette définition était indépendante du choix de « coordonnée curvilignes » dans $\mathbb{R}^n$, c'est-à-dire invariante par « changement de cartes ». A chaque p-forme ω de classe $\mathfrak{C}_s$ sur $\mathfrak{V}$, d associe une $(p+1)$-forme $d\omega$ de classe $\mathfrak{C}_{s-1}$. Si dans des coordonnées locales, ω s'écrit $\omega(x) = \sum a_{\mathrm{I}}(x) dx^{\mathrm{I}}$ alors $d\omega(x) = \sum da_{\mathrm{I}}(x) \wedge dx^{\mathrm{I}}$

Rappelons les propriétés de la différentielle extérieure (§IV.14)

1) $\mathbb{R}$-linéaire : $d(\lambda\alpha + \mu\omega) = \lambda d\alpha + \mu d\omega$, $\forall\lambda, \mu \in \mathbb{R}$,

2) $d^2\omega = d(d\omega) = 0$,

3) $d(\alpha \wedge \omega) = d\alpha \wedge \omega + (-1)^{d_0\alpha}\alpha \wedge d\omega$,

4) si $X_0, X_1, \dots X_p$ sont $(p+1)$ champs de vecteurs et ω une p-forme sur $\mathfrak{V}$ (cf Théorème IV.14.A) :

$$\boxed{\begin{aligned} d\omega(X_0, X_1, \dots X_p) &= \sum_{i=0}^{p} (-1)^i X_i \cdot \omega(X_0, \dots \widehat{X}_i, \dots X_p) \\ &+ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots \widehat{X}_i, ..\widehat{X}_j \dots X_p) \end{aligned}}\,.$$

Définition IX.2.E :

Une variété différentielle $\mathfrak{V}$ de dimension n est dite *orientable* s'il est possible de choisir de façon continue une orientation des espaces tangents en tous ses points : la variété est alors dite orientée. Pour que $\mathfrak{V}$ soit orientable, il faut et il suffit qu'elle possède un atlas dont tous les changements de carte conservent l'orientation, ce qui s'exprime par le fait que les jacobiens :

$\dfrac{D(y^1, y^2, \dots y^n)}{D(x^1, x^2, \dots x^n)} = \text{dét}\left|\dfrac{\partial y^i}{\partial x^j}\right|$ associés aux passages de coordonnées $(x^1, x^2, \dots x^n)$ aux coordonnées $(y^1, y^2, \dots y^n)$ de l'atlas sont tous strictement positifs.

Si $\mathfrak{V}$ est orientable et connexe, le choix d'une orientation de T_{x_0} en un seul point x_0 détermine ainsi une orientation de tous les espaces tangents de $\mathfrak{V}$.

Si la variété *connexe* $\mathfrak{V}$ possède une structure pseudoriemannienne (définition IX.2.C) la métrique est caractérisée par son indice (r, s). On a alors une notion plus précise que la simple orientation de $\mathfrak{V}$: une *orientation complète* (§VII.12) consiste en l'orientation cohérente des sous-espaces tangents positifs maximaux (de dimension r) et une orientation cohérente des sous-espaces tangents négatifs maximaux (de dimension s) de $\mathfrak{V}$, les deux orientations partielles déterminant une orientation de $\mathfrak{V}$. Pour une telle variété $\overline{g} = \text{dét}|g_{ij}|$ a le signe de $(-1)^s$.

Exemples : Les espaces projectifs réels de dimension impaire $\mathbb{R}\mathrm{P}^{2k+1}$ sont des variétés orientables. Les espaces projectifs réels de dimension paire $\mathbb{R}\mathrm{P}^{2k}$ sont des variétés *non-orientables.*

Un *élément de volume* sur $\mathfrak{V}$ est une forme différentielle extérieure ω de degré $n = \dim \mathfrak{V}$ partout non nulle c'est-à-dire une section partout non nulle du fibré en droites $\wedge^n \mathrm{T}^*(\mathfrak{V})$. Un élément de volume oriente de façon continue les droites $\wedge^n \mathrm{T}^*(\mathfrak{V})$ et définit donc une orientation de $\mathfrak{V}$.

Sur une variété orientée, une métrique riemannienne g détermine canoniquement un élément de volume ω, obtenu en prenant l'unité de volume orienté ω_x de chaque espace tangent $\mathrm{T}(v)_x$ (Définition V.3.C). Dans un système de coordonnées locales d'orientation positive, cette forme s'écrit, si $\overline{g} = \text{dét}|g_{ij}|$:
$\omega = \overline{g}^{(1/2)} dx^1 \wedge dx^2 \wedge \ldots dx^n$. Elle est invariante par tout changement de coordonnées *conservant l'orientation*, et a exactement la même expression que l'élément de volume des sous-variétés de l'espace euclidien E_n (cf. VII.9).

IX.3 – Dérivations. Dérivations de Lie. Groupes et algèbres de Lie

Soient $\mathfrak{V}$ une variété différentielle de classe $\mathfrak{C}_r$, $\mathfrak{F}$ l'algèbre des fonctions numériques de classe $\mathfrak{C}_r$ sur $\mathfrak{V}$. On se pose le problème de définir des opérations de dérivations (ou de différentiation) sur les champs de tenseurs sur $\mathfrak{V}$, ou, plus généralement, sur les sections de certains fibrés vectoriels sur $\mathfrak{V}$, supposés de classe $\mathfrak{C}_r$.

Une dérivation a généralement pour effet de diminuer l'ordre de la classe $\mathfrak{C}_r$ d'une unité. Pour avoir des opérations « internes » et aussi pour éviter de jongler avec les ordres de dérivation en alourdissant l'exposé, il est souvent préférable de supposer que les variétés, champs de vecteurs ou de tenseurs *sont tous de classe* $\mathfrak{C}_\infty$, avec donc des dérivées continues de tous ordres dans des coordonnées locales.

Exemples d'opérations de dérivations :

1) la différentielle extérieure d des formes différentielles extérieures sur $\mathfrak{V}$, qui n'existe naturellement que sur la $\mathfrak{F}$-algèbre des sections différentiables du fibré $\wedge \mathrm{T}^*(\mathfrak{V})$; d applique $\mathfrak{F}$ dans les champs de covecteurs ou de formes différentielle de degré un.

2) si X est un champ de vecteurs, il détermine une dérivation i_x, produit intérieur par X, dans l'algèbre des formes différentielles extérieures, de la façon suivante : si $\omega \in \wedge^p \mathrm{T}^*(\mathfrak{V})$ est une forme extérieure de degré p sur $\mathfrak{V}$, $(i_\mathrm{X}\omega)_x$ est le produit intérieur $\mathrm{X}_x \lrcorner\, \omega_x$ (cf. §IV.13) et $i_\mathrm{X}\omega$ est ainsi une forme de degré $(p-1)$.

On a $i^2_\mathrm{X} = 0$ (§IV.13). Cette dérivation, nulle sur $\mathfrak{F}$, est purement algébrique.

3) si u est un champ de tenseurs de variance $(1, 1)$, section du fibré $\mathrm{T}(\mathfrak{V}) \otimes \mathrm{T}^*(\mathfrak{V})$, il détermine une dérivation de la $\mathfrak{F}$-algèbre des champs de tenseurs de toutes variances sur $\mathfrak{V}$, nulle sur $\mathfrak{F}$, de la façon suivante : si t est un champ de tenseurs ; $(\mathrm{D}_u t)_x$ est le tenseur $\mathrm{D}_{u_x} t_x$ dérivé de t_x par l'opérateur linéaire u_x (cf. §IV.13). Si t est de variance $(\mathfrak{v})$, il en résulte que $\mathrm{D}_u t$ est également de variance $(\mathfrak{v})$. Cette dérivation, nulle sur $\mathfrak{F}$, est purement algébrique.

D'une façon générale on peut imposer à une opération de dérivation D les axiomes suivants :

1) D est $\mathbb{R}$-linéaire.

2) D est définie sur $\mathfrak{F}$ (fonctions numériques $\mathfrak{C}_\infty$) et si $f \in \mathfrak{F}$ est plate en x, $(\mathrm{D}f)_x = 0$, ce qui peut être interprété en disant que D est « du premier ordre ». Si D est nulle sur $\mathfrak{F}$, D est dite *purement algébrique*. D est alors $\mathfrak{F}$-linéaire suivant 3).

3) si $\mathrm{D}t$ est défini et $f \in \mathfrak{F}$, $\mathrm{D}(ft)$ est défini et il existe une multiplication $*$ permettant d'écrire :

$$\mathrm{D}(ft)_x = f(x)\mathrm{D}t_x + \mathrm{D}f_x * t_x.$$

4) si pour la multiplication $*$ $\mathrm{D}t$, $\mathrm{D}u$, $\mathrm{D}(t * u)$ sont définis, on a : $\mathrm{D}(t * u) = \mathrm{D}t * u + \varepsilon t * \mathrm{D}u$ avec $\varepsilon = \pm 1$ ce qui est la propriété caractéristique des dérivations vis-à-vis des multiplications.

Ces axiomes entraînent le *caractère local* d'une telle opération D : la valeur du champ $\mathrm{D}t$ en un point x ne dépend que de la restriction de t à un voisinage arbitrairement petit U de x. Si l'on prend en effet, une fonction numérique α, $\mathfrak{C}_\infty$, égale à un sur le voisinage V de x avec $\overline{\mathrm{V}} \subset \mathrm{U}$, donc plate en x, et nulle en dehors de U, on a :

$$\mathrm{D}(\alpha t)_x = (\mathrm{D}\alpha)_x * t_x + \alpha(x)\mathrm{D}t_x = \mathrm{D}t_x.$$

Ce caractère local entraîne qu'une opération de dérivation D est entièrement déterminée par ses valeurs sur les fonctions numériques et sur des champs qui forment des générateurs locaux relativement aux opérations d'addition, de produit et de multiplication par les fonctions numériques.

Applications :

1) Soient D_1 et D_2 deux dérivations définies sur l'algèbre des champs de tenseurs différentiables de toutes variances, respectant la variance (de degré zéro) et commutant avec les contractions. Si D_1 et D_2 coïncident sur les fonctions numériques et les champs de vecteurs différentiables, elles sont identiques.

2) Deux dérivations définies sur l'algèbre des formes différentielles extérieures différentiables qui coïncident sur les fonctions numériques et les formes de degré un sont identiques.

Si D est une dérivation non nulle sur $\mathfrak{F}$, opérant sur certains champs de vecteurs V, on peut souhaiter en avoir une définition géométrique exprimant $(\mathrm{DV})_x$ sous la forme classique d'une limite comme : $\mathrm{DV}_x = \lim\limits_{h\to 0} 1/h(\mathrm{V}(t+h) - \mathrm{V}(t))$. Sous cette forme stricte, cela n'est naturellement possible que si $\mathrm{V}(t)$ et $\mathrm{V}(t+h)$ appartiennent au même espace, V étant alors une fonction différentiable de $\mathfrak{V}$ à valeurs dans un espace vectoriel fixe F. Mais si le champ V est une section d'un fibré vectoriel E, par exemple un champ de tenseurs, les espaces vectoriels que sont les fibres E_{x_1}, E_{x_2} en deux points distincts n'ont *a priori* aucun lien canonique entre eux. Si l'on veut soustraire un vecteur de l'un d'un vecteur de l'autre il faut avoir un moyen de transporter l'un sur l'autre. Nous allons voir deux procédés différents de tels transports qui vont donner, l'un, la dérivation covariante (§suivant IX.4), l'autre la dérivation de Lie, par laquelle nous commencerons.

Une dérivation D qui applique en elle-même l'algèbre $\mathfrak{F}$ des fonctions numériques $\mathfrak{C}_\infty$ sur une variété $\mathfrak{C}_\infty$, définit canoniquement un champ de vecteurs X par $\mathrm{X}_x \cdot f = (\mathrm{D}f)_x$. Le champ X caractérise l'action de D sur les fonctions numériques. On obtient par extension la :

Définition IX.3.A. Dérivation de Lie :

A tout champ de vecteurs continûment différentiable X sur une variété $\mathfrak{V}$ est associé un opérateur de dérivation L_X des champs de tenseurs différentiables sur $\mathfrak{V}$, appelé *dérivée de Lie par rapport à* X, et uniquement déterminé par les conditions :

a) si f est une fonction différentiable, $L_X \cdot f = X \cdot f$ (L_X dérive les fonctions).
b) si Y est un champ de vecteurs différentiable $L_X \cdot Y = [X, Y]$.
c) L_X conserve la variance des champs de tenseurs.
d) si t et u sont deux champs de tenseurs différentiables :

$$L_X(t \otimes u) = L_X t \otimes u + t \otimes L_X u \text{ (}L_X \text{ est une dérivation).}$$

e) L_X commute avec les contractions.

L_X est de caractère local et l'unicité est immédiate sur une carte.

Calculons à titre d'exemple la dérivée $L_X\alpha$ d'un champ de covecteurs α :

$$L_X(\alpha \otimes Y) = L_X\alpha \otimes Y + \alpha \otimes L_X Y.$$

En contractant : $X.\langle\alpha, Y\rangle - \langle\alpha, L_X Y\rangle = \langle L_X\alpha, Y\rangle$ qui donne $L_X\alpha$ comme fonction sur les champs de vecteurs de $\mathfrak{V}$ par : $L_X\alpha = L_X \circ \alpha - \alpha \circ L_X$.

Si X et Y sont deux champs de vecteurs, le crochet des deux dérivations $[L_X, L_Y]$ est une dérivation définie sur les champs de tenseurs différentiables. Sur les fonctions numériques et les champs de vecteurs, on a :

$$[L_X, L_Y]{\cdot}f = L_X(Y{\cdot}f) - L_Y(X{\cdot}f) = X(Y \cdot f) - Y(X{\cdot}f) = [X, Y]{\cdot}f$$
$$[L_X, L_Y]Z = [X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z] = L_{[X,Y]} \cdot Z$$

d'après l'identité de Jacobi.

On a donc : $[L_X, L_Y] = L_{[X,Y]}$.

Nous allons maintenant donner de la dérivation de Lie une définition « géométrique » dans le sens indiqué ci-dessus.

On peut considérer un champ de vecteurs tangents X sur une variété $\mathfrak{V}$ comme le champ de vecteurs-vitesse d'un mouvement de l'ensemble des points $\mathfrak{V}$:

- par exemple le champ de vecteurs constant sur $\mathbb{R}$: $X = ke_1$ est le champ des vitesses de la translation : $x \to \varphi(x) = x + kt$;
- mais si l'on considère ce même champ sur la sous-variété de $\mathbb{R}$ que forme un invertalle ouvert $]a, b[$, on ne peut plus considérer globalement sur $]a, b[$ la translation précédente, car elle fait « sortir » de $]a, b[$. On ne peut la considérer qu'en se restreignant aux intervalles compacts $[c, d] \subset]a, b[$ avec, pour chacun d'eux, un intervalle de temps borné : $|t| < \varepsilon$, par la condition que les points de $[c, d]$ restent dans $]a, b[$ lors des translations $x \to x+kt$, $|t| < \varepsilon$.

Définition IX.3.B :

Groupe à un paramètre de difféomorphismes C_r *d'une variété* $\mathfrak{V}$ *:* c'est une application $\mathfrak{C}_r$ de $\mathbb{R} \times \mathfrak{V}$ dans $\mathfrak{V}$:

$$\mathbb{R} \times \mathfrak{V} \xrightarrow{\varphi} \mathfrak{V}$$

$$(t, x) \longrightarrow \varphi_t(x)$$

telle que :
a) quels que soient $s, t \in \mathbb{R}$: $\varphi_{s+t} = \varphi_s \circ \varphi_t$
b) φ_0 est l'identité

il résulte de *a)* et *b)* que $\varphi_{-t} \circ \varphi_t \cdot x = x$, donc que φ_t est biunivoque, et que φ_t est un difféomorphisme $\mathfrak{C}_r$ de $\mathfrak{V}$ sur elle-même dont le réciproque est φ_{-t}.

Un groupe à un paramètre de difféomorphismes $\mathfrak{C}_r$ sur $\mathfrak{V}$ définit sur $\mathfrak{V}$ un champ de vecteurs tangents $\mathfrak{C}_{r-1}$ par :

$$X_x = \left(\frac{d\varphi_t(x)}{dt}\right)_{t=0}.$$

C'est le vecteur-vitesse au temps $t = 0$ de la trajectoire de x : $t \to \varphi_t(x)$.

Mais, si l'on considère le vecteur-vitesse au temps s quelconque de la trajectoire de x on a :

$$\left(\frac{d\varphi_t(x)}{dt}\right)_{t=s} = \left(\frac{d\varphi_{s+u}(x)}{du}\right)_{u=0} = \left(\frac{d\varphi_u(\varphi_s(x))}{du}\right)_{u=0} = X_{\varphi_s}(x).$$

Le champ X *est donc un champ de vecteurs-vitesse des trajectoires des points de* $\mathfrak{V}$. La trajectoire du point $\varphi_s(x)$:

$u \to \varphi_u(\varphi_s(x)) = \varphi_s(\varphi_u(x))$ est la translatée par le difféomorphisme φ_s de $\mathfrak{V}$ de la trajectoire du point $x : u \to \varphi_u(x)$ et :

$$\mathrm{X}_{\varphi_s}(x) = (\varphi_s)_* \cdot \mathrm{X}_x.$$

Définition IX.3.C :

« Groupe local » à 1 paramètre de difféomorphismes $\mathfrak{C}_r$ (locaux) de $\mathfrak{V}$: c'est une application $\mathfrak{C}_r$, dans $\mathfrak{V}$, d'un voisinage ouvert $\mathfrak{U}$ de $0 \times \mathfrak{V}$ dans $\mathbb{R} \times \mathfrak{V}$ tel que $\mathfrak{U} \cap (\mathbb{R} \times x) =]-\alpha_x, \beta_x[$ où α_x et β_x sont des nombres réels strictement positifs :

$$\begin{array}{ccc} \mathfrak{U} & \xrightarrow{\varphi} & \mathfrak{V} \\ (t, x) & \longrightarrow & \varphi_t(x) \end{array}$$

satisfaisant aux axiomes suivants :

a) si $s, t, s+t \in]-\alpha_x, \beta_x[$, on a $\varphi_{s+t}(x) = \varphi_s(\varphi_t(x))$;
b) φ_0 est l'identité de $\mathfrak{V}$.

Si U est un voisinage compact de $x \in \mathfrak{V}$, il existe $\varepsilon > 0$ tel que $[-\varepsilon, \varepsilon] \times \mathrm{U} \subset \mathfrak{U}$. Il résulte de *a)* et *b)* que pour tout t, $|t| \leqslant \varepsilon$, $\varphi_{-t} \circ \varphi_t$ est l'identité sur U ; φ_t est donc *localement un difféomorphisme.*

Par un raisonnement bien connu un groupe local à 1 paramètre de difféomorphismes de $\mathfrak{V}$ qui est défini sur un ouvert $\mathfrak{U}$ de $\mathbb{R} \times \mathfrak{V}$ contenant $[-\varepsilon, \varepsilon] \times \mathfrak{V}$, pour $\varepsilon > 0$, détermine un groupe *global* à 1 paramètre de difféomorphismes de $\mathfrak{V}$ [cf. Bourbaki, Topologie générale, Ch. V. (groupes à un paramètre) §1, n° 4 Proposition 6]. *C'est toujours le cas lorsque $\mathfrak{V}$ est compacte.*

Un groupe local à 1 paramètre de difféomorphisme $\mathfrak{C}_r$ de $\mathfrak{V}$ détermine un champ de vecteurs tangents $\mathfrak{C}_{r-1}$ sur $\mathfrak{V}$ par :

$$\mathrm{X}_x = \Big(\frac{d\varphi_t(x)}{dt}\Big)_{t=0}.$$

Par le même raisonnement que ci-dessus, on voit que X *est le champ des vecteurs-vitesse des trajectoires* : $t \in]-\alpha_x, \beta_x[\to \varphi_t(x)$ *des points de* $\mathfrak{V}$:

$$\Big(\frac{d\varphi_t(x)}{dt}\Big)_{t=s} = \mathrm{X}_{\varphi_x}(x).$$

Réciproquement, on a le théorème suivant :

Théorème IX.3 : Un champ de vecteurs tangents $\mathfrak{C}_{r-1}$: X sur une variété $\mathfrak{V}$ détermine un groupe local à 1 paramètre de difféomorphismes $\mathfrak{C}_r$ de $\mathfrak{V}$, obtenu en intégrant l'équation différentielle sur $\mathfrak{V}$:

$$\frac{d\varphi_t(x)}{dt} = \mathrm{X}_{\varphi_t}(x).$$

Si le champ X est à support compact, φ_t est un groupe global, défini $\forall t$.

Preuve : Soit W un ouvert de $\mathfrak{V}$ possèdant une carte sur un ouvert W′ de $\mathbb{R}^n$, définissant des coordonnées $x^1, x^2, \dots x^n$ sur W.

L'équation différentielle devient, en écrivant simplement $x^i(t)$ pour $x^i(\varphi_t(x))$:

$$\left\{\frac{dx^i}{dt} = \mathrm{X}^i(x^1, x^2, \dots x^n) \quad i = 1, 2, \dots n; \ (x^1, x^2, \dots x^n) \in \mathrm{W}'\right\}$$

où les X^i sont des fonctions $\mathfrak{C}_{r-1}$ de leur arguments.

Le théorème fondamental d'*existence et unicité* affirme alors que pour tout $x = (x^1, x^2, \dots x^n)$ de W′, on peut trouver un voisinage U de x dans W′, un voisinage I de 0 dans $\mathbb{R}$ et une application $\mathfrak{C}_r$ de $\mathrm{I} \times \mathrm{U}'$ dans :

$$\mathrm{W}' : \ \mathrm{I} \times \mathrm{U}' \xrightarrow{\varphi} \mathrm{W}$$

$$(t, x_0^1, \dots x_0^n) \longrightarrow \{\varphi^1(x_0^1, \dots x_0^n), \dots \varphi^n(t, x_0^1, \dots x_0^n)\}$$

telle que :

$$\frac{d\varphi^i}{dt} = \mathrm{X}^i(\varphi_1, \dots \varphi^n)$$
$$\varphi^i(0, x_0^1, \dots x_0^n) = x_0^i.$$

De plus, toute application Ψ satisfaisant aux équations (1) dans un ouvert quelconque $\mathrm{I}_1 \times \mathrm{U}_1' \subset \mathrm{I} \times \mathrm{U}'$ coïncide avec φ dans cet ouvert.

Puisque les second membres X^i des n première équations (1) ne contiennent pas la variable t, on a :

$$\frac{d\varphi_{s+t}}{dt} = \frac{d\varphi_{s+t}}{d(s+t)}.$$

L'unicité entraîne donc que $\varphi_{s+t}(x_0) = \varphi_t(\varphi_s(x_0))$ c.q.f.d.

Soient maintenant X et Y deux champs de vecteurs de classe $r \geqslant 1$ sur $\mathfrak{V}$, et φ le groupe local à 1 paramètre associé à X. Le vecteur tangent $(\varphi_{-t})_*.\mathrm{Y}_{\varphi_t x}$ est un vecteur tangent au point x et la translation locale φ_{-t} permet alors de comparer les vecteurs Y_x au point x et $\mathrm{Y}_{\varphi_t x}$ au point $\varphi_t x$. On obtient la définition « géométrique » de la dérivation de Lie sur un champ de vecteurs :

$$\mathrm{L_X Y} = [\mathrm{X}, \mathrm{Y}] = \lim_{t\to 0} \frac{1}{t} \{(\varphi_{-t})_* \mathrm{Y}_{\varphi_t x} - \mathrm{Y}_x\} = \frac{d}{dt} \{(\varphi_{-t})_* \mathrm{Y}_{\varphi_t x}\}$$

et plus généralement sur un champ de tenseurs u :

$$\mathrm{L_X} u = \lim_{t\to 0} \frac{1}{t} \{(\varphi_{-t})_* u_{\varphi_t} - u_x\}.$$

Il suffit de le vérifier pour les champs de vecteurs au moyen d'une carte locale. Soit dans une telle carte :

$$\mathrm{Z}^i(t) = \{\{(\varphi_{-t})_* \mathrm{Y}_{\varphi_t x}\}^i = \sum_j \frac{\partial \varphi_i(-t;x)}{\partial x^j} \mathrm{Y}^j(\varphi(t;x))$$

$$\begin{aligned}\Big(\frac{d\mathrm{Z}^i}{dt}\Big)_0 &= -\frac{\partial}{\partial x^j}\Big(\frac{\partial \varphi i(-t,x)}{dt}\Big)_0 \mathrm{Y}^j_x \\ &\quad + \sum_{j,k}\Big(\frac{\partial \varphi i(-t,x_0)}{\partial x^j}\Big)_0 \frac{\partial \mathrm{Y}^j}{\partial x^k}\Big(\frac{d\varphi k(t,x)}{dt}\Big)_0 \\ &= \Big(-\sum_j \frac{\partial \mathrm{X}^i}{\partial x^j}\mathrm{Y}^j + \sum_k \mathrm{X}^k \frac{\partial \mathrm{Y}^i}{\partial x^k}\Big)_x = [\mathrm{X},\mathrm{Y}]^i_x\end{aligned}$$

puisque la matrice $\left|\dfrac{\partial \varphi^i}{\partial x^j}\right|$ a pour limite l'identité pour $t = 0$.

Exemple : Tenseur de déformation métrique.

Soit X un champ de vecteurs sur une variété riemannienne ou pseudoriemanienne $(\mathfrak{V} g)$. Le champ de tenseurs $\mathrm{L_X} g$ est le *tenseur de déformation métrique par* X (ou par le groupe local φ_t associé).

On a $(\mathrm{L_X} g)(\mathrm{Y},\mathrm{Z}) = \mathrm{X}\cdot g(\mathrm{Y},\mathrm{Z}) - g([\mathrm{X},\mathrm{Y}]\mathrm{Z}) - g(\mathrm{Y},[\mathrm{X},\mathrm{Z}])$.

En remplaçant Y par $\dfrac{\partial}{\partial x^i}$, Z par $\dfrac{\partial}{\partial x^j}$ et X par $\sum \mathrm{X}^k \dfrac{\partial}{\partial x^k}$ on obtient :

$$(\mathrm{L_X} g)_{ij} = \sum \mathrm{X}^k \frac{\partial g_{ij}}{\partial x^k} + \sum g_{kj} \frac{\partial \mathrm{X}^k}{\partial x^i} + \sum g_{ik} \frac{\partial \mathrm{X}^k}{\partial x^j}.$$

Si la métrique est euclidienne, les composantes de ce tenseur s'écrivent en coordonnées orthonormales :

$$(\mathrm{L_X} g)_{ij} = \frac{\partial \mathrm{X}^j}{\partial x^i} + \frac{\partial \mathrm{X}^i}{\partial x^j}.$$

On peut comparer avec la fin du §VII.6. où l'on a défini le tenseur de déformation infinitésimale « des longueurs » E. E est la moitié du tenseur de déformation métrique. Celui-ci correspond en effet à la déformation infinitésimale des « carrés des longueurs ».

Les dérivations de Lie opèrent en particulier sur les formes différentielles extérieures. Si ω est une telle forme,

$$(\mathrm{L_X}\omega) = \lim_{t\to 0} \frac{1}{t} \left\{ (\varphi_{-t})_* \omega_{\varphi_t x} - \omega_x \right\} = \frac{d}{dt} \left\{ (\varphi_{-t})_* \omega_{\varphi_t x} \right\}_t = 0.$$

On en déduit :

a) $\mathrm{L_X}(\omega \wedge \Pi) = \mathrm{L_X}\omega \wedge \Pi + \omega \wedge \mathrm{L_X}\Pi$;

b) puisque, quelle que soit l'application différentiable Φ : $\Phi^* d = d\Phi^*$, on a $\mathrm{L_X}(d\omega) = d(\mathrm{L_X}\omega)$: $\mathrm{L_X}$ *commute avec* d. Exemple : $\mathrm{L_X} dx^i = d(\mathrm{X} \cdot x^i) = d\mathrm{X}^i$.

Réciproquement, toute dérivation de l'algèbre des formes différentielles extérieures de degré zéro qui commute avec d, est une dérivation de Lie. En effet, l'action d'une telle dérivation D sur les fonctions numériques différentielles est celle d'un champ de vecteurs X.

On a alors $\mathrm{D}(df) = d(\mathrm{D}f) = d(\mathrm{L_X} f) = \mathrm{L_X}(df)$ et il en résulte que $\mathrm{D} \equiv \mathrm{L_X}$.

Explicitons l'expression de $\mathrm{L_X}\omega$:

$$\boxed{\begin{aligned} (\mathrm{L_X}\omega)(\mathrm{Y}_1, \dots \mathrm{Y}_p) &= \mathrm{X} \cdot \omega(\mathrm{Y}_1, \dots \mathrm{Y}_p) \\ &\quad - \sum_{i=1}^{p} \omega(\mathrm{Y}_1, \dots [\mathrm{X}, \mathrm{Y}_i], \dots \mathrm{Y}_p) \end{aligned}}\,.$$

Si X est un champ de vecteurs et ω une p-forme extérieure sur $\mathfrak{V}$, on a vu ci-dessus que le produit intérieur $i_X\omega$, qui est une $(p-1)$-forme extérieure définie en chaque x par : $(i_X\omega)_x = i_{X_x}\omega_x$ soit : $(i_X\omega)(Y_1, \ldots, Y_{p-1}) = \omega(X, Y_1, \ldots Y_{p1})$ est une dérivation de degré (-1) des formes différentielles extérieures et satisfait aux propriétés suivantes :

a) $i_X^2 = 0$;

b) $i_X(\omega \wedge \Pi) = i_X\omega \wedge \Pi + (-1)^{\deg\cdot\omega}\omega \wedge i_X\Pi$.

Dès lors, le crochet des deux dérivations, d et i_X, soit $d \circ i_X + i_X \circ d$, est une dérivation de degré zéro des formes différentielles extérieures, qui commute avec d puisque $d^2 = 0$, qui est égale à L_X sur les fonctions numériques et qui est donc identique à la dérivation de Lie L_X :

$$\boxed{d \circ i_X + i_X \circ d = L_X}.$$

Cette égalité utilisée dans l'expression de : $d\omega(X_0, X_1, \ldots X_p) = (i_{X_0} \circ d\omega)(X_1, \ldots X_p)$ redonne, à l'aide d'une récurrence, l'expression déjà trouvée de la différentielle extérieure au moyen des crochets des champs de vecteurs.

Si ω est une forme de degré un, la contraction de l'égalité :

$$L_X(\omega \otimes Y) = (L_X\omega) \otimes Y + \omega \otimes L_XY$$

donne : $X \cdot \omega(Y) = (L_X\omega)(Y) + \omega([X, Y])$

Il résulte de cette égalité que la dérivation intérieure $i_{[X,Y]}$ et le crochet des deux dérivations L_X et i_Y sont égales sur les fonctions numériques et les formes de degré un donc identiques sur l'algèbre des formes différentielles extérieures :

$$\boxed{i_{[X,Y]} = [L_X, i_Y] = L_X \circ i_Y - i_Y \circ L_X}.$$

Définition IX.3.D :

Soit $\mathfrak{V}$ une variété différentielle de classe $\mathfrak{C}_r$, $r \geqslant 2$ possédant comme seule structure supplémentaire un élément de volume ω (cf. fin du §IX.2). A tout champs de vecteurs X sur $\mathfrak{V}$ correspond une fonction numérique appelée la *divergence du champ de vecteurs* X, notée $\operatorname{div} X$, définie par l'effet de la dérivée de Lie L_X sur ω :

$$\boxed{L_X\omega = (\operatorname{div} X)\omega}.$$

Puisque $L_X = d \circ i_X + i_X \circ d$ et que $d\omega = 0$, cette définition équivaut à définir $\operatorname{div} X$ par (§IV.10)

$$(\operatorname{div} X)\omega = \mathrm{d}i(X)\omega = (-1)^{n-1} d(X \lrcorner\, \omega)$$

Si f est une fonction numérique différentiable : $\operatorname{div}(fX) = f \operatorname{div} X + Xf$

Localement, dans des coordonnées $x^1, x^2, \dots x^n$, la forme volume s'écrit : $\omega = a dx^1 \wedge dx^2 \wedge \dots dx^n$, et l'on a, puisque $L_X dx^j = dX^j$: $L_X\omega = (X \cdot a) dx^1 \wedge \dots dx^n + a \sum_j dx^1 \wedge \dots dX^j \wedge \dots dx^n$

d'où :
$$\boxed{\operatorname{div} X = \frac{1}{a}\,(X \cdot a) + \sum_j \frac{\partial X^j}{\partial x^j}}.$$

Si $\mathfrak{V}$ est une variété *riemannienne ou pseudoriemannienne orientée* elle possède une forme volume canonique (fin du §IX.2) qui s'écrit localement : $\omega = |\overline{g}|^{1/2} dx^1 \wedge dx^2 \wedge \dots dx^n$ avec $\overline{g} = \text{dét}|g_{ij}|$.

La divergence d'un champ de vecteurs X sur une variété pseudoriemannienne orientée a donc pour expression locale :

$$\boxed{\operatorname{div} X = |g|^{-1/2} \sum_j X^j \frac{\partial |\overline{g}|^{1/2}}{\partial x^j} + \sum_j \frac{\partial X^j}{\partial x^j}},$$

ce que l'on écrit généralement sous la forme condensée déjà trouvée dans le cas des espaces pseudoeuclidiens (§VII.9 et 10) :

$$\boxed{\operatorname{div} X = |\overline{g}|^{-1/2} \sum_j \frac{\partial}{\partial x^j}\,(|\overline{g}|^{1/2} X^j)}.$$

Nous donnerons à la fin du paragraphe suivant une interprétation et une extension de l'opération de divergence.

Si f est une fonction numérique deux fois continûment différentiable sur la variété pseudoriemannienne $\mathfrak{V}$, la métrique g associe au champ de covecteurs df un champ de vecteurs, *le gradient de* f, par $(\operatorname{grad} f \mid Y) = g(\operatorname{grad} f, Y) = df(Y) = Y \cdot f$.

Localement : $\operatorname{grad} f = \sum\limits_j \Big(\sum\limits_k g^{jk} \dfrac{\partial f}{\partial x^k} \Big) \dfrac{\partial}{\partial x^j}$.

Le *Laplacien de* f est la fonction numérique $\Delta f = \operatorname{div}(\operatorname{grad} f)$ qui s'écrit localement :

$$\boxed{\Delta f = |\overline{g}|^{-1/2} \sum_j \frac{\partial}{\partial x^j} \Big(|\overline{g}|^{1/2} \cdot \sum_k g^{ik} \frac{\partial f}{\partial x^k} \Big)}.$$

Nous allons maintenant appliquer ce qui vient d'être vu aux §1, 2, 3 à une situation particulièrement importante :

Définition IX.3.E :

Groupes et algèbres de Lie. Un *groupe de Lie* G est une variété différentielle $\mathfrak{C}_\infty$ munie d'une loi de groupe, la compatibilité entre les deux structures s'exprimant par les conditions que la multiplication : $(g, h) \to gh$ et l'opération τ, prise de l'inverse : $g \to \tau(g) = \overset{-1}{g}$ soient $\mathfrak{C}_\infty$. Une *algèbre de Lie* sur $\mathbb{R}$ est un espace vectoriel sur $\mathbb{R}$ muni d'une multiplication antisymétrique, notée : $(X, Y) \to [X, Y]$, satisfaisant à l'identité de Jacobi :

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0, \quad \forall X, Y, Z.$$

Exemples :

a) les groupes linéaires réels et leurs sous-groupes : orthogonaux, symplectiques, unitaires sont les groupes de Lie classiques;

b) les champs de vecteurs $\mathfrak{C}_\infty$ sur une variété différentielle $\mathfrak{V}$ de classe $\mathfrak{C}_\infty$ forment, avec le crochet, une algèbre de dimension infinie.

Si G est un groupe de Lie, ses translations à gauche $l(g)$ $\forall g \in G : x \to l(g) = gx$, à droite $r(g) : x \to r(g)x = xg$ sont des difféomorphismes $\mathfrak{C}_\infty$. Un champ de vecteurs tangents X sur G est dit invariant à gauche, si $l(g)_* X$, $\forall g$, et X est alors entièrement déterminé par sa valeur en un point, par exemple l'élément neutre e. Réciproquement, si l'on se donne $X_e \in T(G_e)$ et que l'on définit $X_g = l(g)_* X_e$, on a : $l(h)_* X_g = l(h)_* l(g)_* X_e = l(hg)_* X_e = X_{hg}$ et X est un champ $\mathfrak{C}_\infty$ invariant à gauche. De la même façon, la donnée de X_e définit un champ $\mathfrak{C}_\infty$ invariant à droite $\overline{X}$, par $\overline{X}_g = r(g)_* X_e$.

Puisqu'un difféomorphisme respecte le crochet (§IV.2), on a : $l(g)_*[X, Y] = [l(g)_* X,\ l(g)_* Y] = [X, Y]$. Les champs de vecteurs tangents sur G invariants à gauche forment donc une sous-algèbre de Lie de dimension finie égale à la dimension de la variété G, appelée *algèbre de Lie* $\mathfrak{g}$ *de* G. De la même façon de vecteurs tangents invariants à droite forment une algèbre de Lie $\widetilde{\mathfrak{g}}$. Les espaces vectoriels de $\mathfrak{g}$ et $\widetilde{\mathfrak{g}}$ s'identifient à l'espace tangent en e à G.

Le difféomorphisme τ transforme le champ invariant à gauche défini par X_e en le champ invariant à droite défini par $-X_e$. En effet $\tau \circ l(g)(x) = \tau(gx) = \overset{-1}{x}\,\overset{-1}{g} = r(g)^{-1} \tau x$ et $\tau_*(X_g) = \tau_* l(g)_* X_e$ $= r(\overset{-1}{g})_* \tau_* X_e = r(\overset{-1}{g})_*(-X_e) = -\widetilde{X}_g$.

Une translation à droite $r(x)$ de G par un élément x transforme un champ de vecteurs tangents invariants à gauche X en un autre champ invariant à gauche.

En effet, soit $Y = r(x)_* X$, soit $Y_g = r(x)_* X_{g\overset{-1}{x}}$, on a $l(h)_* Y_g$ $= l(h)_* r(x)_* X_{g\overset{-1}{x}} = r(x)_* l(h)_* X_{g\overset{-1}{x}} = r(x)_* X_{hg\overset{-1}{x}} = Y_{hg}$.

L'automorphisme intérieur : $g \to xg\overset{-1}{x} = l(x) r(\overset{-1}{x}) g = \operatorname{ad} x \cdot g$ de G est couramment appelé l'automorphisme adjoint de x.

Son application tangente, notée $\operatorname{Ad} x$ est un automorphisme de $\mathfrak{g} = T(G)_e$ et l'application $x \in G \to \operatorname{Ad} x$ représente G comme groupe d'automorphismes de son algèbre de Lie. Le vecteur de $\mathfrak{g}$, image par $\operatorname{Ad} \overset{-1}{x}$ de X est :

$$\begin{aligned} \operatorname{Ad} \overset{-1}{x} \cdot X_{xg\overset{-1}{x}} &= l(\overset{-1}{x})_* r(x)_* X_{xg\overset{-1}{x}} = r(x)_* l(\overset{-1}{x})_* X_{xg\overset{-1}{x}} \\ &= r(x)_* X_{g\overset{-1}{x}} = (r(x)_* X)_g = Y_g. \end{aligned}$$

On a donc $r(x)_* X = \operatorname{Ad} \overset{-1}{x} \cdot X$.

Un groupe de Lie G *opère à gauche*, resp. à *droite*, sur une variété différentielle $\mathbf{V}$, par la donnée d'une application différentiable :

$$G \times \mathbf{V} \to \mathbf{V} : \ (g, x) \to gx, \ \text{resp. } \mathbf{V} \times G \to \mathbf{V} : \ (x, g) \to xg$$

assujettie à la condition : $g(hx) = (gh)x$, resp. $(xg)h = x(gh)$, $\forall x \in \mathbf{V}, g, h \in \mathrm{G}$.

Il en résulte, en faisant $h = \overset{-1}{g}$, que l'élément neutre de G agit sur $\mathbf{V}$ comme l'identité, et l'opération de G sur $\mathbf{V}$ est dite fidèle si quel que soit $x \in \mathbf{V}$, $xg = x$ entraîne $g = e$.

Par exemple, un groupe de Lie G opère sur lui-même par ses translations à gauche et à droite.

L'opération d'un groupe de Lie G sur une variété $\mathbf{V}$ réalise G comme groupe d'automorphismes de l'algèbre $\mathfrak{F}$ des fonctions $\mathfrak{C}_\infty$ sur $\mathbf{V}$.

Si $g \in \mathrm{G}$ et $f \in \mathfrak{F}$: $(g \cdot f)(x) = f(\overset{-1}{g}\, x)$, resp. $(g \cdot f)(x) = f(xg)$ G opère à gauche sur l'algèbre $\mathfrak{F}$, puisque :

$$(g(hf))(x) = (hf)(\overset{-1}{g}\, x) = f((\overset{-1}{h}\,\overset{-1}{g})x) = ((gh)f)(x)$$

lorsque G opère à gauche sur $\mathbf{V}$, tandis que $g(hf)(x) = (hf)(xg) = f(x(gh)) = ((gh)f)(x)$ lorsque G opère à droite.

Soient Φ un difféomorphisme de $\mathbf{V}$ et $\Phi_*\mathrm{X}$ l'image par Φ du champ de vecteurs X d'un groupe local à un paramètre φ. On a :

$$\begin{aligned}
(\Phi_*\mathrm{X})_x \cdot f &= (\mathrm{X} \cdot (f \circ \Phi))_{\overset{-1}{\Phi}\, x} \\
&= \lim_{t \to 0} \frac{1}{t}\ \{(f \circ \Phi)\varphi_t(\overset{-1}{\Phi}\, x) - (f \circ \Phi)(\overset{-1}{\Phi}\, x)\} \\
&= \lim_{t \to 0} \frac{1}{t}\ \{f(\Phi\varphi_t\, \overset{-1}{\Phi}(x)) - f(x)\} = \Big(\frac{d(\Phi\varphi\, \overset{-1}{\Phi})(t, x)}{dt}\Big)_{t=0}
\end{aligned}$$

d'où il résulte que $\Phi_*\mathrm{X}$ est le champ de vecteurs tangents du groupe local à un paramètre $\Phi\varphi\,\overset{-1}{\Phi}$ et que X est invariant par Φ si et seulement si Φ commute avec φ.

Si X est un champ de vecteurs invariant à gauche du groupe de Lie G, il détermine un groupe global à un paramètre φ de difféomorphismes de G qui commutent avec les translations à gauche. Le raisonnement fait ci-dessus montre immédiatement que la trajectoire $g(t)$ de l'élément neutre est un sous-groupe à un paramètre de G que l'on note usuellement : $g(t) = \exp(t\mathrm{X}_e)$ par extension du cas où G est un groupe de matrices (cf. Exemple 1 ci-dessous). L'exponentielle est donc l'homomorphisme du groupe additif $\{t\mathrm{X}_e\}$, droite de l'espace tangent à G en e sur le sous-groupe à un paramètre $\exp(t\mathrm{X}_e)$ de G.

Le groupe à un paramètre φ est tout simplement celui des translations à droite de G par les éléments de $g(t)$: $\varphi(t;x) = \varphi(t;xe) = x\varphi(t;e) = xg(t) = x\,\exp(t\mathrm{X}_e)$.

Si $\widetilde{\mathrm{X}}$ est le champ invariant à droite défini par X_e le groupe $\widetilde{\varphi}$ de difféomorphismes de G qu'il détermine est celui des translations à gauche par les éléments de $g(t) : \widetilde{\varphi}(t;x) = g(t)x = \exp(t\,\mathrm{X}_e)\cdot x$

Dans le cas où un groupe de Lie G opère sur une variété $\mathbf{V}$, chaque vecteur tangent A en l'élément neutre e de G détermine un sous-groupe à un paramètre $\exp(t\mathrm{A})$ de G, donc un groupe à un paramètre de difféomorphismes de $\mathbf{V}$, d'où un champ de vecteurs $\widetilde{\mathrm{A}}$ sur $\mathbf{B}$: $\widetilde{\mathrm{A}}_x$ est le vecteur-vitesse en $t = 0$ du chemin $\exp(t\,\mathrm{A})x$ si G opère à gauche, ou $x\,\exp t\mathrm{A}$ si G opère à droite.

Considérons sur la variété produit $\mathbf{V}\times\mathrm{G}$, le champ de vecteurs tangents $(0,\overline{\mathrm{A}})$ où $\overline{\mathrm{A}}$ est le champ de vecteurs invariant à gauche sur G défini par $\mathrm{A}\in\mathrm{T(G)}_e$. Au point (x,g) le champ a pour valeur $(\mathrm{O}_x,\overline{\mathrm{A}}_g)$ et est tangent au chemin $(x, g\,\exp\mathrm{A}t)$ pour $t = 0$. Dans l'action μ de G sur $\mathbf{V}$: $\mathrm{V}\times\mathrm{G}\xrightarrow{\mu}\mathbf{V}$, le point $(x, g\,\exp\mathrm{A}t)$ a pour image $xg\,\exp\mathrm{A}t$ et $\mu_*(\mathrm{O}_x,\overline{\mathrm{A}}_g) = \widetilde{\mathrm{A}}_{xg}$. Les champs $(\mathrm{O},\overline{\mathrm{A}})$ sur $\mathbf{V}\times\mathrm{G}$ et $\widetilde{\mathrm{A}}$ sur $\mathbf{V}$ sont donc compatibles pour μ. Lorsque G opère à droite sur $\mathbf{V}$. Il en résulte que $[\overline{\mathrm{A}},\overline{\mathrm{B}}]$ a pour image $[\widetilde{\mathrm{A}},\widetilde{\mathrm{B}}]$ et que l'application $\mathrm{A}\in\mathfrak{g}\longrightarrow\widetilde{\mathrm{A}}\in\mathfrak{X}(\mathbf{V})$ est un homomorphisme d'algèbres de Lie de $\mathfrak{g}$ dans $\mathfrak{X}(\mathbf{V})$.

D'après ce que l'on vient de voir ci-dessus, le champ de vecteurs $r(g)\cdot r(\exp\cdot\mathrm{A}t)r(\overset{-1}{g}) = r(\overset{-1}{g}\cdot\exp\ \mathrm{A}t\cdot g)) = r(\widetilde{\mathrm{ad}\,\overset{-1}{g}}\cdot\exp\ \mathrm{A}t)$ $= r(\exp(\mathrm{Ad}\,\overset{-1}{g}\cdot\mathrm{A})t)$ ce qui montre que $r(g)_*\widetilde{\mathrm{A}} = (\mathrm{Ad}\,\overset{-1}{g}\cdot\mathrm{A})$.

L'algèbre de Lie $\mathfrak{g}$ de G est ainsi représentée comme une algèbre de Lie des dérivations des fonctions sur $\mathbf{V}$, c'est-à-dire comme une *algèbre de Lie d'opérateurs différentiels du premier ordre* sur $\mathbf{V}$. Plus généralement, supposons $\mathbf{V}$ de classe $\mathfrak{C}_\infty$ et soit $\mathbf{A}$ l'algèbre des formes différentielles extérieures $\mathfrak{C}_\infty$ sur $\mathbf{V}$. L'opération à droite d'un groupe de Lie $\mathbf{G}$ sur $\mathbf{V}$ représentant chaque élément X de son algèbre de Lie $\mathfrak{g}$ par un champ de vecteur $\widetilde{\mathrm{X}}$ de $\mathbf{V}$ détermine par la-même des dérivations associées $\mathrm{L}_\mathrm{X} = \mathrm{L}_{\widetilde{\mathrm{X}}}$, de degré 0, et $i_\mathrm{X} = i_{\widetilde{\mathrm{X}}}$, de degré (-1), de l'algèbre $\mathbf{A}$. Ces dérivations vérifient les équations vues ci-dessus :

1) $L_X \cdot f = X \cdot f, \forall f \in \mathfrak{F}$, 2) $L_X d = dL_X$; 3) $L_{[X,Y]} = [L_X, L_Y]$
4) $L_X = [i_X, d] = i_X d + di_X$; 5) $i_{[X,Y]} = [L_X, i_Y] = L_X i_Y - i_Y L_X$.

On dit alors que l'algèbre $\mathbf{A}$ est munie d'une « $\mathfrak{g}$-*structure* ».

Application : Le cas le plus simple est celui d'un sous-groupe fermé G du groupe linéaire $\mathrm{Gl}(n) = \mathrm{Gl}(\mathbb{R}^n)$ des automorphismes de l'espace vectoriel réel $E \approx \mathbb{R}^n$. G opère alors à gauche sur E : $x \to gx$, et à gauche et à droite sur l'espace vectoriel $\mathfrak{L}(E) \approx \mathbb{R}^{n^2}$: $l(g)u = gu$ et $r(g)u = ug$. Les translations à gauche et à droite de G sont les restrictions à G des opérateurs linéaires $l(g)$ et $r(g)$. Les opérateurs linéaires $g \in \mathfrak{L}(E), l(g), r(g) \in \mathfrak{L}(\mathfrak{L}(E))$ sont identiques à leurs applications différentielles en chaque point. Une base (e_j) de E détermine des coordonnées (x^j) dans E et (x^i_j) dans $\mathfrak{L}(E)$, qui fournissent des coordonnées, en général surabondantes, pour G. Un vecteur tangent X_u en $u \in G$ est entièrement déterminé par la matrice $|X^i_j| = |X_u \cdot x^i_j(u)|$. Calculons l'image du vecteur tangent X_u par la translation à gauche $l(g)$.

On a $(l(g)_* X_u) = X_u \cdot (f \circ l(g))$ et sa matrice est donc :

$$\begin{aligned}(l(g)X_u)^i_j &= (l(g)X_u) \cdot x^i_j = X_u(x^i_j(gu)) = X_u\Big(\sum g^i_k x^k_j(u)\Big) \\ &= \sum g^i_k X^k_j = (gX)^i_j\end{aligned}$$

d'où simplement $l(g)_* X_u = g \cdot X_u$. Un champ de vecteurs invariant à gauche X sur G s'obtient donc à partir d'un vecteur tangent à l'élément neutre, défini par la matrice A, par : $X_g = gA$ (produit de matrices).

Le groupe à un paramètre engendré par la matrice A, considérée comme vecteur tangent en l'élément neutre I, est défini par l'équation différentielle : $\dfrac{dM(t)}{dt} = M(t)A$ avec $M(o) = I$.

C'est le groupe multiplicatif de matrices $\exp(tA)$ $(-\infty < t < +\infty)$.

Si Y est un autre champ de vecteurs tangents invariants à gauche sur G, défini par $Y_g = g \cdot B$, calculons le crochet $[X, Y]$ par sa matrice. On a

$$\begin{aligned}X \cdot (Y \cdot x^i_j(g)) &= X \cdot \Big(\sum_k g^i_k b^k_j\Big) \\ &= \sum_{l,k} g^i_l a^l_k b^k_j = (g\ AB)^i_j.\end{aligned}$$

D'où : $[X, Y] = g \cdot [A, B]$ ce qui prouve que le crochet de deux champs de vecteurs tangents invariants à gauche sur G est invariant à gauche et que le crochet de l'algèbre de Lie $\mathfrak{g}$ de G est celui de Lie de $\mathcal{L}(E)$. On peut ainsi identifier $\mathfrak{g}$ à une sous-algèbre de Lie de $\mathcal{L}(E)$.

Calculons maintenant l'opérateur différentiel de E image du champ X invariant à gauche sur G : $X_g = gA$.

Si $g(t) = \exp(At)$ est le sous-groupe à un paramètre de G ayant A pour vecteur tangent en l'élément neutre I et si f est une fonction numérique $\mathfrak{C}^\infty$ sur E :

$$\widetilde{A}_x \cdot f = \lim_{t \to 0} \frac{1}{t}\, f(g(t)x - x) = df_x(Ax) = \Big(\sum_k \frac{\partial f}{\partial x^k}\, dx^k\Big)\big(\sum_{i,j} A^i_j x^j e_i\big)$$
$$= \sum_{i,j} \frac{\partial f}{\partial x^i}\, A^i_j x^j = \Big(\sum_{i,j} A^i_j x^j \frac{\partial}{\partial x^i}\Big) \cdot f = \Big(\sum_i (Ax)^i \frac{\partial}{\partial x^i}\Big) f.$$

L'opérateur différentiel sur E image de $X \in g$ est donc l'opérateur différentiel linéaire $\widetilde{A}_x = \sum_i (Ax)^i \dfrac{\partial}{\partial x^i}$.

Exemples : L'opérateur différentiel de $\mathbb{R}^2$ image de la matrice antisymétrique $A = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$ de la rotation infinitésimale autour de 0 est $\widetilde{A}_{(x,y)} = -y \dfrac{\partial}{\partial x} + x \dfrac{\partial}{\partial y}$.

De même, les opérateurs différentiels de $\mathbb{R}^3$ représentant les rotations infinitésimales autour des axes : O_x, O_y, O_z, base de $O(3)$ sont (cf. VI.3 et VI.5) :

$$\rho_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z},\ \rho_2 = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x},\ \rho_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}.$$

Puisque G opère à gauche sur l'algèbre des fonctions $\mathfrak{C}^\infty$ sur E par $(g \cdot f)(x) = f(\overset{-1}{g}\, x)$, l'application : $A \in \mathfrak{g} \to \widetilde{A} \in \mathfrak{X}(E)$ est une représentation linéaire, fidèle, de $\mathfrak{g}$ comme algèbre de Lie d'opérateurs différentiels sur E.

Dualement une forme différentielle ω de degré p sur un groupe de Lie G est invariante à gauche (resp. à droite) si quelque soit $g \in G$, on a $l(g)^*\omega = \omega$ (resp. $r(g)^*\omega = \omega$). Puisque la différentielle extérieure d commute avec $l(g)^*$ et $r(g)^*$, si ω est invariante à gauche, ou à droite, il en est de même de $d\omega$. Les formes différentielles invariantes à gauche sur G forment donc, avec d, une algèbre différentielle.

L'espace vectoriel des 1-formes invariantes à gauche est le dual $\mathfrak{g}^*$ de l'espace vectoriel $\mathfrak{g}$ des champs de vecteurs invariants à gauche. On définit sur G la *1-forme invariante à gauche* λ *à valeurs dans son algèbre de Lie* $\mathfrak{g}$ par : si $X_x \in T(G)_x$, $\lambda(X_x) = l(\overset{-1}{x})_* X_x$ $\in T(G)_e$ identifié avec $\mathfrak{g}$. On a :

$$(l(y)^*\lambda)(X_x) = \lambda(l(y)_* X_x) = l(\overset{-1}{yx})_* l(y)_* X_x = \lambda(X_x)$$

$$(r(y)^*\lambda)(X_x) = \lambda(r(y)_* X_x) = l(\overset{-1}{xy})_* r(y)_* X_x = \mathrm{Ad}\,\overset{-1}{y}\,\lambda(X_x)$$

soit $r(y)^*\lambda = \mathrm{Ad}\ \overset{-1}{y} \cdot \lambda$

On a de la même façon une 1-forme invariante à droite ρ sur G à valeurs dans $\mathfrak{g}$ définie par : $\rho(X_x) = r(\overset{-1}{g})_* X_x \in T(G_e) \equiv \mathfrak{g}$

Si $\alpha \in \mathfrak{g}^*$, $X, Y \in \mathfrak{g}$, les fonctions $\alpha(X)$ et $\alpha(Y)$ sont constantes, et $d\alpha(X, Y) = X \cdot \alpha(Y) - Y \cdot \alpha(X) - \alpha([X, Y]) = -\alpha([X, Y])$

Si $X_1, X_2, \ldots X_m$ forment une base de $\mathfrak{g}$ dont la base duale de $\mathfrak{g}^*$ est $\alpha^1, \alpha^2, \ldots \alpha^m$, on a $[X_i, X_j] = \sum C_{k,ij} X_k$ d'où $d\alpha^k = -\sum C_{k,ij}\alpha^i \wedge \alpha^j$.

Les formes différentielles invariantes à gauche sur G forment l'algèbre $\wedge\mathfrak{g}^*$ qui est une algèbre différentielle munie d'une $\mathfrak{g}$-structure au sens ci-dessus et d'une représentation du groupe G par $\mathfrak{g} \longrightarrow \mathrm{Ad}\ g$. Les formes invariantes à droite *et* à gauche sont les invariants de la représentation adjointe dans $\wedge g^*$. Ces formes π sont fermées : $d\pi = 0$.

Nous allons maintenant examiner comment la présence d'une métrique pseudoriemannienne g sur la variété $\mathfrak{V}$ détermine sur l'algèbre de ses formes différentielles extérieures des opérateurs exprimant la dualité : l'opérateur d'adjonction $*$ (appelé étoile, star, ou « de Hodge ») venant de l'opérateur $*$ du §VI.5, et la codifférentielle δ. Supposons $\mathfrak{V}$ orientée, et soit ω l'élément de volume, qui s'écrit sous sa forme covariante en coordonnées locales : $\omega = |\overline{g}|^{1/2} dx^1 \wedge \ldots \wedge dx^n$.

Soient α et β deux p-formes différentielles extérieures sur $\mathfrak{V}$. On peut les considérer comme des sections de l'espace fibré $\wedge^p T(\mathfrak{V})^*$. En chaque point x de $\mathfrak{V}$ est défini l'automorphisme de dualité $*$ par : $\beta_x \wedge *_x \alpha_x = (\beta_x \mid \alpha_x)\omega_x$.

Nous allons calculer l'expression locale de $*$ en supposant donné localement un champ de repères $e = (e_1, e_2, \dots e_n)$ de $\mathrm{T}(\mathfrak{V})$ dont le dual $e^* = (e^{*1}, e^{*2}, \dots e^{*n})$ est un champ de repères de $\mathrm{T}(\mathfrak{V})^*$. Avec les notations usuelles (chap. III) :
$\beta = \sum \beta_{\mathrm{H}} e^{*\mathrm{H}}$, $\alpha = \sum \alpha_{\mathrm{I}} e^{*\mathrm{I}}$, où I et H parcourent les suites $(1 \leqslant i_1 < \dots i_p \leqslant n)$

$$*\alpha = \textstyle\sum (*\alpha)_{\mathrm{J}} e^{*\mathrm{J}}, \quad \mathrm{J} = (1 \leqslant j_1 < \dots j_{n-p} \leqslant n).$$

$$\begin{aligned}\beta \wedge *\alpha &= \textstyle\sum \beta_{\mathrm{H}} (*\alpha)_{\mathrm{J}} e^{*\mathrm{H}} \wedge e^{*\mathrm{J}} = \sum \beta_{\mathrm{H}} (*\alpha)_{\mathrm{H}'} e^{*\mathrm{H}} \wedge e^{*\mathrm{H}'} \\ &= \textstyle\sum \beta_{\mathrm{H}} (*\alpha)_{\mathrm{H}'} \varepsilon(\mathrm{H}) \overline{e}^*\end{aligned}$$

tandis que $(\beta \mid \alpha) = \sum g^{\mathrm{HI}} \beta_{\mathrm{H}} \alpha_{\mathrm{I}}$

L'égalité $\beta \wedge *\alpha = (\beta \mid \alpha)\omega$ définissant $*\alpha$ en écrivant qu'elle est vraie quel que soit β, on obtient localement :

$$\boxed{(*\alpha)_{\mathrm{H}'} = \varepsilon(\mathrm{H}) |\overline{g}|^{1/2} \textstyle\sum g^{\mathrm{HI}} \alpha_{\mathrm{I}} = \varepsilon(\mathrm{H}) |\overline{g}|^{1/2} \alpha^{\mathrm{H}}},$$

ce qui prouve que $*\alpha$ dépend différentiablement de α. Le produit scalaire $(\beta \mid \alpha)$ étant une fonction numérique sur $\mathfrak{V}$, l'opérateur $*$ est ainsi défini sur $\mathfrak{V}$ et satisfait à :

$$\boxed{\beta \wedge *\alpha = (\beta \mid \alpha)\omega},$$

en coordonnées locales : $*\alpha = |\overline{g}|^{1/2} \sum \varepsilon(\mathrm{H}) \alpha^{\mathrm{H}} dx^{\mathrm{H}'}$.

Définition IX.3.F :

Sur une variété pseudoriemannienne orientée $\mathfrak{V}$, on appelle *codifférentielle* l'opérateur δ sur les formes différentielles extérieures « dual » de la différentielle extérieure. Sur les formes de degré p :
$\delta = (-1)^p \overset{-1}{*} d*$

δ diminue le degré d'une unité, et $\delta^2 = \delta\delta = 0$. Le *Laplacien* sur les formes extérieures est l'opérateur $\Delta = d\delta + \delta d$; Δ est homogène de degré zéro. Une forme α est dite *cofermée* si $\delta\alpha = 0$, *harmonique* si $\Delta\alpha = 0$.

Remarque : Sur une variété pseudoriemannienne $\mathfrak{V}$ orientée, le *produit scalaire global* de deux p-formes α, β dont l'une est à support compact est l'intégrale sur $\mathfrak{V}$ de la fonction numérique obtenue par leur produit scalaire ponctuel (cf. §IX.10) :

$$\{\alpha \mid \beta\} = \int (\alpha \mid \beta)\omega = \int \alpha \wedge *\beta = \int \beta \wedge *\alpha,$$

On a alors, si α est de degré p et β de degré $(p+1)$:

$$\{d\alpha \mid \beta\} = \int d\alpha \wedge *\beta = \int d(\alpha \wedge *\beta) - (-1)^p \int \alpha \wedge d * \beta.$$

D'après le théorème de Stokes, l'intégrale de la différentielle extérieure d'une forme à support compact est nulle. On a donc :

$$\{d\alpha \mid \beta\} = (-1)^{p+1} \int (\alpha \mid \overset{-1}{*}\, d * \beta)\omega = \{\alpha \mid \delta\beta\}$$

d et δ sont donc des opérateurs adjoints relativement au produit scalaire global. Il en résulte que le Laplacien est son propre adjoint :

$$\begin{aligned}\{\Delta\alpha \mid \beta\} &= \{d\delta\alpha \mid \beta\} + \{\delta d\alpha \mid \beta\} = \{\delta\alpha \mid \delta\beta\} + \{d\alpha \mid d\beta\} \\ &= \{\alpha \mid \Delta\beta\}\end{aligned}$$

Sur une variété $\mathfrak{V}$ *riemannienne* orientée compacte, on a donc $\{\Delta\alpha \mid \alpha\} \geqslant 0$. Dans ce cas, si α est harmonique, $\Delta\alpha = 0$ impose $\{\delta\alpha \mid \delta\alpha\} = 0$ et $\{d\alpha \mid d\alpha\} = 0$ soit $d\alpha = 0$ et $\delta\alpha = 0$: α est à la fois fermée et cofermée.

Application : divergence d'un champ de vecteurs X. On a vu (Définition IX.3.D) que :

$$(\mathrm{div}\ \mathrm{X})\omega = di(\mathrm{X})\omega = (-1)^{n-1} d(\mathrm{X} \lrcorner\ \omega).$$

Si $\zeta = \overset{-1}{\rho}\,\mathrm{X}$ est la forme différentielle correspondant à X, on a vu (Définition VI.5.A) que $\mathrm{X} \lrcorner\ \omega = \zeta \lrcorner\ \omega$ et que, puisque $\dim \zeta = 1$, $*\zeta = (-1)^{n-1}\zeta \lrcorner\ \omega$. On obtient donc : $(\mathrm{div}\ \mathrm{X})\omega = d{*}\zeta$, en appliquant aux deux membres $\overset{-1}{*}$: $\mathrm{div}\ \mathrm{X} = \overset{-1}{*}\, d * \zeta = -\delta\zeta$. On retrouve l'expression du §VII.10.

$$\boxed{\mathrm{div}\ \mathrm{X} = -\delta\zeta}.$$

Si f est une fonction numérique différentiable sur $\mathfrak{V}$, la forme df a pour représentant contravariant le champ de vecteurs $\overset{-1}{\rho}\, df =$ grad f. Si Δ est le laplacien sur les formes :

$$\Delta f = d\delta f + \delta df = \delta df = -\mathrm{div} \cdot \mathrm{grad}\ f$$

et l'on constate qu'il est l'opposé du Laplacien usuel.

Remarque : on note une certaine perversité de la codifférentielle. Le signe intervenant dans sa définition (IX.3.E), choisi pour assurer la dualité $\{d\alpha \mid \beta\} = \{\alpha \mid \delta\beta\}$ entraîne l'apparition de signes $(-)$ dans la comparaison avec la divergence et le laplacien usuel! il faut s'y résigner (certains ouvrages préfèrent l'erreur de signe) et considérer indépendamment de δ une notion plus générale de divergence, ce qui sera fait à la fin du paragraphe suivant IX.4.

IX.4 – Dérivation covariante et transport parallèle. Connexions

Rien dans la structure de variété différentielle ne semble déterminer une comparaison des espaces tangents en des points distincts, par un « transport parallèle » analogue à celui qui permet, dans $\mathbb{R}^n$, de définir la dérivation des champs de vecteurs.

Il est en effet, nécessaire pour cela, d'ajouter à la structure de variété différentielle une structure supplémentaire qui est indifféremment appelée *connexion sur le fibré vectoriel tangent* parce qu'elle relie entre eux les espaces tangents en des points distincts, ou *dérivation covariante*, parce qu'elle est équivalente à la possibilité de dériver les champs de vecteurs tangents et plus généralement les champs de tenseurs. A ces deux aspects correspondent deux façons équivalentes, de définir cette structure. Nous verrons d'abord la seconde qui est la plus simple. Bien entendu, il suffit de définir une telle dérivation sur les champs de vecteurs, l'extension aux champs de tenseurs ou aux formes extérieures s'effectuant de façon automatique suivant les principes énoncés au § IX.3.

Définition IX.4.A :

Une dérivation covariante en un point x d'une variété différentielle $\mathfrak{V}$ de classe $\geqslant 2$ est une fonction D qui fait correspondre à tout vecteur tangent A en x, et à tout champ de vecteurs tangents différentiables Y défini au voisinage de x, un vecteur tangent $\mathrm{D_A Y}$ en x appelé dérivée covariante de Y par A de telle sorte que les axiomes suivants soient satisfaits :

1) $\mathrm{D_A Y}$ est $\mathbb{R}$-linéaire en A et en Y

2) quelle que soit la fonction numérique différentiable f sur $\mathfrak{V}$.

$$\mathrm{D_A}(f\mathrm{Y}) = (\mathrm{A} \cdot f)\mathrm{Y}_x + f(x)\mathrm{D_A} \cdot \mathrm{Y}.$$

Une *dérivation covariante sur* $\mathfrak{V}$ est la donnée d'une dérivation covariante en chaque point de $\mathfrak{V}$ astreinte à la condition que quels que soient les champs de vecteurs X et Y de classe s, le champ de vecteurs $\mathrm{D_X Y}$ soit de classe $(s-1)$. Le champ de tenseurs d'ordre deux de variance $(.*)$ défini par : $\mathrm{X} \to \mathrm{D_X Y}$ est appelé *différentielle covariante* de Y et noté DY. C'est donc une section de $\mathrm{T} \otimes \mathrm{T}^*$.

La condition 2) s'écrit $\mathrm{D_X}(f \cdot \mathrm{Y}) = (\mathrm{X} \cdot f)\mathrm{Y} + f\mathrm{D_X Y}$ ou encore $\mathrm{D}(f\mathrm{Y}) = df \otimes \mathrm{Y} + f\mathrm{DY}$ et entraîne que D est de caractère local (cf. IX.3).

Analysons localement l'opération de dérivation sur un ouvert de $\mathfrak{V}$ possédant des coordonnées $x^1 x^2, \dots x^n$. Pour chaque $i = 1, 2, \dots, n$, $\mathrm{D}_{\frac{\partial}{\partial x^i}}$ est un opérateur linéaire de l'espace tangent en chaque point, dont la matrice, notée (Γ^k_{ij}) permet d'écrire la dérivation sous la forme :

$$\mathrm{D}_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j}\right) = \sum \Gamma^k_{ij} \, \frac{\partial}{\partial x^k} \quad (\Gamma^k_{ij} \text{ fonctions numériques}).$$

Si $\mathrm{X} = \sum \mathrm{X}^i \, \frac{\partial}{\partial x^i}$ et $\mathrm{Y} = \sum \mathrm{Y}^j \, \frac{\partial}{\partial x^j}$:

$$\mathrm{D_X Y} = \sum \mathrm{X}^i \frac{\partial \mathrm{Y}^j}{\partial x^i} \, \frac{\partial}{\partial x^j} + \sum \mathrm{X}^i \mathrm{Y}^j \Gamma^k_{ij} \frac{\partial}{\partial x^k}$$

$\omega = \sum \Gamma^k_{ij} \, \frac{\partial}{\partial x^k} \otimes dx^i \otimes dx^j$ est la forme de la connexion dans le système de coordonnées.

Si $\mathfrak{V}$ est de classe $\mathfrak{C}_r$, les $\frac{\partial}{\partial x^k}$ sont de classe $(r-1)$. Pour que leurs dérivées soient de classe $(r-2)$, il est nécessaire et suffisant que les Γ^k_{ij} soient de classe $\geqslant (r-2)$.

Les n^3 fonctions Γ^k_{ij}, liées aux coordonnées locales choisies déterminent ainsi complètement la dérivation sur *l'ouvert* U. On peut les fixer *arbitrairement*, pourvu qu'elles soient de classe $(r-2)$, et la formule ci-dessus définit alors une dérivation covariante sur U, comme on le vérifie aisément.

On peut réécrire l'expression trouvée pour $\mathrm{D_X Y}$ sous la forme :

$$\boxed{\mathrm{D_X Y} = \sum \mathrm{X}^i \cdot \left(\frac{\partial \mathrm{Y}^k}{\partial x^i} + \sum \Gamma^k_{ij}\mathrm{Y}^j\right) \frac{\partial}{\partial x^k}}.$$

La k-ème composante de la dérivée $\mathrm{D}_{\frac{\partial}{\partial x_i}} \mathrm{Y}$ se note (la barre verticale devant le i rappelant la dérivation)

$$\mathrm{Y}^k_{|i} = \frac{\partial \mathrm{Y}^k}{\partial x^i} + \sum \Gamma^k_{ij}\mathrm{Y}^j \text{ soit } \boxed{\mathrm{D}_{\frac{\partial}{\partial x_i}} \mathrm{Y} = \sum \mathrm{Y}^k_{|i} \frac{\partial}{\partial x^k}}.$$

Les $\mathrm{Y}^k_{|i}$ *sont les composantes du champ de tenseurs* DY *:*

$$\mathrm{DY} = \sum \mathrm{Y}^k_{|i} \frac{\partial}{\partial x^k} \otimes dx^i \text{ et } \mathrm{D_X Y} = \sum \dot{\mathrm{X}}^i \mathrm{Y}^k_{|i} \frac{\partial}{\partial x^k}$$

Exemples : La dérivation ordinaire de $\mathbb{R}^n$ correspond dans le systéme de coordonnées standard, à des fonctions Γ^k_{ij} identiquement nulles, *mais qui ne le sont plus en coordonnées curvilignes.* Dans la dérivation covariante des courbes, surfaces et sous-variétés des espaces euclidiens, les Γ^k_{ij} sont les symboles de Christoffel (§VII.8 et 9).

La dérivation dans $\mathbb{R}^n$ peut s'écrire $\mathrm{DY} = d\mathrm{Y}$, soit $\mathrm{D_X Y} = d\mathrm{Y}(\mathrm{X})$, et on donc : $\mathrm{D_X Y} - \mathrm{D_Y X} = [\mathrm{X}, \mathrm{Y}]$ ce qui traduit le fait que la connexion de $\mathbb{R}^n$ est « symétrique » suivant la :

Définition IX.4.B :

Le critère de tensorialité (Corollaire IX.2) permet de vérifier que l'application T :

$$(\mathrm{X}, \mathrm{Y}) \to \mathrm{T}(\mathrm{X}, \mathrm{Y}) = \mathrm{D_X Y} - \mathrm{D_Y X} - [\mathrm{X}, \mathrm{Y}]$$

définit un champ de tenseurs de variance $(.**)$ appelé *torsion* de la connexion. T est une 2-forme extérieure « *à valeurs dans l'espace* *fibré tangent* $\mathrm{T}(\mathfrak{V})$ ». Sa signification sera de nouveau examinée au §IX.6. La connexion est dite *symétrique* si sa torsion est nulle.

Exprimée dans les coordonnées locales, cette condition s'écrit : $\mathrm{D}_{\frac{\partial}{\partial x_i}}\left(\frac{\partial}{\partial x_j}\right) = \mathrm{D}_{\frac{\partial}{\partial x_j}}\left(\frac{\partial}{\partial x_i}\right)$ et se traduit par la symétrie : $\Gamma^k_{ij} = \Gamma^k_{ji}$.

Dire, que la connexion est symétrique signifie que l'opération crochet de deux champs de vecteurs $[\mathrm{X}, \mathrm{Y}]$ peut s'exprimer à l'aide de la dérivation covariante comme antisymétrisé de $\mathrm{D}_{\mathrm{X}}\mathrm{Y}$. Nous verrons de la même façon que la différentielle extérieure peut alors s'exprimer comme antisymétrisée de la différentielle covariante.

La première question qui se pose est celle de la possibilité de définir une dérivation covariante sur une variété différentielle donnée.

Théorème IX.4.A : On peut définir une connexion, (ou dérivation covariante) sur toute variété différentielle $\mathfrak{V}$, de classe $\geqslant 2$ paracompacte, et même une connexion symétrique.

Preuve : c'est dans les théorèmes d'existence de structures globales sur les variétés différentielles qu'intervient la paracompacité, qui s'exprime ainsi : $\mathfrak{V}$ est *paracompacte* si elle est séparée et si pour tout recouvrement ouvert $(\mathrm{U}_i, i \in \mathrm{I})$ de $\mathfrak{V}$, il existe un recouvrement ouvert *localement fini* $(\mathrm{V}_j,\ j \in \mathrm{J})$ tel que chaque V_j soit contenu dans un U_i. Il existe alors une partition $\mathfrak{C}_r$ de l'unité $\{f_j; j \in \mathrm{J}\}$ où les f_j sont des fonctions numériques $\mathfrak{C}_r$ prenant leurs valeurs dans l'intervalle $[0, 1]$ de $\mathbb{R}$, telles que :

1) l'adhérence F_j de l'ensemble des points où f_j est non nulle (support de f_j) et contenue dans V_j ;

2) la somme (localement finie) : $\sum f_j = 1$.

Les partitions de l'unité servent à « coller » des morceaux de structure locale pour obtenir une structure globale. C'est ce que l'on va faire en supposant choisies, sur chaque ouvert U_α d'un atlas $\{\mathrm{U}_\alpha, \alpha \in \mathfrak{A}\}$, n^3 fonctions $\mathfrak{C}_r$: $\Gamma^{(\alpha)k}_{ij}$ déterminant une dérivation covariante sur U_α. On prend alors un recouvrement localement fini $(\mathrm{V}_\beta,\ \beta \in \mathfrak{B})$ tel que chaque V_β soit contenu dans un U_α. Les restrictions $\Gamma^{(\beta)k}_{ij}$ à V_β des fonctions $\Gamma^{(\alpha)k}_{ij}$, définissent sur V_β une dérivation covariante D^β.

Une partition $\mathfrak{C}_r$ de l'unité : $(f_\beta,\ \beta \in \mathfrak{B})$ subordonnée au recouvrement localement fini $(\mathrm{V}_\beta,\ \beta \in \mathfrak{B})$ permet alors de définir sur $\mathfrak{V}$ tout entière la dérivation covariante $\mathrm{D} = \sum f_\beta \mathrm{D}^\beta$, soit en un point x :

$$\mathrm{D}_\mathrm{X}\mathrm{Y} = \sum f_\beta(x)\mathrm{D}^\beta_\mathrm{X}\mathrm{Y} = \sum \mathrm{D}^\beta_{f_\beta(x)\mathrm{X}}\mathrm{Y}.$$

On a bien, puisque $\sum f_\beta = 1$:

$$\begin{aligned}\mathrm{D}_\mathrm{X}(f\mathrm{Y}) &= (\sum_\beta f_\beta(x))(\mathrm{X}\cdot f)\mathrm{Y} + f(x)\sum f_\beta(x)\mathrm{D}^\beta_\mathrm{X}\mathrm{Y}\\ &= (\mathrm{X}\cdot f)\mathrm{Y} + f(x)\mathrm{D}_\mathrm{X}\mathrm{Y}.\end{aligned}$$

Remarque : il n'est pas inutile de mesurer l'arbitraire du choix d'une connexion sur $\mathfrak{V}$. Si D et D′ sont deux dérivations covariantes, la fonction :

$$u(\mathrm{X},\mathrm{Y}) = \mathrm{D}'_\mathrm{X}\mathrm{Y} - \mathrm{D}_\mathrm{X}\mathrm{Y}$$

fait correspondre à 2 champs de vecteurs X,Y un champ de vecteurs $u(\mathrm{X},\mathrm{Y})$. Si f et g sont deux fonctions différentiables, on a $u(f\mathrm{X}, g\mathrm{Y}) = fgu(\mathrm{X},\mathrm{Y})$.

Il en résulte, d'après le Corollaire IX.2, que u est un champ de tenseurs de classe $\geqslant r-2$, de variance $(.\,**)$. Réciproquement, si u est un tel champ de tenseurs et D une dérivation covariante $\mathrm{D}+u$ en est également une.

Si D et u sont symétriques. $\mathrm{D}+u$ l'est également.

Si D′ est définie par : $\mathrm{D}'_\mathrm{X}\mathrm{Y} = \mathrm{D}_\mathrm{X}\mathrm{Y} - \dfrac{1}{2}\,\mathrm{T}(\mathrm{X},\mathrm{Y})$, D′ est symétrique, et D *s'écrit de façon unique :* $\mathrm{D} = \mathrm{D}' + u$ *avec* D *symétrique et* u *antisymétrique.*

Théorème IX.4.B : On peut définir une structure riemannienne sur toute variété différentielle paracompacte.

Preuve : Elle est analogue à celle du théorème précédent, en remarquant que si $g_1, g_2, \ldots g_p$ sont p formes quadratiques positives non-dégénérées sur un espace vectoriel réel E, et $\beta_1, \beta_2, \ldots \beta_p$, p nombres réels positifs de somme $\sum \beta_j = 1$, $\beta_1 g_1 + \ldots \beta_p g_p$ est encore une forme quadratique positive non-dégénérée sur E (propriété de convexité).

Théorème IX.4.C : Sur toute variété riemannienne, ou pseudo-riemannienne, il existe une dérivation covariante unique, appelée *connexion riemannienne, ou pseudoriemannienne*, satisfaisant aux deux conditions suivantes :

1) elle est symétrique (Définition IX.4.B)

2) elle respecte le produit scalaire, ce qui se traduit, comme on le verra ci-après par l'identité :

$$\mathrm{X} \cdot (\mathrm{Y} \mid \mathrm{Z}) = (\mathrm{D_X Y} \mid \mathrm{Z}) + (\mathrm{Y} \mid \mathrm{D_X Z})$$

vérifiée quels que soient les champs de vecteurs X, Y, Z.

Preuve : Avant de prouver l'existence de D, montrons que les conditions 1) et 2) la déterminent de façon unique.

Soit g le champ de tenseurs définissant le produit scalaire sur l'espace tangent : $g(\mathrm{X}, \mathrm{Y}) = (\mathrm{X} \mid \mathrm{Y})$. Dans des coordonnées locales, on a $g = \sum g_{ij} dx^i \otimes dx^j$ avec $g_{ij} = \left(\frac{\partial}{\partial x^i} \mid \frac{\partial}{\partial x^j}\right) = g_{ji}$.

Notons, pour simplifier $\mathrm{D}_{ij} = \mathrm{D}_{\frac{\partial}{\partial x^i}}\left(\frac{\partial}{\partial x_j}\right)$ les champs de vecteurs locaux définissant localement la dérivation covariante D. On doit avoir d'après 1) : $\mathrm{D}_{ij} = \mathrm{D}_{ji}$. La condition 2) donne sur les champs de base $\frac{\partial}{\partial x^i}$ par permutation circulaire de (ijk)

$$\begin{aligned}
\frac{\partial}{\partial x^i}\, g_{jk} &= \left(\mathrm{D}_{ij} \mid \frac{\partial}{\partial x^k}\right) + \left(\frac{\partial}{\partial x^j} \mid \mathrm{D}_{ik}\right)\\
\frac{\partial}{\partial x^j}\, g_{ki} &= \left(\mathrm{D}_{jk} \mid \frac{\partial}{\partial x^i}\right) + \left(\frac{\partial}{\partial x^k} \mid \mathrm{D}_{ji}\right)\\
\frac{\partial}{\partial x^k}\, g_{ij} &= \left(\mathrm{D}_{ki} \mid \frac{\partial}{\partial x^j}\right) + \left(\frac{\partial}{\partial x^i} \mid \mathrm{D}_{kj}\right)
\end{aligned}$$

La combinaison : (1$^{\text{re}}$ ligne) + (2$^{\text{e}}$ ligne) − (3$^{\text{e}}$ ligne) donne :

$$\left(\mathrm{D}_{ij} \mid \frac{\partial}{\partial x^k}\right) = \frac{1}{2}\left(\frac{\partial}{\partial x^i}\, g_{jk} + \frac{\partial}{\partial x^j}\, g_{ik} - \frac{\partial}{\partial x^k}\, g_{ij}\right).$$

Or $\left(\mathrm{D}_{ij} \mid \frac{\partial}{\partial x^k}\right) = \sum \Gamma^l_{ij} g_{lk} = \Gamma_{ijk}$ et on obtient Γ^k_{ij} à l'aide de la matrice inverse de $|g_{ij}|$ soit $|g^{ij}|$:

$$\Gamma^k_{ij} = \sum g^{km}\Gamma_{ijm} = \frac{1}{2}\sum g^{km}\left(\frac{\partial}{\partial x^i}\, g_{jm} + \frac{\partial}{\partial x^j}\, g_{im} - \frac{\partial}{\partial x^m}\, g_{ij}\right).$$

En remontant les calculs, on vérifie réciproquement que la connexion définie par les Γ_{ij}^k, évidemment symétrique, respecte le produit scalaire. Comme les conditions 1) et 2) sont intrinsèques et que la connexion trouvée localement est unique, on obtient une connexion unique sur $\mathfrak{V}$ tout entière.

Exemple : La dérivation covariante des sous-variétés de l'espace euclidien E_n, définie au §VII.9.

Nous revenons *au cas général* d'une dérivation covariante D définie sur une variété différentielle nue, donc en dehors de toute métrique.

Nous allons construire le « transport parallèle » dont nous avons parlé au début de ce § par « intégration » de la structure de dérivation covariante.

Soient c un chemin continûment différentiable dans $\mathfrak{V}$, x^1, x^2, $\ldots$ x^n des coordonnées locales au voisinage de $x = c(t)$.

Si $\mathrm{Y} = \sum \mathrm{Y}^j \dfrac{\partial}{\partial x^j}$ est un champ de vecteurs défini au voisinage de x, on a :

$$\boxed{\mathrm{D}_{\frac{dc}{dt}}\, \mathrm{Y} = \sum \frac{d\mathrm{Y}^j(c(t))}{dt} \frac{\partial}{\partial x^j} + \sum \left(\frac{dc}{dt}\right)^i \mathrm{Y}^j(c(t))\Gamma_{ij}^k \frac{\partial}{\partial x^k}}\,.$$

On constate que la dérivée du champ Y par rapport à un vecteur $\mathrm{X} = \dfrac{dc}{dt}$ est entièrement déterminée par la restriction du champ Y à un chemin différentiable $c(t)$ ayant le vecteur X pour vecteur tangent.

On constate également sur cette égalité que l'opération de dérivation covariante permet de définir la dérivée d'un champ de vecteurs différentiable défini *seulement* le long de c. De façon précise, un tel champ est une application continûment différentiable de l'intervalle de définition de c à valeurs dans l'espace tangent $\mathrm{T}(\mathfrak{V})$, telle que, $\forall t$, $\mathrm{V}(t) \in \mathrm{T}(\mathfrak{V})_{c(t)}$. On note alors simplement $\dfrac{\mathrm{DV}}{dt}$ la dérivée covariante $\mathrm{D}_{\frac{dc}{dt}}\, \mathrm{V}(t)$.

Plus généralement, au lieu de rapporter les champs de vecteurs tangents au champ de repères $\left(\dfrac{\partial}{\partial x^1}, \ldots \dfrac{\partial}{\partial x^n}\right)$ défini par des coordonnées locales, on peut supposer que l'on a sur un ouvert U un champ de repères $e = (e_1, e_2, \dots e_n)$ formé de n champs de vecteurs sur U, linéairement indépendants en chaque point de U. L'expression de la dérivation covariante dans e est analogue à la précédente.

On définit les fonctions Γ_{ij}^k par : $\mathrm{D}_{e_i} e_j = \sum \Gamma_{ij}^k e_k$ d'où, si $\mathrm{X} = \sum \mathrm{X}^i e_i$ et $\mathrm{Y} = \sum \mathrm{Y}^j e_j$: $\mathrm{D_X Y} = \sum \mathrm{X}^i (e_i \cdot \mathrm{Y}^j) e_j + \sum \mathrm{X}^i \mathrm{Y}^j \Gamma_{ij}^k e_k$.

Définition IX.4.B : Transport parallèle et géodésique

Un champ de vecteurs V de classe $s \geqslant 1$ le long d'un chemin différentiable $c(t)$, dans une variété différentielle $\mathfrak{V}$ munie d'une dérivation covariante D, est dit *parallèle le long de c* si sa dérivée covariante $\dfrac{\mathrm{DV}}{dt}$ est identiquement nulle. Un chemin différentiable $c(t)$ est appelé une *géodésique* de $(\mathfrak{V}, \mathrm{D})$ si le champ de ses vecteurs tangents $\dfrac{dc}{dt}$ le long de c est parallèle.

Le système d'équations différentielles traduisant localement le parallélisme de V, soit $\dfrac{\mathrm{DV}}{dt} = 0$ s'écrit :

(P)

$$\boxed{\frac{d\mathrm{V}(t)^k}{dt} + \sum \left(\frac{dc}{dt}\right)^i \mathrm{V}^j(t) \Gamma_{ij}^k = 0} \quad k = 1, 2, \dots n \textit{ (parallélisme)}.$$

L'équation différentielle des *géodésiques* $\dfrac{\mathrm{D}}{dt}\left(\dfrac{dc}{dt}\right) = 0$ signifie que le *vecteur accélération* $\dfrac{\mathrm{D}}{dt}\left(\dfrac{dc}{dt}\right)$ du chemin $c(t)$ est nul. Cette équation s'écrit localement sous la forme du système d'équations différentielles :

$$\text{(G)} \quad \boxed{\frac{d^2 c^k}{dt^2} + \sum \frac{dc^i}{dt}\frac{dc^j}{dt} \Gamma_{ij}^k = 0} \quad k = 1, 2, \dots n \quad \textit{(géodésiques)}.$$

Théorème IX.4.D : Soient $\mathfrak{V}$ une variété $\mathfrak{C}_r$ $(r \geqslant 2)$ munie d'une dérivation covariante D, c un chemin continu et continûment différentiable par morceaux : $t \in [0, a]$ $(a > 0$ ou $a = +\infty) \to c(t)$: et V_0 un vecteur tangent de $\mathfrak{V}$ au point $c(0)$. Il existe alors un champ de vecteurs parallèle le long de c : $\mathrm{V}(t)$, tel que $\mathrm{V}(0) = \mathrm{V}_0$, défini pour toutes les valeurs de t. Ce champ de vecteurs, appelé le *transport parallèle de* V_0 *le long de* c est unique.

Preuve : Le théorème n'est qu'une reformulation géométrique du théorème classique d'existence, unicité, et prolongement de la solution pour toutes les valeurs de t, pour un système différentiel linéaire homogène tel que (P).

Soient $V_1(t)$, $V_2(t)$, ..., $V_n(t)$ n solutions du système différentiel linéaire homogène (P) qui s'écrit sous forme matricielle :

$$\frac{dV}{dt} = A(t)V \quad A(t) : \text{ matrice } (n \times n).$$

Soit $\Phi(t)$ la matrice $n \times n$ de leurs composantes : $|\Phi_p^k| = |V_p(t)^k|$. $\Phi(t)$ vérifie l'équation matricielle : $\Phi'(t) = A(t)\Phi(t)$ d'où l'on déduit par un calcul classique :

$$[\text{dét } \Phi(t)]' = \operatorname{Tr} A(t) \text{ dét } \Phi(t)$$

soit

$$\text{dét } \Phi(t) = \text{dét } \Phi(t_0) \exp \int_{t_0}^{t} \operatorname{Tr} A(t) \cdot dt.$$

Il en résulte que, si $V_1(t_0), V_2(t_0), \ldots V_n(t_0)$ sont linéairement indépendant, formant une base de $T(\mathfrak{V})_{c(t_0)}$: dét $\Phi(t_0) \neq 0$, $V_1(t), V_2(t), \ldots V_n(t)$ forment alors une base de $T(\mathfrak{V})_{c(t)}$quel que soit t, car dét $\Phi(t) \neq 0$ d'où :

Théorème IX.4.E :

Le transport parallèle τ_{tt_0} de $c(t_0)$ à $c(t)$ le long du chemin c continu et continûment différentiable par morceaux détermine un isomorphisme linéaire de l'espace tangent $T(\mathfrak{V})_{c(t_0)}$ sur l'espace tangent $T(\mathfrak{V})_{c(t)}$ quel que soit t.

Soient maintenant $e_1, e_2, \ldots e_n$ une base de $T(\mathfrak{V})_{c(t_0)}$, $e_1(t), \ldots$ $e_n(t)$ les transports parallèles de $e_1, \ldots, e_n$ le long de $c(t)$. Il résulte du théorème ci-dessus que $e_1(t), \ldots e_n(t)$ forment une base de $T(\mathfrak{V})_{c(t)}$. Pour un champ V quelconque le long de c :

$$V(t) = \sum V^j(t) e_j(t)$$

$$\frac{DV}{dt} = \sum \frac{dV^j}{dt} e_j(t) \text{ puisque } \frac{De_j}{dt} \equiv 0, \ \forall j.$$

Le transport parallèle de $V(t)$ en $V(t_0)$ s'écrit donc :

$$\overset{-1}{\tau}{}_{tt_0}V(t) = \tau_{t_0t}V(t) = \sum_j V^j(t)e_j(t_0)$$

d'où : $$\boxed{\left(\frac{DV}{dt}\right)_{t_0} = \lim_{t\to t_0} \frac{1}{t-t_0}\,\{\tau_{t_0t}V(t) - V(t_0)\}},$$

ce qui *détermine réciproquement la dérivation covariante à partir du transport parallèle.*

Remarque : Si l'on transporte parallèlement un vecteur le long d'un chemin fermé, même « petit », le vecteur ne revient pas coïncider avec le vecteur de départ.

Cet effet traduit la « *courbure* » *de l'espace* $(\mathfrak{V}, D)$. Cette courbure va se manifester au moyen d'un champ de tenseurs d'ordre quatre, le *tenseur de courbure*, expression concentrée de la géométrie de la variété $\mathfrak{V}$ munie de D.

Corollaire IX.4 : Le transport parallèle associé à la connexion canonique d'une variété riemannienne ou pseudoriemannienne *conserve le produit scalaire.*

Preuve : En effet, de par la définition de cette connexion, la dérivée du produit scalaire $(V(t) \mid W(t))$ de deux champs de vecteurs *parallèles* le long du chemin $c(t)$ s'écrit :

$$\begin{aligned}\frac{d}{dt}\,(V(t) \mid W(t)) &= \frac{dc}{dt}\cdot(V(t) \mid W(t))\\ &= \left(\frac{DV}{dt} \mid W(t)\right) + \left(V(t)) \mid \frac{DW}{dt}\right) = 0.\end{aligned}$$

Il en résulte en particulier que le vecteur vitesse $\dfrac{dc}{dt}$ d'une géodésique garde une « longueur » constante puisque :

$$\frac{d}{dt}\left(\frac{dc}{dt}\Big|\frac{dc}{dt}\right) = 2\left(D_{\frac{dc}{dt}}\frac{dc}{dt}\Big|\frac{dc}{dt}\right) = 0$$

Le transport parallèle le long du chemin fermé γ d'origine et extrêmité x_0 détermine un automorphisme linéaire τ_γ de l'espace tangent T_{x_0}, orthogonal si la variété est riemannienne ou pseudoriemannienne.

Si l'on parcourt successivement deux tels chemin γ et γ', on a : $\tau_{\gamma' \circ \gamma} = \tau_{\gamma'} \circ \tau_\gamma$ tandis que $\tau_{\overset{-1}{\gamma}} = \overset{-1}{\underset{\gamma}{\tau}}$.

Les automorphismes linéaires de T_{x_0} déterminés par le transport parallèle le long des chemins fermés forment ainsi un groupe : le *groupe d'holonomie* en x_0, sous-groupe de $\mathrm{Gl}(T_{x_0})$, de $\mathrm{O}(T_{x_0})$ si $\mathfrak{V}$ est riemannienne ou pseudoriemannienne).

Les groupes d'holonomie aux différents points d'une variété connexe sont isomorphes.

Le transport parallèle τ_{tt_0} le long du chemin c, de $c(t_0)$ à $c(t)$ étant un isomorphisme de $T(\mathfrak{V})_{c(t_0)}$ sur $T(\mathfrak{V})_{c(t)}$ détermine un *isomorphisme contragrédient* $\check{\tau}_{tt_0}$ de $T(\mathfrak{V})^*_{c(t_0)}$ sur $T(\mathfrak{V})^*_{c(t)}$ et un *isomorphisme* $\overset{(v)}{\tau}_{tt_0}$ de la puissance tensorielle $\otimes^{(v)}T(\mathfrak{V})_{c(t_0)}$ sur $\otimes^{(v)}T(\mathfrak{V})_{c(t)}$.

Pour définir la dérivée covariante $D_X u$ d'un champ de tenseurs différentiable u, de variance (v), relativement à un vecteur tangent X en un point x de $\mathfrak{V}$, il suffit de construire un chemin différentiable $c(t)$, $t \in [0, a]$ tel que $c(0) = x$ et $\left(\frac{dc}{dt}\right)_0 = X$, et de poser :

$$D_X u = \lim_{t \to 0} \frac{1}{t}(\overset{(v)}{\tau}_{t_0 t} u_{c(t)} - u_x).$$

Il faut alors démontrer que $D_X u$ ne dépend pas du chemin $c(t)$ choisi. Ce fait est une conséquence des propriétés algébriques de cette dérivation dont la démonstration est immédiate :

1) si u et u' sont de même variance (v), $D_X(\lambda u + \lambda' u') = \lambda D_X u + \lambda' D_X u'$, $\forall \lambda, \lambda' \in \mathbb{R}$;

2) lorsque u est une fonction numérique , $D_X f = X \cdot f$, soit $Df = df$;

3) lorsque u est un champ de vecteurs Y, la dérivée ainsi définie coïncide avec la dérivée covariante initiale $D_X Y$;

4) $D_X(u \otimes u') = D_X u \otimes u' + u \otimes D_X u'$;

5) D_X commute avec les contractions : $D_X(c^i_j u) = c^i_j D_X u$.

De ces propriétés algébriques; il résulte que, la dérivation D_X en un point, des champs de tenseurs, est entièrement déterminée par son action sur les fonctions numériques et les champs de vecteurs et ne dépend donc que de X.

L'opérateur de différentiation covariante s'exprime localement par $\mathrm{D} = \sum dx^i \otimes \mathrm{D}\dfrac{\partial}{\partial x^i}$.

La différentielle $\mathrm{D}u$ d'un champ de tenseurs u de variance (p, q) est le champ de tenseurs de variance $(p, q+1)$ que définit la fonction qui à chaque champ de vecteurs X associe le champ de tenseurs $\mathrm{D}_\mathrm{X}u$.

Remarquons que la propriété 4) entraîne que la dérivation D_X commute avec les opérateurs de permutation. En particulier si u est un champ de tenseurs symétriques (ou antisymétriques), et X un vecteur tangent en un point x, $\mathrm{D}_\mathrm{X}u$ est un tenseur symétrique (ou antisymétrique) en x. Cette propriété *n'est pas* partagée par la différentielle D, qui ajoute un indice covariant.

En utilisant les propriétés ci-dessus de la dérivée D_X on calcule la dérivée covariante d'un champ de covecteurs α en contractant :

$$\mathrm{D}_\mathrm{X}(\alpha \otimes \mathrm{Y}) = \mathrm{D}_\mathrm{X}\alpha \otimes \mathrm{Y} + \alpha \otimes \mathrm{D}_\mathrm{X}\mathrm{Y}$$

soit : $\langle \mathrm{D}_\mathrm{X}\alpha, \mathrm{Y}\rangle = \mathrm{X}.\ \langle \alpha, \mathrm{Y}\rangle - \langle \alpha, \mathrm{D}_\mathrm{X}\mathrm{Y}\rangle$ où $\langle\quad\rangle$ est le crochet de dualité entre champs de covecteurs et de vecteurs.

Applications : 1) Le *Hessien* $(\mathrm{H}f)$ d'une fonction numérique f de classe $\mathfrak{C}_2$ sur une variété munie d'une connexion est le champ de tenseurs deux fois covariants : $\mathrm{H}f = \mathrm{D}^2 f = \mathrm{D}(df)$, soit, si X et Y sont deux champs de vecteurs :

$$\begin{aligned}(\mathrm{H}f)(\mathrm{X}, \mathrm{Y}) &= (\mathrm{D}_\mathrm{X}(df))(\mathrm{Y}) = \mathrm{X} \cdot df(\mathrm{Y}) - df(\mathrm{D}_\mathrm{X}\mathrm{Y}) \\ &= \mathrm{X} \cdot (\mathrm{Y} \cdot f) - (\mathrm{D}_\mathrm{X}\mathrm{Y}) \cdot f\end{aligned}$$

que l'on peut aussi écrire :

$$(\mathrm{H}f)(\mathrm{X}, \mathrm{Y}) = (\mathrm{D}_\mathrm{X}\mathrm{D}_\mathrm{Y} - \mathrm{D}_{\mathrm{D}_\mathrm{X}\mathrm{Y}}) \cdot f.$$

Si la connexion est symétrique, $\mathrm{H}f$ est symétrique. En effet :

$$\begin{aligned}(\mathrm{H}f)(\mathrm{X}, \mathrm{Y}) - (\mathrm{H}f)(\mathrm{Y}, \mathrm{X}) &= \mathrm{X} \cdot (\mathrm{Y} \cdot f) - \mathrm{Y}(\mathrm{X} \cdot f) \\ &\quad - \{\mathrm{D}_\mathrm{X}\mathrm{Y} - \mathrm{D}_\mathrm{Y}\mathrm{X}\} \cdot f = -\mathrm{T}(\mathrm{X}, \mathrm{Y}) \cdot f = 0.\end{aligned}$$

(ou T est le tenseur de torsion)

Remarque : On définit parfois le hessien d'une fonction f sur une variété différentielle nue (sans connexion) par la fonction $\Phi(\mathrm{X}, \mathrm{Y}) = \mathrm{X} \cdot (\mathrm{Y} \cdot f)$, qui n'est évidemment pas un champ de tenseurs, n'étant pas $\mathfrak{F}$-linéaire en Y, et, qui n'est symétrique en X et Y qu'aux points x de la variété où f est plate $(df_x = 0)$.

2) Plus généralement, le champ d'opérateurs $D_{X,Y}$ associé à deux champs de vecteurs X, Y défini par : $D_{X,Y} = D_X D_Y - D_{D_X Y}$ est appelé l'*opérateur hessien*, ou de dérivée seconde covariante relativement à X et Y.

Il dépend « tensoriellement » de X et Y, c'est-à-dire, est $\mathfrak{F}$-linéaire en X et Y. En effet, si f est une fonction numérique :

$$\begin{aligned} D_{X,fY} &= D_X D_{fY} - D_{D_X(f\cdot Y)} = D_X(f \cdot D_Y) - D_{(X\cdot f)Y + fD_X Y} \\ &= (X \cdot f) \cdot D_Y + f D_X D_Y - (X \cdot f) D_Y - f D_{D_X Y} \\ &= f \cdot D_{X,Y}. \end{aligned}$$

L'antisymétrisé du hessien mesure la non-commutativité des dérivations covariantes successives, *c'est-à-dire la courbure* de $(\mathbf{V}, D)$

$$\mathfrak{R}_{X,Y} = D_{X,Y} - D_{Y,X} = D_X D_Y - D_Y D_X - D_{D_X Y - D_Y X}.$$

Lorsque la connexion est symétrique, par exemple dans le cas riemannien, c'est l'opposé $R_{X,Y} = -\mathfrak{R}_{X,Y}$ qui est utilisé, et ceci pour des raisons géométriques, soit :

$$\boxed{R_{X,Y} = D_{[X,Y]} - (D_X D_Y - D_Y D_X)}.$$

Ce dernier opérateur $R_{X,Y}$, nul sur $\mathfrak{F}$: $(R_{X,Y}\, f = 0)$ est appelé l'*opérateur de courbure* de $(\mathbf{V}, D)$. *Il est $\mathfrak{F}$-linéaire en* X *et* Y donc ne dépend que des valeurs de X et Y en un point donné. (voir plus bas le théorème IX.6.A).

Si $(x^1, x^2, \ldots x^n)$ sont des coordonnées locales, tout *champ de covecteurs* α s'exprime en fonction des différentielles des fonctions coordonnées par : $\alpha = \sum \alpha_k dx^k$. Les opérateurs $D_{\frac{\partial}{\partial x^i}}$ opérant sur les espaces cotangents sont les transposés changés de signe des mêmes opérateurs définis sur les espaces tangents :

$$D_{\frac{\partial}{\partial x^i}}\, dx^k = -\sum \Gamma^k_{ij} dx^j$$

$$D_X \alpha = \sum X^i \Big(\frac{\partial \alpha_j}{\partial x^i} - \sum \Gamma^k_{ij} \alpha_k \Big) dx^j.$$

En notant comme précédemment : $\alpha_{j|i} = \dfrac{\partial \alpha_j}{\partial x^i} - \sum \Gamma^k_{ij}\alpha_k$ les $\alpha_{j|i}$ sont les composantes du champ de tenseurs deux fois covariants $\mathrm{D}\alpha$, différentielle covariante de α

$$\mathrm{D}\alpha = \sum \alpha_{j|i} dx^j \otimes dx^i.$$

D'une façon générale, soit t un champ de tenseurs de variance (p, q) :

$$t = \sum t^{j_1 \ldots j_p}_{k_1 \ldots k_q} \frac{\partial}{\partial x^{j_1}} \otimes \ldots \frac{\partial}{\partial x^{j_p}} \otimes dx^{k_1} \otimes \ldots dx^{k_q}$$

On a : $\mathrm{D}t = \sum dx^i \otimes \mathrm{D}_{\frac{\partial}{\partial x^i}} t$, et ses composantes s'écrivent, en notant que le premier indice covariant i de Γ^k_{ij} est celui de la dérivation et reste inchangé :

$$\begin{aligned} t^{j_1 \ldots j_p}_{k_1 \ldots k_q|i} = \frac{\partial}{\partial x^i} t^{j_1 \ldots j_p}_{k_1 \ldots k_q} + \sum \Gamma^{j_1}_{im} t^{m j_2 \ldots j_p}_{k_1 \ldots k_q} + \ldots \sum \Gamma^{j_p}_{im} t^{j_1 \ldots j_{p-1} m}_{k_1 \ldots k_q} \\ - \sum \Gamma^m_{ik_1} t^{j_1 \ldots j_p}_{m k_2 \ldots k_q} - \ldots - \sum \Gamma^m_{ik_q} t^{j_1 \ldots j_p}_{k_1 \ldots k_{q-1} m} \end{aligned}$$

tandis que $(\mathrm{D_X} t)^{j_1 \ldots j_p}_{k_1 \ldots k_q} = \sum \mathrm{X}^i t^{j_1 \ldots j_p}_{k_1 \ldots k_q|i}$.

Remarque : Les composantes Γ^k_{ij} d'une dérivation covariante dans un système de coordonnées locales $(x^1, x^2, \ldots x^n)$ sont les composantes de la fonction : $(\mathrm{X}, \mathrm{Y}) \to \mathrm{D_X Y}$, et puisque $\mathrm{D_X}(f\mathrm{Y}) = (f \cdot \mathrm{Y}) + f\mathrm{D_X}(\mathrm{Y})$, ces composantes *ne peuvent pas* être celles d'un tenseur. En effet, si $y^1, y^2, \ldots y^n$ sont d'autres coordonnées locales au voisinage du même point, en reportant : $\dfrac{\partial}{\partial y^j} = \sum \dfrac{\partial x^m}{\partial y^j} \dfrac{\partial}{\partial x^m}$ dans l'expression de la dérivée on obtient :

$$\sum \overline{\Gamma}^k_{ij} \frac{\partial}{\partial y^k} = \mathrm{D}_{\cdot\frac{\partial}{\partial y^i}} \frac{\partial}{\partial y^j} = \sum \frac{\partial^2 x^m}{\partial y^i \partial y^j} \frac{\partial}{\partial x^m} + \sum \frac{\partial x^m}{\partial y^j} \mathrm{D}_{\cdot\frac{\partial}{\partial y^i}} \frac{\partial}{\partial x^m}$$

$$\text{avec} \quad \mathrm{D}_{\cdot\frac{\partial}{\partial y^i}} \frac{\partial}{\partial x^m} = \sum \frac{\partial x^l}{\partial y^i} \mathrm{D}_{\cdot\frac{\partial}{\partial y^l}} \frac{\partial}{\partial x^m} = \sum \Gamma^p_{lm} \frac{\partial x^l}{\partial x^i} \frac{\partial}{\partial x^p}$$

$$= \sum \Gamma^p_{lm} \frac{\partial x^l}{\partial y^i} \frac{\partial y^k}{\partial x^p} \frac{\partial}{\partial y^k}$$

d'où : $$\overline{\Gamma}^k_{ij} = \sum \frac{\partial^2 x^r}{\partial y^i \partial y^j}\frac{\partial y^k}{\partial x^r} + \sum \frac{\partial x^l}{\partial y^i}\frac{\partial x^m}{\partial y^j}\frac{\partial y^k}{\partial x^p}\Gamma^p_{lm}.$$

C'est la présence des dérivées secondes qui traduit le caractère non tensoriel des Γ^k_{ij} et qui fait que même la dérivation standard de $\mathbb{R}^n$ a des composantes Γ^k_{ij} non nulles dans un système de coordonnées non affines (« curvilignes »).

On note que cette formule reste inchangée si l'on permute à la fois (i et j) et (l et m). Cela traduit le fait que les $\check{\Gamma}^k_{ij} = \Gamma^k_{ji}$ sont les composantes d'une autre connexion, « symétrique » de la connexion donnée, dont la dérivation D est définie par :

$$\check{\mathrm{D}}_{\mathrm{X}}\mathrm{Y} = \mathrm{D}_{\mathrm{Y}}\mathrm{X} + [\mathrm{X}, \mathrm{Y}] = \mathrm{D}_{\mathrm{X}}\mathrm{Y} - \mathrm{T}(\mathrm{X}, \mathrm{Y})$$

Les $\Gamma^k_{ij} - \Gamma^k_{ji}$ sont les composantes du tenseur de torsion (Définition IX.4.B)

On peut se demander quelle relation il y a entre la différentielle extérieure d des formes différentielles extérieures et une dérivation covariante D choisie sur la variété. La *différentielle covariante* ajoute un indice covariant. Il doit être placé *devant* les indices avec lesquels on veut l'antisymétriser, comme pour la différentielle extérieure.

D transforme un tenseur p fois covariant t en un tenseur $(p+1)$ fois covariant Dt. Si t est antisymétrique, Dt ne l'est que dans ses p derniers indices. On le rend antisymétrique en lui appliquant l'opérateur d'antisymétrisation partielle $a_{1,p}$, en définissant une différentielle covariante antisymétrisée $\overline{D} = a_{1,p}\mathrm{D}$ sur les *seuls tenseurs covariants antisymétriques*. On peut maintenant comparer $\overline{d}t$ et $\overline{\mathrm{D}}t = a_{1,p}\mathrm{D}t$ ($\overline{d}$ désignant la différentielle antisymétrisée des champs de tenseurs covariants antisymétriques correspondant à la différentielle extérieure). Dans le cas le plus simple $p = 2$, on a : si $\omega = \sum a_i dx^i$

$$\overline{d}\omega = \sum_{i<j}\left(\frac{\partial a_j}{\partial x^i} - \frac{\partial a_i}{\partial x^i}\right)(dx^i \otimes dx^j - dx^j \otimes dx^i)$$

$$\begin{aligned}\overline{\mathrm{D}}\omega = a_{1,1}\ \mathrm{D}\omega &= \sum_{i<j}(a_{|i,j} - a_{|j,i})(dx^i \otimes dx^j - dx^j \otimes dx^i)\\ &= \overline{d}\omega - \sum(\Gamma^k_{ij} - \Gamma^k_{ji})a_k(dx^i \otimes dx^j - dx^j \otimes dx^i)\end{aligned}$$

On a donc dans ce cas $\overline{\mathrm{D}}\omega = \overline{d}\omega$ si et seulement si la connexion est symétrique. On remarquera également que si l'on avait pas placé l'indice covariant de dérivation devant, on aurait eu $\overline{D}\omega = -\overline{d}\omega$ ce qui eut été fâcheux.

On aurait pu obtenir ce résultat de façon intrinsèque : plongeons deux vecteurs tangents X, Y en un point dans des champs de vecteurs localement définis au voisinage de ce point.

Puisque $\mathrm{D_X} = \mathrm{L_X}$ sur les fonctions $\mathfrak{C}_1$, (§VI.3) :

$$\begin{aligned} d\omega(\mathrm{X}, \mathrm{Y}) &= \mathrm{D_X} \cdot \omega(\mathrm{Y}) - \mathrm{D_Y} \cdot \omega(\mathrm{X}) - \omega([\mathrm{X}, \mathrm{Y}]) \\ &= \mathrm{D_X} \cdot \omega(\mathrm{Y}) - \mathrm{D_Y} \cdot \omega(\mathrm{X}) - \omega(\mathrm{D_X Y} - \mathrm{D_Y X} - \mathrm{T}(\mathrm{X}, \mathrm{Y})) \\ &= (\mathrm{D_X}\omega)(\mathrm{Y}) - (\mathrm{D_Y}\omega)(\mathrm{X}) + \omega(\mathrm{T}(\mathrm{X}, \mathrm{Y})) \\ &= (a_{1,1}\mathrm{D}_\omega)(\mathrm{X}, \mathrm{Y}) + \omega(\mathrm{T}(\mathrm{X}, \mathrm{Y})) \end{aligned}$$

soit, si $\mathrm{T} = 0$: $d\omega = \overline{\mathrm{D}}\omega$.

La formule générale s'obtient exactement de la même façon à partir du développement de $d\omega(\mathrm{X}_0, \mathrm{X}_1, \ldots \mathrm{X}_p)$ ce qui donne $d\omega = \overline{\mathrm{D}}\omega$ lorsque la connexion est *symétrique* :

$$d\omega(\mathrm{X}_0, \mathrm{X}_1, \ldots \mathrm{X}_p) = \sum_1 (-1)^i \{\mathrm{D}_{\mathrm{X}_i} \cdot \omega(\mathrm{X}_0, \ldots \widehat{\mathrm{X}}_i, \ldots \mathrm{X}_p) - \sum_{j \neq i} \omega(\mathrm{X}_0, \ldots \mathrm{D}_{\mathrm{X}_i}\mathrm{X}_j, \ldots \widehat{\mathrm{X}}_i, \ldots \mathrm{X}_p)\},$$ soit :

$$\boxed{d\omega(\mathrm{X}_0, \mathrm{X}_1 \ldots \mathrm{X}_p) = \sum_i (-1)^i (\mathrm{D}_{\mathrm{X}_i} \cdot \omega)(\mathrm{X}_0, \ldots \widehat{\mathrm{X}}_i, \ldots \mathrm{X}_p)}$$

(si D est symétrique).

Si D n'est pas symétrique, la torsion ajoute un terme au second membre. Si maintenant X est un champ de vecteurs tangents de classe $\mathfrak{C}_1$ sur $\mathfrak{V}$, la dérivation $\mathrm{D_X} - \mathrm{L_X}$ est définie par le champ de tenseurs $(1, 1)$: $\mathrm{DX} + \mathrm{T}(\mathrm{X}, .)$, puisque, d'après la définition même de la torsion, on a :

$$\begin{aligned} \{\mathrm{DX} + \mathrm{T}(\mathrm{X}, .)\}(\mathrm{Y}) = \mathrm{D_Y X} + \mathrm{T}(\mathrm{X}, \mathrm{Y}) &= \mathrm{D_X Y} - [\mathrm{X}, \mathrm{Y}] \\ &= (\mathrm{D_X} - \mathrm{L_X}) \cdot \mathrm{Y}. \end{aligned}$$

Comme application, on a la :

Proposition IX.4 :

Soient $\mathfrak{V}$ une variété différentielle sur laquelle on a choisi une dérivation covariante *symétrique* D. Quels que soient l'élément de volume ω *parallèle* (ou *invariant*) *relativement à la connexion*, et le champ de vecteur X, la fonction div X (cf. définition IX.3.D) est égale en chaque point de $\mathfrak{V}$ à la trace de l'opérateur linéaire $(\mathrm{DX})_x$.

Preuve : Puisque la torsion est nulle, d'où $\mathrm{L_X} = \mathrm{D_X} - \mathrm{DX}$, et que ω est parallèle ($\mathrm{D_X}\omega = 0$), il en résulte que : $(\mathrm{div\ X})_\omega = \mathrm{L_X}\omega = \mathrm{D_X}\omega - (\widetilde{\mathrm{DX}})\omega = -(\widetilde{\mathrm{DX}})\omega$ où $(\widetilde{\mathrm{DX}})_\omega$ désigne la dérivation purement algébrique de ω par le champ d'opérateurs linéaires DX, qui s'écrit (cf. §IX.3) :

$$\boxed{(\widetilde{\mathrm{DX}})\omega(\mathrm{Y}_1, \mathrm{Y}_2, \ldots \mathrm{Y}_n) = -\sum_j \omega(\mathrm{Y}_1, \ldots \mathrm{D}_{\mathrm{Y}_j}\mathrm{X}, \ldots \mathrm{Y}_n)}\,.$$

Or (fin du §IV.13) la dérivation de l'algèbre extérieure sur un espace vectoriel associée à un opérateur linéaire u agit sur un n-vecteur en le multipliant par la trace Tr u. On a donc :

$$\boxed{(\mathrm{div\ X})_x = \mathrm{Tr}(\mathrm{DX})_x} \quad (\mathrm{D}\ \textit{symétrique}),$$

qui généralise la formule du §IV.14 dans l'espace $\mathbb{R}^n$, et qui s'écrit dans un système de coordonnées locales :

$$\boxed{\mathrm{div\ X} = \sum \mathrm{X}^i_{|i} = \sum \left(\frac{\partial \mathrm{X}^i}{\partial x^i} + \sum \Gamma^j_{ij}\mathrm{X}^i\right)} \quad \text{(connexion symétrique).}$$

On obtient directement cette formule à partir de la formule du §précédent IX.3.D en explicitant la condition de parallèlisme de la forme de volume :

$$\begin{aligned} 0 = \mathrm{D_X}\omega &= \mathrm{D_X}(a\, dx^1 \wedge dx^2 \wedge \ldots dx^n) \\ &= (\mathrm{X} \cdot a)\omega + a \sum_j dx^1 \wedge \ldots \mathrm{D_X} dx^j \wedge \ldots dx^n. \end{aligned}$$

Puisque $\mathrm{D}_{\frac{\partial}{\partial x^i}} dx^j = -\sum \Gamma^j_{ik} dx^k$, d'où $\mathrm{D_X} dx^j = -\sum \mathrm{X}^i \Gamma^j_{ik} dx^k$ on obtient finalement : $\dfrac{1}{a}\,(\mathrm{X} \cdot a) = \sum \mathrm{X}^i \Gamma^j_{ij}$ c.q.f.d.

La forme volume pseudoriemannienne étant parallèle pour la connexion canonique, le calcul précédent donne dans ce cas la formule suivante :

$$\mathrm{X}\cdot|\overline{g}|^{1/2}=\left(\sum \mathrm{X}^i\Gamma^j_{ij}\right)|\overline{g}|^{1/2}.$$

Lors de la comparaison entre les opérateurs divergence et codifférentielle de la fin du paragraphe précédent, où avait été obtenue l'égalité div $\mathrm{X}=-\delta\zeta$, on avait mentionné l'utilité d'une notion générale de divergence. Celle-ci possède une *définition directe*, et non « duale » comme celle de δ, *donc plus simple*. Elle généralise la proposition précédente IX.4.

Définition IX.4.C :

Soit t un champ de tenseurs d'ordre p sur la variété riemannienne ou pseudoriemannienne $(\mathfrak{V},g)$ munie de sa connexion canonique. On appel *divergence de* t tout contracté du champ $\mathrm{D}t$ sur l'indice covariant introduit par la dérivation et un indice j venant de t, que l'on note $\mathrm{div}_j t$. Les $\mathrm{div}_j t$ sont donc des traces du champ de tenseurs $\mathrm{D}t$. S'il n'y a pas d'ambiguïté, on note simplement div t la divergence $\mathrm{div}_1 t$ correspond au premier indice. Si t est symétrique, $\mathrm{div}_i t=\mathrm{div}\ t$, $\forall i$ et le champ div t est symétrique. Si t est antisymétrique, $\mathrm{div}_i t=\pm\mathrm{div}\ t$, et div t est antisymétrique. Si α est un champ différentiable de p-formes extérieures sur $\mathfrak{V}$, la divergence de α, notée div_α est le champ de $(p-1)$-formes obtenu par *contraction antisymétrisée* de la différentielle covariante : $\mathrm{D}\alpha\in\Gamma\ (\mathrm{T}^*(\mathfrak{V})\otimes\wedge^p\mathrm{T}^*(\mathfrak{V}))$ sur l'indice de dérivation et chacun des différents indices de α.
(Remarquons que puisque D commute avec les contractions, il en est de même des divergences)

Exemple : Si t est un tenseur d'ordre deux de composantes contravariantes (t^{ij}) :

$$t^{ij}_{|k}=\frac{\partial t^{ij}}{\partial x^k}+\sum t^{hj}\Gamma^i_{kh}+\sum t^{ih}\Gamma^j_{kh}$$

et $(\mathrm{div}_1 t)^j=\sum t^{ij}_{|i}=\sum\dfrac{\partial t^{ij}}{\partial x^i}+\sum t^{hj}\Gamma^i_{ih}+\sum t^{ih}\Gamma^j_{ih}.$

D'après le calcul fait au §VII.10, également valable pour une métrique pseudoriemannienne, on a : $\sum \Gamma^i_{ih} = |\overline{g}|^{-1/2} \dfrac{\partial |\overline{g}|^{1/2}}{\partial x^h}$ et on obtient :

$$(\mathrm{div}_1 t)^j = |\overline{g}|^{-1/2} \sum \frac{\partial}{\partial x^i} \left(|\overline{g}|^{1/2} t^{ij}\right) + \sum t^{ih} \Gamma^j_{ih}$$

Le dernier terme du second membre disparaît si t est antisymétrique.

Théorème IX.4.F : Sur les formes différentielles extérieures de la variété pseudoriemannienne $(\mathfrak{V}, g)$, l'opérateur divergence est l'opposé de la codifférentielle :

$$\boxed{\mathrm{div}\ \alpha = -\delta\alpha}\,.$$

Preuve : Vérifions d'abord cette égalité en coordonnées quelconque pour une 1-forme $\alpha = \sum \alpha_j dx^j$. On a (§IX.3) :

$$*\alpha = |\overline{g}|^{-1/2} \sum (-1)^{j-1} \alpha^j dx^1 \wedge \ldots \widehat{dx^j} \wedge \ldots dx^n$$

où les $\alpha^j = \sum g^{jk} \alpha_k$ sont les composantes contravariantes de α.

$$\begin{aligned} d * \alpha &= \sum \frac{\partial}{\partial x^j} \left(|\overline{g}|^{1/2} \alpha^j\right) dx^1 \wedge dx^2 \wedge \ldots dx^n \\ &= |\overline{g}|^{-1/2} \sum \frac{\partial}{\partial x^j} \left(|\overline{g}|^{1/2} \alpha^j\right) \omega \end{aligned}$$

d'où $\delta\alpha = -\overset{-1}{*} d{*}\alpha = -(|\overline{g}|^{-1/2} . \sum \dfrac{\partial}{\partial x^j} |\overline{g}|^{1/2} \alpha^j$ (puisque $*1 = \omega$) ce qui s'écrit en effectuant la dérivation :

$$\delta\alpha = -\Big\{ \sum \frac{\partial \alpha^j}{\partial x^j} + |\overline{g}|^{-1/2} \sum \frac{\partial |\overline{g}|^{1/2}}{\partial x^j} \cdot \alpha^j \Big\}.$$

Mais d'après le calcul de l'exemple précédent (cf. §VII.10), on a

$\bar{g}^{-1/2}\dfrac{\partial|\bar{g}|^{1/2}}{\partial x^k} = \sum \Gamma^i_{ik}$. D'où :

$$\delta\alpha = -\Big\{\sum \frac{\partial \alpha^j}{\partial x^j} + \sum \Gamma^i_{ij}\alpha^j\Big\} = -\sum \mathrm{D}_j\alpha^j = -\mathrm{div}\ \alpha.$$

Nous laissons au lecteur qui en aurait envie la tâche héroïque de faire le calcul pour une p-forme en coordonnées quelconques. En effet, il suffit de vérifier l'égalité dans un système de coordonnées particulières, pour lequel on prendra bien entendu, des coordonnées normales $(x^1, \dots x^n)$ au point x (cf. §IX.9 ci-dessous).

De plus, l'égalité étant linéaire en α, il suffit de prendre pour celle-ci la p-forme décomposable : $\alpha = f\ dx^1 \wedge dx^2 \wedge \dots dx^p$.

Pour des coordonnées normales en x, $(dx^i \mid dx^j) = g^{ij} = \pm\delta^{ij}$, $|\bar{g}| = 1$, et puisque les symboles de Christoffel sont alors nuls en x, les dérivées covariantes en x sont les dérivées ordinaires.

$$*\alpha = f\ dx^{p+1} \wedge \dots dx^n;\ d(*\alpha) = \sum_{j=1}^{p} \frac{\partial f}{\partial x^j}\ dx^j \wedge dx^{p+1} \wedge \dots dx^n$$

$$\delta\alpha = (-1)^p \overset{-1}{*}\, d * \alpha = \sum_{j=1}^{p} (-1)^j \frac{\partial f}{\partial x^j}\ dx^1 \wedge \dots \widehat{dx^j} \wedge \dots dx^p$$

tandis que $\mathrm{D}\alpha = \Big(\sum \dfrac{\partial f}{\partial x^k}\ dx^k\Big) \otimes (dx^1 \wedge \dots dx^p)$ dont la contraction antisymétrisée de k et des indices de α s'écrit :

$$\mathrm{div}\ \alpha = \sum (-1)^{j-1} \frac{\partial f}{\partial x^j}\ dx^1 \wedge \dots \widehat{dx^j} \wedge \dots dx^p,\ \text{c.q.f.d.}$$

IX.5 – Connexions sur les fibrés vectoriels avec groupe structural

Les développements de la géométrie différentielle et de ses applications, à la physique par exemple, ont très vite montré que la structure de dérivation covariante « naïve » construite au paragraphe précédent était insuffisante. Elle doit être généralisée et enrichie dans deux directions :

a) définition d'une dérivation covariante des sections de fibrés vectoriels, c'est-à-dire de champs de vecteurs qui ne sont plus nécessairement des vecteurs tangents.

Cette expression est immédiate et on y retrouve les analogues de la définition IX.4.A et d'une partie de ses conséquences.

Définition IX.5.A :

Une dérivation covariante de classe $\mathfrak{C}_r$ sur un espace fibré vectoriel E de base $\mathfrak{V}$, est une fonction D qui fait correspondre à tout vecteur tangent A de $\mathfrak{V}$ en un point x_0 et à toute section différentiable V de E un vecteur $\mathrm{D_A V}$ de E_x appelé dérivée (covariante) de V par A de telle sorte que les axiomes suivants soient satisfaits :

1) $\mathrm{D_A}$ est $\mathbb{R}$-linéaire en A et V.

2) Si f est une fonction numérique différentiable sur $\mathfrak{V}$, $\mathrm{D_A}f(\mathrm{V}) = (\mathrm{A}f)\mathrm{V} + f(x)\mathrm{D_A V}$.

3) Quels que soient le champ X sur $\mathfrak{V}$, la section V de E, de classe $\mathfrak{C}_r$, la section $\mathrm{D_X V}$ de E est de classe $\mathfrak{C}_r$.

La forme différentielle à valeurs dans l'espace fibré vectoriel $\mathrm{E} \otimes \mathrm{T}^* : \mathrm{X} \to \mathrm{D_X V}$ est la différentielle DV de V.

Si $e = (e_1, e_2, \dots e_n)$ est un champ de repères de $\mathrm{T}(\mathfrak{V})$, et $\varepsilon_\alpha = (\varepsilon_{\alpha 1}, \varepsilon_{\alpha 2}, \dots \varepsilon_{\alpha p})$ un champ de repères de E au-dessus d'un ouvert U_α de $\mathfrak{V}$, on note Γ^k_{ij} la k-ème composante dans ε de $\mathrm{D}_{e_i}\varepsilon_j$. On obtient en développant :

$$\mathrm{D_X V} = \mathrm{D}_{\sum e_i \mathrm{X}^i}(\textstyle\sum \mathrm{V}^j \varepsilon_j) = \sum \mathrm{X}^i(e_i \mathrm{V}^j)\varepsilon_j + \sum \Gamma^k_{ij}\mathrm{X}^i\mathrm{V}^j\varepsilon_k.$$

Si $t \to c(t)$ est un chemin différentiable de U_α, on a, en posant $\mathrm{V}(t) = \mathrm{V}(c(t))$:

$$\frac{\mathrm{DV}}{dt} = \mathrm{D}_{\frac{dc}{dt}}\mathrm{V} = \mathrm{D}_{\frac{dc}{dt}}(\textstyle\sum \mathrm{V}^j\varepsilon_j) = \sum \dfrac{d\mathrm{V}^j}{dt}\,\varepsilon_j + \sum \Gamma^k_{ij}\,\dfrac{dc^i}{dt}\,\mathrm{V}^j\varepsilon_k$$

qui ne dépend que de la restriction de V au chemin c. Une section V de E au-dessus du chemin c (que l'on peut appeler un champ de vecteurs de E le long de c) est parallèle si $\mathrm{D}_{\frac{dc}{dt}}\mathrm{V} = 0$ quel que soit t ce qui se traduit par le système d'équations différentielles linéaire du premier ordre :

$$\frac{d\mathrm{V}^k}{dt} + \sum \Gamma^k_{ij}\frac{dc^i}{dt}\,\mathrm{V}^j = 0 \quad (k = 1, 2, \dots p).$$

Les notions de connexion symétrique et de torsion n'ont évidemment de sens que si $\mathrm{E} = \mathrm{T}(\mathfrak{V})$. Par contre on a toujours un opérateur de courbure qui à chaque couple X, Y, de vecteurs tangents en un point x de $\mathfrak{V}$: fait correspondre l'opérateur $\mathrm{R}_{\mathrm{X,Y}} = \mathrm{D}_{[\mathrm{X,Y}]} - (\mathrm{D_X D_Y} - \mathrm{D_Y D_X})$ dont l'expression suppose les vecteurs tangents en x plongés dans des champs locaux de vecteurs tangents mais qui, étant $\mathfrak{F}$-linéaires en X et Y ne dépend que des vecteurs X et Y en x.

b) Intervention de groupes structuraux, ce qui sera expliqué ci-dessous, mais dont l'importance peut aisément être appréciée par l'exemple des connexions riemanniennes. On a vu en effet que, par construction, le transport parallèle associé à une structure riemannienne conservait la métrique, ce qui localement se traduit par le fait que la différentielle covariante du tenseur métrique est nulle. On est donc conduit à trouver une manière de définir une connexion sur un fibré vectoriel E mettant en évidence dès le départ la conservation par le transport parallèle associé de structures sur les fibres de E. Les plus simples, et les plus importantes consistent en la donnée sur les fibres de E d'une « géométrie » caractérisée par un groupe structural G (G est le groupe orthogonal dans le cas d'une métrique riemannienne sur l'espace tangent).

Si G est le groupe d'invariance d'un ou plusieurs tenseurs sur les fibres, la dérivation covariante associée peut naturellement être définie en imposant qu'elle annule ces tenseurs. Si l'on veut par contre, se placer uniquement du point de vue du groupe lui-même, considéré comme une donnée, il faut trouver une autre façon de définir une connexion que par simple dérivation des champs de vecteurs.

Nous allons rechercher une telle définition en réexaminant le cas de l'espace tangent. Soit : $t \longrightarrow c(t)$ un chemin différentiable dans la variété $\mathfrak{V}$ munie d'une connexion D. Soient $e = (e_1, e_2, \dots e_n)$ un champ de repères de $\mathrm{T}(\mathfrak{V})$ le long de c et Y un champ de vecteurs le long de c. On note $\mathrm{Y}(t) = \mathrm{Y}(c(t))$, $e(t) = e(c(t))$.

La dérivée covariante de Y : $\mathrm{D}_{\frac{dc}{dt}}\, \mathrm{Y} = \dfrac{\mathrm{DY}}{dt}$ peut être considérée comme la vitesse absolue du point $\mathrm{Y}(t) = \sum \mathrm{Y}^j(t) e_j(t)$ de l'espace tangent $\mathrm{T}(\mathbf{V})$. La décomposition :

$$\frac{\mathrm{DY}}{dt} = \sum \frac{d\mathrm{Y}^j}{dt}\, e_j + \sum \frac{dc^i}{dt}\, \mathrm{Y}^j \Gamma^k_{ij} e_k$$

exprime cette vitesse absolue comme *somme de la vitesse relative du point* $\mathrm{Y}(t)$ *dans le repère* e *et de sa vitesse d'entraînement* quand on le considère comme lié à ce repère (Y^j constants). On peut donc considérer l'intervention d'une connexion comme opérant sur les repères (point de vue du *repère mobile*) plutôt que sur les vecteurs. Cette opération s'effectue alors au moyen de la forme locale de la connexion, qui dans des champs de repères $e = (e_1, e_2, \dots e_n)$ $e^* = (\omega^1, \omega^2, \dots \omega^n)$ s'écrit : $\gamma = \sum \Gamma^k_{ij} \omega^i \otimes e_k \otimes \omega^j$.

Nous allons préciser ce point de vue.

Si $e(t+u)$ est le repère au point $c(t+u)$, notons $\overline{e}(t+u) = \tau_{t+u,t} e(t+u)$ le repère au point $c(t)$ transporté parallèle le long de c de $e(t+u)$ (cf. théorème IX.4.E).

$\overline{e}(t+u) = e(t)\mathrm{S}(u)$ où $\mathrm{S}(u)$ est un chemin différentiable issu de l'identité ($\mathrm{S}(0) = \mathrm{Id}$) dans le groupe linéaire $\mathrm{Gl}(n)$ des matrices réelles inversibles $(n \times n)$. Soit $\mathrm{V} = |v^j|$ une matrice colonne $(n \times 1)$ fixe. Elle détermine un champ de vecteurs le long de c ayant des composantes fixes dans les repères $e(t)$, soit : $v(t) = e(t) \cdot \mathrm{V}$. Le transporté parallèle dans $e(t)$ de $\mathrm{V}(t+u)$ est $\overline{\mathrm{V}}(t+u) = \overline{e}(t+u) \cdot \mathrm{V}$.

La dérivée covariante de V au point $c(t)$ est donc (cf. théorème IX.4.E) :

$$\begin{aligned}\frac{\mathrm{DV}}{dt} &= \frac{d}{du}\,(\overline{e}(t+u)\mathrm{V})_{u=0} = \frac{d}{du}\,(e(t)\mathrm{S}(u)\mathrm{V})_{u=0}\\ &= e(t)\mathrm{S}'(o)\mathrm{V}\end{aligned}$$

La vitesse d'entraînement du repère $e(t)$ en $c(t)$ définie par la connexion est ainsi :

$$\mathrm{S}'(o) = \sum \frac{dc^i}{dt}\,\Gamma^k_{ij}\, e_k \otimes \omega^j.$$

Les éléments de matrice sont : $\mathrm{S}'(o)^k_j = \sum\limits_i \dfrac{dc^i}{dt}\,\Gamma^k_{ij}$.

Les vecteurs tangents en l'identité du groupe linéaire $\mathrm{Gl}(n)$ forment son algèbre de Lie $gl(n)$: espace vectoriel de toutes les matrices réelles $n \times n$ muni du crochet, qui est ici le commutateur : $[a, b] = ab - ba$, et de la représentation adjointe de $\mathrm{Gl}(n)$: $(\mathrm{ad}\, g) \cdot a = ga \overset{-1}{g}$.

Nous voyons donc que la forme locale de la connexion rapportée à un champ de repères e sur un ouvert U est une « forme différentielle à valeurs dans l'algèbre de Lie $gl(n)$ » fonction linéaire des vecteurs tangents : $X = \sum X^i e_i \to \sum X^i \Gamma_{ij}^k e_k \otimes \omega^j$ à valeurs dans $gl(n)$.

Si la connexion est riemannienne, le transport parallèle conserve la métrique. En ne considérant que des champs de repères orthonormés sur U, les matrices précédentes $S(u)$ sont des matrices orthogonales, et la forme de la connexion est maintenant à valeurs dans l'algèbre de Lie $o(n)$ du groupe $O(n)$.

On voit ainsi que la forme locale de la connexion manifeste la présence d'un groupe de structure sur les espaces tangents, si les repères sont adaptés, et celà par l'algèbre de Lie de ce groupe.

Puisqu'une connexion sur l'espace tangent $T(\mathfrak{V})$ agit essentiellement sur ses repères, il faut mettre ceux-ci en évidence. En considérant systématiquement des « espaces de repères » on va alors pouvoir remplacer les formes locales d'une connexion qui sont des formes différentielles sur des ouverts assez petits de $\mathfrak{V}$, à valeurs dans une algèbre de Lie, par une *forme différentielle unique*, dite forme de la connexion, mais, cette fois définie sur la variété des repères ! Voyons les explications détaillées.

Soit P_x l'ensemble des repères de l'espace tangent $T(\mathfrak{V})_x$ en un point x de $\mathfrak{V}$. Le groupe linéaire $Gl(n)$ des matrices inversibles d'ordre n opère transitivement à droite sur P_x. En effet, si $e = (e_1, e_2, \ldots e_n) \in P_x$ et $A = |a_{i,j}|$, $e' = eA$ avec $e'_j = \sum e_i a^i_j$ tandis que, si $B = |b^j_k|$ et $e'' = e'B$ avec $e''_k = \sum e'_j b^j_k$, on a bien $e' = eAB$, soit $r(B)(r(A)e) = r(AB)e$.

On peut considérer l'ensemble P des P_x, $\forall x \in \mathfrak{V}$. Des cartes de $\mathfrak{V}$ définissent de façon évidente des cartes et une structure différentielle sur P, de classe $\mathfrak{C}_{r-1}$ (comme $T(\mathfrak{V})$) si $\mathfrak{V}$ est de classe de $\mathfrak{C}_r$. Un champ de repères sur un ouvert U de $\mathfrak{V}$ est une section de P au-dessus de U. De plus, on vient de voir que le groupe linéaire $G = Gl(n)$ opère différentiablement et fidèlement à droite sur P. Cet exemple appelé *fibré des repères* conduit à la définition suivante :

Définition IX.5.B :

Un espace *fibré principal* P, de *groupe* (de Lie) G, au dessus d'une variété différentielle $\mathfrak{V}$, dite *base de* P, est une variété différentielle sur laquelle le groupe G opère à droite, par une multiplication $P \times G \to P$ différentiable de telle sorte que les axiomes suivants soient satisfaits :

1) G opère fidèlement sur P : si $g \neq h$ et $p \in \mathrm{P}$, $pg \neq ph$.

2) il existe une projection différentiable π de P sur $\mathfrak{V}$ telle que, quel que soit $x \in \mathfrak{V}$, $\overset{-1}{\pi}\, x$ soit une trajectoire de G : $\pi(pg) = \pi(p)$, $\forall g$ et si $p, q \in \overset{-1}{\pi}\, x$, il existe $g \in \mathrm{G}$, unique, tel que $q = pg$.

3) il existe un recouvrement ouvert (U_α) de $\mathfrak{V}$ et, pour chaque U_α une section différentiable σ_α de P au dessus de U_α, d'où résulte un difféomorphisme respectant les fibres et les opérations de G de $\overset{-1}{\pi}\, \mathrm{U}_\alpha$ sur $\mathrm{U}_\alpha \times \mathrm{G}$. (structure locale de produit).

Si $\mathrm{U}_\alpha \cap \mathrm{U}_\beta \neq \emptyset$, la fonction différentiable, $g_{\alpha\beta}$ de $\mathrm{U}_\alpha \cap \mathrm{U}_\beta$ dans G définie par $\sigma_\beta = \sigma_\alpha g_{\alpha\beta}$ est dite *fonction de transition*. On a évidemment, si $\mathrm{U}_\alpha \cap \mathrm{U}_\beta \cap \mathrm{U}_\gamma \neq \emptyset$, $g_{\alpha\gamma} = g_{\alpha\beta} g_{\beta\gamma}$.

Réciproquement, la donnée d'un tel ensemble de fonctions de transition, sur un recouvrement ouvert $\{\mathrm{U}_\alpha\}$ de $\mathfrak{V}$, à valeurs dans un groupe G, détermine, par « collage » des « morceaux » : $\mathrm{U}_\alpha \times \mathrm{G}$, un fibré G-principal P sur $\mathfrak{V}$. Les $g_{\alpha\beta}$ déterminent donc P complètement.

Remarques :

1) Lorsque P est un fibré principal de repères, un chemin $t \to p(t)$ dans P est un repère mobile le long du chemin $c(t) = \pi p(t)$ dans $\mathfrak{V}$. Le vecteur tangent $p'(t)$ représente la « vitesse » de ce repère en $p(t)$. Un chemin « vertical » dans P : $c(t) = \pi p(t) = c(o) = x$ fixe, représente alors évidemment un repère mobile d'origine fixe dans $\mathfrak{V}$.

2) Puisque $g \to pg$ est un difféomorphisme de G sur la fibre $\mathrm{P}_{\pi p}$, il détermine un isomorphisme de l'espace tangent $\mathrm{T}(\mathrm{G})_e$, identifié à l'algèbre de Lie $\mathfrak{g}$ de G, sur l'espace tangent « vertical » de P en p : $\mathrm{V}_p = \mathrm{V}(\mathrm{P}_{\pi p})_p$, le vecteur $\mathrm{A} \in \mathfrak{g}$ ayant pour image le vecteur tangent $\widetilde{\mathrm{A}}_p$ au chemin $p \cdot \exp \mathrm{A}t$ dans P. Dans le cas d'un fibré de repères, il représente la « vitesse » du repère mobile $p \cdot \exp \mathrm{A}t$ d'origine fixe πp. La « vitesse », dans ce même mouvement, d'un autre repère $p' = pg$, soit $p'(t) = p(t)g = p \cdot \exp \mathrm{A}t \cdot g = pg \overset{-1}{g} \exp \mathrm{A}t \cdot g = pg \cdot \exp(\mathrm{Ad}.\, \overset{-1}{g}\, \mathrm{A}t)$ est donc, $(\widetilde{\mathrm{Ad}\, \overset{-1}{g} \cdot \mathrm{A}})_{pg}$.

En s'inspirant de la signification d'une connexion sur le fibré des repères tangents comme « définition de la vitesse d'entraînement d'un repère mobile » on obtient la définition générale (et finale) suivante :

Définition IX.5.C :

Une connexion sur un fibré principal P, de groupe G, sur une variété $\mathfrak{V}$, est une 1-forme différentielle ω sur P, à valeurs dans l'algèbre de Lie $\mathfrak{g}$ de G, satisfaisant aux deux conditions suivantes :

1) pour tout $\mathrm{A} \in \mathfrak{g}$, $\omega(\widetilde{\mathrm{A}}_p) = \mathrm{A}$.

2) pour tout $g \in \mathrm{G}$, $r(g)^*\omega = \mathrm{ad}\,\overset{-1}{g}\cdot\omega$, soit, $\forall \mathrm{X} \in \mathrm{T(P)}$, $\omega(r(g)_*\mathrm{X}) = \mathrm{ad}\,\overset{-1}{g}\cdot\omega(\mathrm{X})$.

Remarques :

1) signifie qu'en chaque point $p \in \mathrm{P}$, ω_p peut être considéré comme un projecteur de l'espace tangent $\mathrm{T(P)}_p$ sur le sous-espace « vertical » V_p tangent à la fibre $\mathrm{P}_{\pi p}$ si on identifie V_p à $\mathfrak{g}$ par $\mathrm{A} \longrightarrow \widetilde{\mathrm{A}}_p$. L'ensemble des vecteurs X de $\mathrm{T(P)}_p$ tels que $\omega(\mathrm{X}) = 0$ est un sous-espace H_p, naturellement appelé « horizontal », que la projection π_* applique isomorphiquement sur l'espace tangent à $\mathfrak{V}$ en πp : $\mathrm{T}(\mathfrak{V})_{\pi p}$, et l'on a $\mathrm{H}_{pg} = r(g)_*\mathrm{H}_p$. Réciproquement, la donnée d'une famille différentiable de sous-espace horizontaux de T(P) satisfaisant à cette dernière condition (invariance par les opérations de G) détermine complètement la forme ω d'une connexion sur P.

Une section différentiable σ de P au-dessus d'un ouvert U de $\mathfrak{V}$ peut être considérée comme un « champ différentiable de repères ». L'image réciproque de la forme ω d'une connexion sur P par σ, soit $\gamma = \sigma^*\omega = \omega \circ \sigma_*$ est la *forme locale de la connexion* relative au champ de repères σ.

2) Les vecteurs tangents en $p \in \mathrm{P}$ appartenant à $\mathrm{H}_p = \mathrm{Ker}\,\omega_p$ correspondent à une « vitesse d'entraînement » nulle, autrement dit, sont tangents à un « transport parallèle ». Un chemin $p(t)$ tel qu'en tout point le vecteur tangent $\dfrac{dp}{dt}$ soit horizontal (appartient à $\mathrm{H}_{p(t)}$) représente un transport parallèle le long du chemin $c(t) = \pi p(t)$ de $\mathfrak{V}$. Si l'on se donne l'origine p_o, un chemin différentiable $c(t)$ de $\mathfrak{V}$ tel que $c(o) = \pi p_o$ possède un relèvement horizontal unique $p(t)$ dans P d'où l'unicité du transport parallèle le long de c.

Dans ce §IX.5, nous somme partis des fibrés vectoriels, et même des plus simples d'entre eux, les espaces fibrés tangents, pour aboutir au concept de fibré principal. Il faut maintenant montrer comment s'effectue la démarche inverse.

Soient P un fibré différentiable principal de groupe (de Lie) G sur $\mathfrak{V}$, F un espace vectoriel réel de dimension finie muni d'une représentation linéaire de G (plus généralement, F peut être une variété différentielle sur laquelle G opère *à gauche*, différentiablement).

Sur le produit P × F, faisons opérer G à droite par :

$g \cdot (p \times v) = pg \times \overset{-1}{g}\, v$ et soit $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ le quotient de $\mathrm{P} \times \mathrm{F}$ par cette action (espace des trajectoires de G). L'application Π de $\mathrm{P} \times \mathrm{F}$ sur $\mathfrak{V}$ définie par : $\Pi(p \times v) = \Pi p$ passe au quotient et détermine une application Π_{F} de $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ sur $\mathfrak{V}$. On ne peut avoir $(p \times_{\mathrm{G}} v) = (p \times_{\mathrm{G}} w)$ dans $\mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ que si $v = w$. Un élément de la fibre $\overset{-1}{\Pi}_{\mathrm{F}} x$ ou $x = \Pi p$ est l'image de $(q \times w)$ avec $\Pi q = \Pi p$, d'où $q = pg$ et $(q \times_{\mathrm{G}} w) = (p \times_{\mathrm{G}} \overset{-1}{g}\, w)$. Chaque élément $p \in \mathrm{P}$ détermine ainsi une bijection de F sur la fibre $\overset{-1}{\Pi}_{\mathrm{F}} \cdot (\Pi p) = \mathrm{E}_{\pi p}$ par : $u \in \mathrm{F} \to v = p \times_{\mathrm{G}} u \in \mathrm{E}_{\pi p}$ et lorsqu'on considère $p \in \mathrm{P}$ comme une application de F sur $\mathrm{E}_{\pi p}$ on peut écrire : $v = p \times_{\mathrm{G}} u = p \cdot u$, v est la représentation de u dans le « repère » p (voir ci-dessous).

Une carte locale de P au-dessus de l'ouvert U de $\mathfrak{V}$ est un G-difféomorphisme de P_{U} sur $\mathrm{U} \times \mathrm{G}$ et $(\mathrm{U} \times \mathrm{G}) \times \mathrm{F}/\mathrm{G} = (\mathrm{U} \times \mathrm{G}) \times_{\mathrm{G}} \mathrm{F}$ s'applique bijectivement sur $\mathrm{U} \times \mathrm{F}$ par : $(x \times e) \times_{\mathrm{G}} u \to (x \times u)$. Les cartes de $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ qui sont ainsi construites, définissent sur $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ une *structure d'espace fibré vectoriel différentiable de groupe structural* G *sur* $\mathfrak{V}$*, associé à* P. Les points p de P peuvent être conçus, au sens large, comme des *repères* de $(\mathrm{P} \times_{\mathrm{G}} \mathrm{F})_{\pi p} = \overset{-1}{\Pi}_{\mathrm{F}}(\Pi p) = \mathrm{E}_{\pi p}$ puisque ce sont des isomorphismes de la fibre type F sur la fibre $\mathrm{E}_{\pi p}$. Le transport parallèle le long d'un chemin différentiable $c(t)$ à partir de $c_0 = c(o)$ dans E se détermine à l'aide du transport parallèle de ses repères dans P à partir de p_0, $\Pi p_0 = c_0$ par le relèvement horizontal unique $p(t)$ dans P issu de p_0 du chemin $c(t)$. Considérons un champ de vecteurs : $t \to v(t)$, section du fibré vectoriel $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ le long du chemin $t \to c(t)$ de $\mathfrak{V}$. On peut définir $v(t)$ au moyen d'un champ de repères : $t \to p(t)$, section de P au-dessus de c, par : $c(t) = p(t) \times_{\mathrm{G}} u(t)$ où $t \to u(t)$ est un chemin dans F qui peut être considéré comme décrivant le mouvement relatif de $v(t)$ par rapport au repère.

La *dérivée covariante de* $v(t)$ est alors la somme de la vitesse relative et de la vitesse d'entraînement de $v(t)$, soit :

$$\boxed{\frac{\mathrm{D}v}{dt} = p(t) \times_{\mathrm{G}} \left\{\frac{du}{dt} + \omega\Big(\frac{dp}{dt}\Big) \cdot u(t)\right\} = p(t) \cdot \left\{\frac{du}{dt} + \omega\Big(\frac{dp}{dt}\Big) u(t)\right\}}$$

De la même façon un champ de vecteurs v sur $\mathfrak{V}$ peut s'exprimer localement à l'aide d'un champ de repères σ_{U}, section de P au-dessus d'un ouvert U de $\mathfrak{V}$, soit : $v = \sigma_{\mathrm{U}} \times_{\mathrm{G}} u$, où u est une fonction différentiable de U dans la fibre type F. Soit γ_{U} la forme locale de la connexion : $\gamma_{\mathrm{U}} = \sigma_{\mathrm{U}}^{*}\omega$. Si X est un vecteur tangent à $\mathfrak{V}$ en $x \in \mathrm{U}$, on a alors (avec $\mathrm{L}_{\mathrm{X}}u = \mathrm{X} \cdot u$) :

$$\boxed{\mathrm{D}_{\mathrm{X}}v = \sigma_{\mathrm{U}}\{\mathrm{L}_{\mathrm{X}}u + \gamma_{\mathrm{U}}(\mathrm{X}) \cdot u\} = \sigma_{\mathrm{U}}\{\mathrm{L}_{\mathrm{X}}u + \omega(\sigma_{\mathrm{U}*}\mathrm{X}) \cdot u\}}$$

Lorsque le groupe de Lie G est un sous-groupe du groupe linéaire Gl(F) de l'espace vectoriel F on peut donner des éléments de G et de son algèbre de Lie $\mathfrak{g}$ des expressions tensorielles puisque $\mathrm{Gl(F)} \subset \mathfrak{L}(\mathrm{F}) \equiv \mathrm{F} \otimes \mathrm{F}^{*}$. Si au-dessus d'un ouvert U de $\mathfrak{V}$, on a des champs de repères $e = (e_1, e_2, \dots e_n)$ de $\mathrm{T}(\mathfrak{V})$, $(\omega^1, \dots \omega^n)$ dual de e dans $\mathrm{T}^{*}(\mathfrak{V})$, $\varepsilon = (\varepsilon_1, \dots \varepsilon_p)$ de E, $(\lambda^1, \dots \lambda^p)$ dual de ε, la forme locale de la connexion s'écrit : $\gamma_{\mathrm{U}} = \sum \Gamma_{ij}^{k} \omega^i \otimes \varepsilon_k \otimes \lambda^j$, et l'on retrouve ainsi les expressions initiales.

La dérivation covariante ainsi définie sur les champs de vecteurs de E s'étend naturellement aux champs de tenseurs de toutes natures sur E, sections des espaces fibrés vectoriels $\otimes^{(v)}\mathrm{E}$.

Remarques :

1) Si la représentation de G dans F est la représentation identique le fibré vectoriel associé $\mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ se réduit évidemment au produit $\mathfrak{V} \times \mathrm{F}$.

2) Si P est le fibré des repères du fibré tangent $\mathrm{T}(\mathfrak{V})$ d'une variété différentielle $\mathfrak{V}$ de dimension n, on a évidemment : $\mathrm{T}(\mathfrak{V}) \equiv \mathrm{P} \times_{\mathrm{Gl}(n)} \mathbb{R}^n$.

Au début de ce paragraphe, une intervention de groupes structuraux avait été souhaitée pour tenir compte de structures éventuelle sur les fibres d'un espace fibré vectoriel conservées par transport parallèle.

La plus importante est l'existence sur E d'un produit scalaire sur les fibres : $(|)_x$ sur E_x, dépendant différentiablement de x : $(|)$ est une section différentiable de $E^* \otimes E^*$.

Cette structure est équivalente à une structure quadratique (cf. [DR]) caractérisée pour un espace vectoriel réel par son indice (r, s). Soit donc $F = \mathbb{R}^m$ muni de la forme quadratique standard $q(x) = (x^1)^2 + \ldots (x^r)^2 - (x^{r+1})^2 - \ldots - (x^{r+s})^2$ avec $r + s = m$. L'existence d'une structure quadratique de type (r, s) sur E entraîne l'existence d'un sous-fibré P_0 du fibré P des repères de E : si P est l'ensemble des isomorphismes linéaires de $\mathbb{R}^m$ sur les fibres de E, P_0 est le sous-ensemble des isomorphismes quadratiques de $(\mathbb{R}^m; q)$ sur les fibres de E. Le groupe de structure de P_0 est le groupe des automorphismes de $(\mathbb{R}^m; q)$; c'est-à-dire le groupe orthogonal $G_0 = O(r, s)$, sous-groupe de $Gl(m)$. Son algèbre de Lie $\mathfrak{g}_0$ est l'algèbre de Lie des opérateurs linéaires de $\mathbb{R}^m$ « antisymétriques » relativement au produit scalaire :

$$a \in \mathfrak{g}_0 \iff (ax \mid y) + (x \mid ay), \ \forall x, y \in \mathbb{R}^m.$$

On a alors le théorème suivant :

Théorème IX.5 : Le transport parallèle associé à une connexion ω sur le fibré principal P conserve le produit scalaire du fibré vectoriel associé E si et seulement si ω est à valeurs dans l'algèbre de Lie $\mathfrak{g}_0$ du groupe orthogonal G_0.

Preuve : Exprimés dans un champ local de repères, soient $u(t)$, $v(t)$ des champs de vecteurs de E le long du chemin $c(t)$ de $\mathfrak{V}$. Si γ est la forme locale de la connexion, u et v sont parallèles le long de c si et seulement si :

$$\frac{\mathrm{D}u}{dt} = \frac{du}{dt} + \gamma\Big(\frac{dc}{dt}\Big)u = 0 \text{ et } \frac{\mathrm{D}v}{dt} = \frac{dv}{dt} + \gamma\Big(\frac{dc}{dt}\Big)v = 0.$$

On a donc :

$$\begin{aligned}\frac{d}{dt}\,(u(t) \mid v(t)) &= \Big(\frac{du}{dt} \mid v\Big) + \Big(u \mid \frac{dv}{dt}\Big)\\ &= -\Big\{\Big(\gamma\Big(\frac{dc}{dt}\Big)u \mid v\Big) + \Big(u \mid \gamma\Big(\frac{dc}{dt}\Big)v\Big)\Big\}.\end{aligned}$$

Il en résulte que $(u(t) \mid v(t))$ est constant si et seulement si $\gamma\left(\frac{dc}{dt}\right)$ est antisymétrique relativement au produit scalaire, c'est-à-dire appartient à $\mathfrak{g}_0$, c.q.f.d.

IX.6 – Formes de courbure. Formes extérieures tensorielles

Rappelons que, par définition, la restriction de la forme ω de la connexion à une fibre P_x du fibré principal P est une 1-forme différentielle à valeurs dans l'algèbre de Lie $\mathfrak{g}$ du groupe structural G de P, telle que : $\omega(\widetilde{A}) = A$, quel que soit $A \in \mathfrak{g}$ où $\widetilde{A}$ est le champ de vecteurs tangents défini sur P par le sous-groupe à un paramètre exp At de G.

Pour exprimer la restriction de la différentielle extérieure de ω aux fibres, il suffit d'utiliser les champs verticaux $\widetilde{A}$, d'où :

$$d\omega(\widetilde{A}, \widetilde{B}) = \widetilde{A} \cdot \omega(\widetilde{B}) - \widetilde{B} \cdot \omega(\widetilde{A}) - \omega([\widetilde{A}, \widetilde{B}]).$$

Puisque $\omega(\widetilde{A})$ et $\omega(\widetilde{B})$ sont constants, les deux premiers termes du second membre son nuls. De plus : $\omega([\widetilde{A}, \widetilde{B}] = \omega(\widetilde{[A, B]}) = [A, B] = [\omega(\widetilde{A}), \omega(\widetilde{B})]$ où les trois derniers crochets sont les crochets de l'algèbre de Lie $\mathfrak{g}$.

On obtient finalement : $d\omega(\widetilde{A}, \widetilde{B}) = -[\omega(\widetilde{A}), \omega(\widetilde{B})]$. Or, si X et Y sont deux vecteurs tangents en un point quelconque p de P, la fonction : $(X, Y) \to [\omega(X), \omega(Y)]$ est bilinéaire alternée. C'est une 2-forme extérieure sur P que l'on peut désigner par $[\omega, \omega]$ et qui s'exprime par l'opérateur $\omega \wedge \omega$ dans les espaces où $\mathfrak{g}$ est représentée linéairement. Pour obtenir la partie intéressante de $d\omega$, on va lui soustraire sa partie verticale, qui est la même quelle que soit la connexion, représentée par $-[\omega, \omega]$. D'où :

Définition IX.6.A :

On appelle *forme de courbure de la connexion* ω sur le fibré G-principal P, la 2-forme différentielle extérieure ω sur P définie par :

$$\boxed{\Omega(X, Y) = d\omega(X, Y) + [\omega(X), \omega(Y)] \quad \forall X, Y \in T(P)_p}$$

soit :

$$\boxed{\Omega = d\omega + [\omega, \omega]}.$$

D'après ce qui précède, la restriction de Ω aux fibres de P est nulle. Mais il y a plus : $\Omega(\mathrm{X}, \mathrm{Y})$ est nul dès que *l'un* des arguments X ou Y est vertical, autrement dit, $\Omega(\mathrm{X}, \mathrm{Y})$ ne dépend que des composantes horizontales $h\mathrm{X}$, $h\mathrm{Y}$ des vecteurs $\mathrm{X}, \mathrm{Y} \in \mathrm{T(P)}_p$. C'est une conséquence du lemme suivant :

Lemme IX.6 : Soit Z un champ de vecteurs tangents différentiable *horizontal* sur $\mathrm{P}(\mathrm{Z}_p \in \mathrm{H}_p, \forall p)$, *invariant par* G : $r(g)_*\mathrm{Z}_p = \mathrm{Z}_{pg}$. Tous les vecteurs Z_p d'une même fibre P_x ont donc même projection z_x dans $\mathrm{T}(\mathfrak{V})_x$ et réciproquement, tout champ de vecteurs tangents différentiable z sur $\mathfrak{V}$ détermine, par relèvement dans les sous-espaces horizontaux de P un tel champ Z. On a alors la propriété : quel que soit le champ $\widetilde{\mathrm{A}}$, champ vertical défini par $\mathrm{A} \in \mathfrak{g}$, le crochet $[\widetilde{\mathrm{A}}, \mathrm{Z}]$ est nul.

Preuve : C'est une conséquence directe de l'invariance de Z par G. En effet, d'après ce qui a été vu au §IX.3 :

$$[\widetilde{\mathrm{A}}, \mathrm{Z}]_p = \lim_{t \to 0} \frac{1}{t} \left\{ r(\exp(-\mathrm{A}t))_* \mathrm{Z}_{\exp \mathrm{A}t} - \mathrm{Z}_p \right\} = 0.$$

Soient donc X, Y deux vecteurs tangents à P en p que l'on décompose en leurs composantes horizontales : $h\mathrm{X}$, $h\mathrm{Y}$ et verticales $v\mathrm{X}, v\mathrm{Y}$. On peut plonger les composantes horizontales $h\mathrm{X}$, $h\mathrm{Y}$ dans des champs horizontaux invariants par G et les composantes verticales dans des champs $\widetilde{\mathrm{A}}$ et $\widetilde{\mathrm{B}}$. En remplaçant X et Y par leurs décompositions dans l'expression de Ω et en éliminant les termes évidemment nuls, il reste :

$$\begin{aligned}\Omega(\mathrm{X}, \mathrm{Y}) = {} & d\omega(h\mathrm{X}, h\mathrm{Y}) + d\omega(h\mathrm{X}, \widetilde{\mathrm{B}}) + d\omega(\widetilde{\mathrm{A}}, h\mathrm{Y}) + d\omega(\widetilde{\mathrm{A}}, \widetilde{\mathrm{B}}) \\ & + [\omega(\widetilde{\mathrm{A}}), \omega(\widetilde{\mathrm{B}})].\end{aligned}$$

Dans le second membre, la somme des deux derniers termes est nulle, le second et le troisième sont nuls d'après le lemme, d'où :

$$\boxed{\Omega(\mathrm{X}, \mathrm{Y}) = d\omega(h\mathrm{X}, h\mathrm{Y}) = -\omega([h\mathrm{X}, h\mathrm{Y}])}.$$

Proposition IX.6.A :

La différentielle extérieure $d\Omega$ de la forme de courbure d'une connexion ω sur le fibré G-principal P a une restriction nulle sur les sous-espaces horizontaux de T(P) : $d\Omega(h\mathrm{X}, h\mathrm{Y}, h\mathrm{Z}) = 0$ (seconde identité de Bianchi).

Preuve : $\Omega = d\omega + [\omega, \omega]$, d'où $d\Omega = d[\omega, \omega]$.

$$\begin{aligned} d[\omega,\omega](\mathrm{X},\mathrm{Y},\mathrm{Z}) = {} & \mathrm{X}\cdot[\omega(\mathrm{Y}),\omega(\mathrm{Z})] + \mathrm{Y}\cdot[\omega(\mathrm{Z}),\omega(\mathrm{X})] \\ & + \mathrm{Z}\cdot[\omega(\mathrm{X}),\omega(\mathrm{Y})] - [\omega([\mathrm{X},\mathrm{Y}],\omega(\mathrm{Z})] \\ & - [\omega([\mathrm{Y},\mathrm{Z}]),\omega(\mathrm{X})] - [\omega([\mathrm{Z},\mathrm{X}]),\omega(\mathrm{Y})]. \end{aligned}$$

Si l'on considère, comme ci-dessus, que les champs de vecteurs $\mathrm{X}, \mathrm{Y}, \mathrm{Z}$ sont sommes de champs horizontaux invariants par G, $h\mathrm{X}, h\mathrm{Y}, h\mathrm{Z}$ et de champs verticaux images de champs invariants à gauche sur G, on obtient finalement :

$$d\Omega(\mathrm{X},\mathrm{Y},\mathrm{Z}) = [\Omega(\mathrm{X},\mathrm{Y}),\omega(\mathrm{Z})] + [\Omega(\mathrm{Y},\mathrm{Z}),\omega(\mathrm{X})] + [\Omega(\mathrm{Z},\mathrm{X}),\omega(\mathrm{Y})]$$

qui est bien nul si X, Y et Z sont horizontaux.

Les propriétés particulières des formes de connexion et de courbure suggèrent la définition générale suivante :

Définition IX.6.B :

Soient P un espace fibré principal différentiable, de groupe structural G, au-dessus de la variété $\mathfrak{V}$, F un espace vectoriel réel de dimension finie muni d'une représentation linéaire de G, $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ le fibré vectoriel associé.

Une m-forme *pseudotensorielle* φ sur P est une m-forme différentielle extérieure à valeurs dans F qui possède, vis-à-vis des opérations de G, la propriété de covariance : $r(g)^*\varphi = \overset{-1}{g} \cdot \varphi$.

Une m-forme *tensorielle* sur P est une m-forme pseudotensorielle qui est nulle, dès que l'un des arguments est vertical.

Ces définitions sont naturellement *indépendantes* du choix d'une connexion sur P. La connaissance d'une forme pseudotensorielle φ en point p de P la détermine sur toute la fibre : $\varphi_{pg} = r(\overset{-1}{g})^*(\overset{-1}{g} \cdot \varphi_p)$.

S'il existe au-dessus d'un ouvert U de $\mathfrak{V}$ une section différentiable σ_{U} de P, φ est donc déterminée par sa restriction à l'image

$\sigma_U(U)$ dans P, et on peut donc se donner une forme pseudotensorielle φ sur P par ses expressions locales sur $\mathfrak{V}$: $\sigma_U^*(\varphi)$ est une m-forme différentielle extérieure sur U à valeurs dans F.

Si φ et ψ sont des formes pseudotensorielles sur P, φ d'ordre r à valeurs dans F, ψ d'ordre s à valeurs dans H, leur *produit* $\varphi \wedge \psi$ est une $(r+s)$-forme pseudotensorielle à valeurs dans le produit tensoriel $F \otimes H$ muni du produit tensoriel des représentations de G dans F et H. $\varphi \wedge \psi$ est tensorielle si φ et ψ le sont.

Exemples IX.6.A :

1) La forme ω d'une connexion sur P est une 1-forme pseudotensorielle à valeurs dans l'algèbre de Lie $\mathfrak{g}$ de G, puisque $r(g)^*\omega = \mathrm{ad}\,\overset{-1}{g} \cdot \omega$.

2) La forme de courbure Ω d'une connexion sur P est une 2-forme tensorielle. En effet :

$$\begin{aligned}(r(g)^*\Omega)(X,Y) &= \Omega(r(g)_*X, r(g)_*Y)\\ &= -\omega([hr(g)_*X, hr(g)_*Y])\\ &= -\omega[r(g)_*hX, r(g)_*hY])\\ &= -\omega(r(g)_*.[hX,hY])\\ &= -\mathrm{ad}\,\overset{-1}{g} \cdot \omega([hX,hY])\\ &= \mathrm{ad}\,\overset{-1}{g} \cdot \Omega(X,Y).\end{aligned}$$

Proposition IX.6.B :

Une m-forme tensorielle $\breve{\alpha}$ sur le fibré principal P, à valeurs dans F, s'identifie naturellement à une m-forme différentielle extérieure sur $\mathfrak{V}$ à valeurs dans le fibré vectoriel associé $E = P \times_G F$, section du fibré vectoriel $\mathrm{Hom}(\wedge^m T(\mathfrak{V}); E) \equiv E \otimes \wedge^m T(\mathfrak{V})^*$, $\widetilde{\alpha}$ est appelée le *développement* de α.

Preuve : Tout élément p de P peut être interprété comme un isomorphisme de F sur la fibre $E_{\pi p} = p \times_G F$ (p est un « repère » de F). Si $\breve{\alpha}$ est une m-forme tensorielle sur P, il lui correspond $\alpha \in \mathrm{Hom}(\wedge^m T(\mathfrak{V}); E)$ ainsi définie :

Si $X_1, X_2, \dots X_m \in T(\mathfrak{V})_x$, soient $\check{X}_1, \check{X}_2, \dots \check{X}_m \in T(P)_p$ des relèvements quelconques de ces vecteurs dans l'espace tangent de

P en un point p au-dessus de $x(\Pi p = x)$. On définit alors :

$$\alpha(\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_m) = p \cdot \breve{\alpha}_p(\check{\mathrm{X}}_1, \check{\mathrm{X}}_2, \dots \check{\mathrm{X}}_m).$$

Cette définition est indépendant du choix de relèvements en p puisque $\breve{\alpha}$ est tensorielle, et si l'on se place en un autre point pg au-dessus de x, on a :

$$\begin{aligned} pg \cdot \breve{\alpha}_{pg}(r(g)_* \check{\mathrm{X}}_1, \dots r(g)_* \check{\mathrm{X}}_m) &= pg \cdot (r(g)^* \breve{\alpha}_{pg}(\check{\mathrm{X}}_1, \dots \check{\mathrm{X}}_m) \\ &= pg \cdot \overset{-1}{g}\, \breve{\alpha}_p(\check{\mathrm{X}}_1, \dots \check{\mathrm{X}}_m) \\ &= p \cdot \breve{\alpha}_p(\check{\mathrm{X}}_1, \dots \check{\mathrm{X}}_m) \end{aligned}$$

α est donc parfaitement déterminée.

Réciproquement, le même calcul montre que si c'est α qui est donnée, la m-forme extérieure $\breve{\alpha}$ sur P à valeurs dans F définie par :

$$\breve{\alpha}_p(\mathrm{Y}_1, \mathrm{Y}_2, \dots \mathrm{Y}_m) = \overset{-1}{p} \cdot \alpha(\Pi_* \mathrm{Y}_1, \dots \Pi_* \mathrm{Y}_m)$$

est tensorielle.

Exemples IX.6.B :

1) La forme identique θ *de l'espace tangent* $\mathrm{T}(\mathfrak{V})$ d'une variété différentielle $\mathfrak{V}$ est la 1-forme tensorielle sur le fibré P des repères de $\mathrm{T}(\mathfrak{V})$ développement de la 1-forme $\mathfrak{I} \in \mathrm{Hom}(\mathrm{T}(\mathfrak{V}); \mathrm{T}(\mathfrak{V}))$ égale en chaque point x de $\mathfrak{V}$ à l'opérateur identique de $\mathrm{T}(\mathfrak{V})_x$. On a :

$$\theta(\mathrm{X}_p) = \overset{-1}{p} \cdot (\Pi_* \mathrm{X}_p)$$

$$r(g)^* \theta(\mathrm{X}_p) = \theta(r(g)_* \mathrm{X}_p) = (\overset{-1}{pg})(\Pi_*(r(g)_* \mathrm{X}_p)$$

et
$$= \overset{-1}{g}\,\overset{-1}{p} \cdot (\Pi_* t \mathrm{X}_p)$$

soit
$$r(g)^* \theta = \overset{-1}{g} \cdot \theta$$

2) Un champ de vecteurs v du fibré vectoriel $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ peut être considéré comme une 0-forme tensorielle et se relève sur P en une fonction u de P à valeurs dans F : $u_p = \overset{-1}{p} \cdot v$. On a : $u_{pg} = \overset{-1}{g} \cdot u_p$ et v s'écrit : $v = p \cdot u_p$.

Dans le cas de l'espace tangent $\mathrm{T}(\mathfrak{V})$, le développement sur le fibré P des repères d'un champ de vecteurs X sur $\mathfrak{V}$ peut aussi s'exprimer avec la forme identique θ. En effet, si $h\mathrm{X}$ est le champ horizontal invariant par G relevant X sur P, on a : $\theta(h\mathrm{X}_p) = \overset{-1}{p} \cdot \Pi_*(h\mathrm{X}_p) = \overset{-1}{p} \cdot \mathrm{X}$.

De façon analogue, un champ de tenseurs de E, section d'un filtré tensoriel sur E : $\otimes^{(v)}\mathrm{E}$ peut être considéré comme une 0-forme tensorielle et se relève en une fonction sur P à valeurs dans $\otimes^{(p)}\mathrm{F}$, où opère G.

3) Une forme extérieure tensorielle associée à l'espace tangent $\mathrm{T}(\mathfrak{V})$, de type (r, s), est évidemment un champ de tenseurs sur $\mathfrak{V}$, de variance $(r, s+p)$, antisymétrique par rapport aux p derniers indices covariants, qui sont distingués des autres indices par choix. Par exemple une 1-forme différentielle α sur $\mathfrak{V}$ peut aussi être considérée comme une 0-forme tensorielle, à valeurs dans les covecteurs.

4) Lorsque G opère identiquement sur F, $\mathrm{P} \times_{\mathrm{G}} \mathrm{F} \equiv \mathfrak{V} \times \mathrm{F}$ (fin du §IX.5) et les formes tensorielles se réduisent aux formes différentielles extérieures sur $\mathfrak{V}$ à valeurs dans F.

Corollaire IX.6.A : La forme de courbure Ω d'une connexion sur le fibré principal P détermine sur la base $\mathfrak{V}$ de P une 2-forme différentielle extérieure $\mathfrak{R}$ à valeurs dans le fibré vectoriel $\mathrm{L} = \mathrm{P} \times_{\mathrm{G}} \mathfrak{g}$ associé à l'algèbre de Lie $\mathfrak{g}$ munie de la représentation adjointe : $\mathfrak{R} \in \mathrm{Hom}(\wedge^2\mathrm{T}(\mathfrak{V}) : \mathrm{P} \times_{\mathrm{G}} \mathfrak{g}) = (\mathrm{P} \times_{\mathrm{G}} \mathfrak{g}) \otimes \wedge^2\mathrm{T}(\mathfrak{V})^*$. Ω est le développement de $\mathfrak{R}$.

Remarque : Si F est un espace vectoriel réel muni d'une représentation linéaire de G, donc de son algèbre de Lie $\mathfrak{g}$, *l'espace fibré* $\mathrm{L} = \mathrm{P} \times_{\mathrm{G}} \mathfrak{g}$ *opère dans le fibré vectoriel* $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$. En effet, soient $\mathfrak{R} = p \times_{\mathrm{G}} a \in \mathrm{L}_{\Pi p}$ où $a \in \mathfrak{g}$, et $\mathfrak{v} = p \times_{\mathrm{G}} u \in \mathrm{E}_{\Pi p}$ où $u \in \mathrm{F}$. Si l'on définit $\mathfrak{R}\mathfrak{v} = p \times_{\mathrm{G}} au$, cet élément de $\mathrm{E}_{\Pi p}$ ne dépend pas du choix du « repère » p. En effet dans un autre « repère » pg, $\mathfrak{R}$ et $\mathfrak{v}$ s'écrivent : $\mathfrak{R} = pg \times_{\mathrm{G}} \mathrm{ad}\, \overset{-1}{g} \cdot a$ et $\mathfrak{v} = pg \times_{\mathrm{G}} \cdot \overset{-1}{g} u$. Dès lors $pg \times (\mathrm{ad}\, \overset{-1}{g})a \cdot \overset{-1}{g} u = pg \times \overset{-1}{g} ag \overset{-1}{g} u = pg \times_{\mathrm{G}} \overset{-1}{g}(au) = p \times_{\mathrm{G}} au$, c.q.f.d.

On a vu ci-dessus comment l'existence d'une section différentiable σ_{U} de P permet de définir localement les formes pseudotensorielles sur P. C'est ce qui a été fait au paragraphe précédent où l'on a défini les formes locales de connexion ω et de courbure Ω par : $\gamma_{\mathrm{U}} = \sigma_{\mathrm{U}}^* \omega$ et $\Omega_{\mathrm{U}} = \sigma_{\mathrm{U}}^* \Omega$, avec :

$$\Omega_{\mathrm{U}} = \sigma_{\mathrm{U}}^* \Omega = d(\sigma_{\mathrm{U}}^* \omega) + [\sigma_{\mathrm{U}}^* \omega, \sigma_{\mathrm{U}}^* \omega] = d\gamma_{\mathrm{U}} + [\gamma_{\mathrm{U}}, \gamma_{\mathrm{U}}].$$

Ces formes locales $\gamma_{\mathrm{U}}, \Omega_{\mathrm{U}}$ sont, comme ω et Ω, à valeurs dans l'algèbre de Lie $\mathfrak{g}$ qui opère, comme G, dans l'espace vectoriel F, fibre type du fibré vectoriel associé $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$.

La 2-forme extérieure $\mathfrak{R} \in \mathrm{Hom}(\wedge^2 \mathrm{T}(\mathfrak{V}) : \mathrm{P} \times_{\mathrm{G}} \mathfrak{g})$ associé à Ω d'après le corollaire IX.6.A s'écrit à l'aide de la forme locale Ω_{U} :

$$\mathfrak{R}(\mathrm{X}, \mathrm{Y})_x = \sigma_{\mathrm{U}}(x) \Omega(\sigma_{\mathrm{U}*}\mathrm{X}, \sigma_{\mathrm{U}*}\mathrm{Y}) = \sigma_{\mathrm{U}}(x) \Omega_{\mathrm{U}}(\mathrm{X}, \mathrm{Y})$$

$\mathfrak{R}(\mathrm{X}, \mathrm{Y})$ opère dans E d'après la remarque précédente. Le théorème remarquable suivant relie la courbure à la dérivée covariante :

Théorème IX.6.A : Soient v un champ de vecteurs sur $\mathfrak{V}$, section du fibré vectoriel $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$, ω une connexion sur P, Ω sa forme de courbure, D la dérivation covariante associée. Soient U un ouvert de $\mathfrak{V}$ où est définie une section différentiable σ_{U} de P qui permet de représenter v sous la forme : $v = \sigma_{\mathrm{U}} u$, $\gamma_{\mathrm{U}} = \sigma_{\mathrm{U}}^* \omega$ la forme locale de la connexion sur U et $\Omega_{\mathrm{U}} = \sigma_{\mathrm{U}}^* \Omega$. On a alors pour tout couple X, Y de vecteurs tangents à $\mathfrak{V}$ en un point x (cf. §IX.4).

$$\boxed{\mathfrak{R}(\mathrm{X}, \mathrm{Y}) v = \sigma_{\mathrm{U}}(\Omega_{\mathrm{U}}(\mathrm{X}, \mathrm{Y}) u) = (\mathrm{D_X D_Y} - \mathrm{D_Y D_X} - \mathrm{D}_{[\mathrm{X},\mathrm{Y}]}) v}\,.$$

Cette égalité sera démontré de façon directe ci-dessous.

Preuve : On a (fin du §IX.5) :

$$\begin{aligned}
&\mathrm{D_Y} v = \sigma_{\mathrm{U}}[\mathrm{L_Y} u + \gamma_{\mathrm{U}}(\mathrm{Y}) u\}\\
&\mathrm{D_X D_Y} v = \sigma_{\mathrm{U}}\{\mathrm{L_X}(\mathrm{L_Y} u + \gamma_{\mathrm{U}}(\mathrm{Y}) u) + \gamma_{\mathrm{U}}(\mathrm{X}) \cdot (\mathrm{L_Y} \cdot u + \gamma_{\mathrm{U}}(\mathrm{Y}) u)\}\\
&= \sigma_{\mathrm{U}}\{\mathrm{L_X L_Y} u + (\mathrm{L_X} \cdot \gamma_{\mathrm{U}}(\mathrm{Y}) + \gamma_{\mathrm{U}}(\mathrm{Y}) \mathrm{L_X} u + \gamma_{\mathrm{U}}(\mathrm{X}) \mathrm{L_Y} u\\
&+ \gamma_{\mathrm{U}}(\mathrm{X}) \gamma_{\mathrm{U}}(\mathrm{Y}) u\}
\end{aligned}$$

ce qui donne après simplifications :

$$\begin{aligned}
&(\mathrm{D_X D_Y} - \mathrm{D_Y D_X} - \mathrm{D}_{[\mathrm{X},\mathrm{Y}]}) v = \sigma_{\mathrm{U}}\{(\mathrm{L_X L_Y} - \mathrm{L_Y L_X} - \mathrm{L}_{[\mathrm{X},\mathrm{Y}]})\\
&+ (\mathrm{L_X} \gamma_{\mathrm{U}}(\mathrm{Y}) - \mathrm{L_Y} \gamma_{\mathrm{U}}(\mathrm{X}) - \gamma_{\mathrm{U}}([\mathrm{X}, \mathrm{Y}]) + [\gamma_{\mathrm{U}}(\mathrm{X}), \gamma_{\mathrm{U}}(\mathrm{Y})]\} u\\
&= \sigma_{\mathrm{U}}\{d\gamma_{\mathrm{U}}(\mathrm{X}, \mathrm{Y}) + [\gamma_{\mathrm{U}}(\mathrm{X}), \gamma_{\mathrm{U}}(\mathrm{Y})]\} u \quad \text{c.q.f.d.}
\end{aligned}$$

(Le crochet d'opérateurs dans F : $[\gamma_U(X), \gamma_U(Y)]$ est l'image du crochet dans l'algèbre de Lie $\mathfrak{g}$ des éléments $\gamma_U(X)$ et $\gamma_U(Y)$ puisque $\mathfrak{g} \to \mathcal{L}(F)$ est une représentation linéaire de $\mathfrak{g}$ dans F issue de celle de G).

Nous allons finalement généraliser la dérivation covariante des champs de vecteurs et de tenseurs en une dérivation covariante des formes différentielles extérieures tensorielles et pseudotensorielles. Cette nouvelle définition, a le grand avantage d'être *globale*. Elle s'exprime « géométriquement » par la définition ci-dessous, et « algébriquement » par le théorème IX.6.B qui en résulte.

Définition IX.6.C :

Soit φ une m-forme pseudotensorielle sur le fibré G-principal P, à valeurs dans l'espace vectoriel réel F où opère G. La *différentielle covariante* $D\varphi$ de φ associée au choix d'une connexion sur P est alors la $(m+1)$-*forme tensorielle* définie pour $X_0, X_1, \ldots X_m \in T(P)_p$ par :

$$\boxed{D\varphi(X_0, X_1, \ldots X_m) = d\varphi(hX_0, hX_1, \ldots hX_m)}$$

où h désigne le projecteur vertical du plan $T(P)_p$ sur le sous-espace horizontal H_p.

Si φ et ψ sont des formes pseudotensorielles sur P, φ de degré r à valeurs dans F, ψ de degré s à valeurs dans H, la différentielle covariante de leur produit $\varphi \wedge \psi$, de degré $(r+s)$ à valeurs dans $F \otimes H$, satisfait à l'équation classique : $D(\varphi \wedge \psi) = D\varphi \wedge \psi + (-1)^r \varphi \wedge D\psi$.

Exemples IX.6.C :

1) La forme de courbure Ω est la différentielle covariante de la forme ω de la connexion $\Omega(X, Y) = d\omega(hX, hY)$.

2) La seconde identité de Bianchi (Proposition IX.6.A) exprime que la différentielle covariante de Ω est nulle.

3) La différentielle covariante de la 1-forme identique θ de l'espace tangent $T(\mathfrak{V})$, associée au choix d'une connexion sur le fibré principal P des repères de $T(\mathfrak{V})$ est une 2-forme tensorielle Θ appelée *forme de torsion* de la connexion : $\Theta(X, Y) = d\theta(hX, hY)$ dont on verra ci-dessous que le champ de tenseurs correspondant sur $\mathfrak{V}$ est la torsion T de la définition IX.4.B. Dire que la connexion

sur $\mathrm{T}(\mathfrak{V})$ est symétrique revient donc à dire que le champ de tenseurs formé par les opérateurs identiques des espaces tangents est parallèle.

La première démarche à la suite de la définition ci-dessus, doit être de vérifier que pour les 0-formes tensorielles, c'est-à-dire pour les champs de vecteurs ou de tenseurs sur le fibré vectoriel E, on retrouve bien la définition initiale.

Soit donc U un ouvert de $\mathfrak{V}$ sur lequel est définie une section différentiable σ du fibré principal P. Si v est un champ de vecteurs du fibré vectoriel $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$, et u le développement de v sur P, à valeurs dans F : $u_p = \overset{-1}{p} \cdot v$ (Exemple IX.6.B n° 2), on a : $v_x = p \cdot u_p$ et sur U, $v_x = \sigma(x)(u \circ \sigma)_x$.

Soient X un vecteur tangent à $\mathfrak{V}$ en un point x de U et $\sigma_*\mathrm{X}$ son image par σ. Elle se décompose en ses composantes horizontale et verticale : $h\sigma_*\mathrm{X}$ et $v\sigma_*\mathrm{X} = \omega(\widetilde{\sigma}_*\mathrm{X})$ où ω est la forme de la connexion. Or :

$$(\widetilde{\mathrm{A}} \cdot u)_p = \frac{d}{dt}(u_{p \exp \mathrm{A}t})_{t=0} = \frac{d}{dt}(\exp(-\mathrm{A}t) \cdot u_p)_{t=0} = -\mathrm{A}u.$$

On a donc :

$$\begin{aligned}\mathrm{D}u(\sigma_*\mathrm{X}) &= du(h\sigma_*\mathrm{X}) = du(\sigma_*\mathrm{X}) - du(v\sigma_*\mathrm{X}) \\ &= \mathrm{X} \cdot (u \circ \sigma) + \omega(\sigma_*\mathrm{X})u\end{aligned}$$

et $\mathrm{D}_{\mathrm{X}}v = \sigma \mathrm{D}u(\sigma_*\mathrm{X}) = \sigma\{\mathrm{L}_{\mathrm{X}} \cdot (u \circ \sigma) + \omega(\sigma_*\mathrm{X})u\}$

On retrouve bien la définition du § précédent IX.5.

Remarque : Si on relève le vecteur tangent $\mathrm{X} \in \mathrm{T}(\mathfrak{V})_x$ en un vecteur horizontal $(h\mathrm{X})_p \in \mathrm{T}(\mathrm{P})_p$, ce dernier annule la forme de connexion. La *dérivée covariante de v par rapport à* X *sur* $\mathfrak{V}$ s'exprime donc tout simplement sur le développement u de v sur P comme la *dérivée de Lie* par rapport à $(h\mathrm{X})_p$ en un point p quelconque de P :

$$\boxed{\mathrm{D}_{\mathrm{X}}v = p \cdot \{(h\mathrm{X})_p \cdot u\}}.$$

Si maintenant X est un champ de vecteurs sur $\mathfrak{V}$, $\mathrm{D}_{\mathrm{X}}v$ est un champ de vecteurs de E et la formule précédente montre que son développement sur P est le champ $(h\mathrm{X}) \cdot u$.

Applications : Dans ce qui suit $X, Y, X_0, X_1, \ldots$ désignent des champs de vecteurs tangents sur $\mathfrak{V}$, $hX, hY, hX_0, hX_1, \ldots$ leurs relèvements horizontaux invariants par G dans l'espace tangent T(P) du fibré G-principal P sur $\mathfrak{V}$, que l'on suppose muni d'une connexion de forme ω (cf. Lemme IX.6)

1) *Expression de la différentielle covariante d'une m-forme extérieure tensorielle α sur $\mathfrak{V}$:* $\alpha \in E \otimes \wedge^m T(\mathfrak{V})^*$ (où $E = P \times_G F$).

Si φ est le développement de α sur P, $D\alpha$ est la $(m+1)$-forme tensorielle sur $\mathfrak{V}$ dont le développement sur P est $D\varphi$. $D\alpha$ est donc définie, quel que soit $p \in P$, par :

$$\begin{aligned} D\alpha(X_0, X_1, \ldots X_m) &= p \cdot d\varphi(hX_0, hX_1, \ldots hX_m)_p \\ &= p\{\sum_{i<j} (-1)^j (hX_j) \cdot \varphi(hX_0, \ldots h\widehat{X}_j, \ldots hX_m)_p \\ &+ \sum (-1)^{i+j} \varphi(h[X_i, X_j], \ldots \widehat{hX_i}, \ldots \widehat{hX_j}, \ldots hX_m\} \end{aligned}$$

ce qui s'écrit, vue la remarque précédente :

$$\boxed{\begin{aligned} (D\alpha)(X_0, \ldots X_m) = \sum (-1)^j \cdot D_{X_j} \cdot (\alpha(X_0, \ldots, \widehat{X}_j, \ldots X_m)) \\ + \sum_{i<j} (-1)^{i+j} \cdot \alpha([X_i, X_j], \ldots \widehat{X}_i, \ldots \widehat{X}_j, \ldots X_m) \end{aligned}}.$$

Cette égalité généralise l'expression de la différentielle extérieure d'une forme différentielle scalaire.

Si l'on a localement une expression de α à l'aide de m-formes extérieures scalaires, soit :
$\alpha = \sum_k u_k \otimes \omega_k$, où les u_k sont des sections de E alors : $D\alpha = \sum_k (Du_k \wedge \omega_k + u_k \otimes d\omega_k)$ où Du_k est une 1-forme à valeurs dans E et où $Du_k \wedge \omega_k$ est obtenu en effectuant l'antisymétrisation partielle $a_{1,m}$ entre le nouvel indice covariant introduit par la différentielle D et les m indices covariants de ω_k.

2) Si Ω est la forme de courbure de la connexion sur P c'est le développement sur P du tenseur de courbure $\mathfrak{R} \in (P \times_G \mathfrak{g}) \otimes \wedge^2 T(\mathfrak{V})^*$ et

$$\mathfrak{R}(X, Y)_x = p \cdot \Omega(hX_p, hY_p), \quad \forall p \in P_x.$$

Si v est un champ de vecteurs de E, de développement u sur P :

$$\mathfrak{R}(X, Y)v = p\{\Omega(hX_p, hY_p) \cdot u\}.$$

On a, d'après ce que l'on vient de voir :

$$p\cdot\{(hX)_p\cdot(hY)u-(hY)_p(hX)u-(h[X,Y]_p)\cdot u\}$$
$$=D_XD_Yv-D_YD_Xv-D_{[X,Y]}v.$$

Mais $(hX_p\cdot hY-hY_p\cdot hX)u=[hX,hY]_pu$.

La différence $[hX,hY]-h[X,Y]$ est un vecteur vertical qui peut être écrit $\omega(\widetilde{[hX,hY]})$ dont l'opération sur la 0-forme tensorielle u a été calculée au cours de la preuve du théorème IX.6.A et est égale à $\Omega(X,Y)\cdot u$.

On retrouve la formule du théorème IX.6.A mais sans avoir utilisé cette fois, de repères locaux.

3) il n'est pas inutile de remarquer que dans le cas particulier où E est tout simplement l'espace tangent $T(\mathfrak{V})$ et P le fibré principal des repères, la différentielle covariante, DZ d'un champ de vecteurs tangents Z peut être considérée :

a) soit comme une 1-forme tensorielle à valeurs dans $T(\mathfrak{V})$: $X\to D_XZ$ dont la différentielle covariante est alors la 2-forme extérieure tensorielle que l'on vient de voir :

$$(D^2Z)(X,Y)=\mathfrak{R}(X,Y)Z=D_XD_YZ-D_YD_XZ-D_{[X,Y]}Z$$

b) soit comme une 0-forme tensorielle à valeurs dans $\mathfrak{L}(T(\mathfrak{V})=T(\mathfrak{V})\otimes T(\mathfrak{V})^*$. Or si u est une telle 0-forme sa différentielle covariante est $Du=D\circ u-u\circ D$, soit : $(D_Xu)(Y)=D_X(u(Y))-u(D_XY)$ ce qui donne si $u=DZ$: $D(DZ)(X,Y)=D_XD_YZ-D_{D_XY}Z$.

L'antisymétrisation par rapport aux deux arguments s'écrit alors :

$$D(DZ)(X,Y)-D(DZ)(Y,X)=D_XD_YZ-D_YD_XZ-D_{D_XY-D_YX}Z$$

qui se trouve être égal à D^2Z lorsque la connexion est symétrique.

4) P étant toujours le fibré principal des repères de l'espace tangent $T(\mathfrak{V})$ la forme de torsion $\Theta=D\theta$ d'une connexion sur P est le développement de $T\in T(\mathfrak{V})\otimes\wedge^2(\mathfrak{V})^*$, et :

$$T(X,Y)=p\Theta(hX_p,hY_p)=p\cdot d\theta(hX_p,hY_p)$$
$$=p\cdot\{(hX)_p\cdot\theta(hY_p)-(hY)_p\cdot\theta(hX_p)-\theta(h[X_p,Y_p]\}$$

soit $$\boxed{\mathrm{T}(\mathrm{X},\mathrm{Y}) = \mathrm{D}_\mathrm{X}\mathrm{Y} - \mathrm{D}_\mathrm{Y}\mathrm{X} - [\mathrm{X},\mathrm{Y}]}.$$

Le théorème ci-dessous donne une expression algébrique simple de la différentielle covariante.

Théorème IX.6.B : Soient P un fibré principal différentiable de groupe structural G, muni d'une connexion ω, F un espace vectoriel réel muni d'une représentation linéaire de G.

La différentielle covariante $\mathrm{D}\varphi$ *d'une m-forme tensorielle* φ *sur* P à valeurs dans F peut s'écrire :

$$\boxed{\mathrm{D}\varphi = d\varphi + \omega \wedge \varphi},$$

où, dans le produit extérieur $\omega \wedge \varphi$, il est entendu que ω, à valeurs dans $\mathfrak{g}$, algèbre de Lie de G, opère dans F, espace des valeurs de φ. ($d\varphi$ est la différentielle « relative » de φ, $\omega \wedge \varphi$ sa différentielle « d'entraînement »).

Preuve : Soient $\mathrm{X}_0, \mathrm{X}_1, \ldots \mathrm{X}_m$, $(m+1)$ vecteurs tangents en $p \in \mathrm{P}$, $h\mathrm{X}_0, \ldots h\mathrm{X}_m$, et $v\mathrm{X}_0, \ldots v\mathrm{X}_m$ leurs composantes horizontales et verticales, que l'on plonge, les premières dans les champs horizontaux invariants par G, les secondes dans les champs $\widetilde{\mathrm{A}}$ images de champs A invariants à gauche sur G.

$$\begin{aligned} d\varphi(\mathrm{X}_0, \ldots \mathrm{X}_m) = {} & \sum (-1)^j \mathrm{X}_j \cdot \varphi(\mathrm{X}_0, \ldots \widehat{\mathrm{X}}_j, \ldots \mathrm{X}_m) \\ & + \sum_{i<j} (-1)^{i+j} \varphi([\mathrm{X}_i, \mathrm{X}_j], \ldots \widehat{\mathrm{X}}_i, \ldots \widehat{\mathrm{X}}_j, \ldots \mathrm{X}_m). \end{aligned}$$

Or, φ est nulle dès que *l'un* des arguments est vertical. Après décomposition des vecteurs en leurs composantes, il ne reste que les termes suivants :

$$\begin{aligned} d\varphi(\mathrm{X}_0, \ldots \mathrm{X}_m) = {} & \sum (-1)^j (h\mathrm{X}_j) \cdot \varphi(h\mathrm{X}_0, \ldots h\widehat{\mathrm{X}}_j, \ldots h\mathrm{X}_m) \\ & + \sum_{i<j} (-1)^{i+j} \varphi([h\mathrm{X}_i, h\mathrm{X}_j], \ldots \widehat{h}\mathrm{X}_i, \ldots \widehat{h}\mathrm{X}_j, \ldots h\mathrm{X}_m) \\ & + \sum (-1)^j (v\mathrm{X}_j) \cdot \varphi(h\mathrm{X}_0, \ldots h\widehat{\mathrm{X}}_j, \ldots h\mathrm{X}_m). \end{aligned}$$

En effet, d'après le lemme IX.6 $[v\mathrm{X}_i, h\mathrm{X}_j] = 0$, $\forall i, j$.

Mais, si l'on note $\widetilde{\mathrm{A}}_j$ le champ $v\mathrm{X}_j$, on a, puisque φ est tensorielle :

$$\begin{aligned}\widetilde{\mathrm{A}}_j \cdot \varphi(h\mathrm{X}_0, \ldots h\widehat{\mathrm{X}}_j, \ldots h\mathrm{X}_m) &= \frac{d}{dt}\, \varphi(r(\exp \mathrm{A}_j t) h\mathrm{X}_0, \ldots) \\ &= \frac{d}{dt} \exp(-\mathrm{A}_j t) \cdot \varphi(h\mathrm{X}_0, \ldots) \\ &= -\mathrm{A}_j \cdot \varphi(h\mathrm{X}_0, \ldots) \\ &= -\omega(\mathrm{X}_j) \cdot \varphi(h\mathrm{X}_0, \ldots \widehat{h\mathrm{X}_j}, \ldots h\mathrm{X}_m).\end{aligned}$$

On obtient donc finalement :

$$d\varphi(\mathrm{X}_0, \ldots \mathrm{X}_m) = d\varphi(h\mathrm{X}_0, \ldots h\mathrm{X}_m) - (\omega \wedge \varphi)(\mathrm{X}_0, \ldots \mathrm{X}_m) \quad \text{c.q.f.d.}$$

Remarque : Lorsque G opère identiquement sur F, la représentation associée de $\mathfrak{g}$ est nulle, et $\mathrm{D}\varphi = d\varphi$.

Corollaire IX.6.B : La différentielle covariante seconde d'une m-forme tensorielle φ sur P ou α sur $\mathfrak{V}$ s'écrit :

$$\boxed{\mathrm{D}^2\varphi = \Omega \wedge \varphi}, \qquad \boxed{\mathrm{D}^2\alpha = \mathfrak{R} \wedge \alpha}.$$

Preuve : $d(\mathrm{D}\varphi) = d\omega \wedge \varphi - \omega \wedge d\varphi$.

La différentielle covariante $\mathrm{D}(\mathrm{D}\varphi)$ s'obtient en prenant la valeur de $d(\mathrm{D}\varphi)$ sur les composantes horizontales des arguments. Puisque celles-ci annulent ω, on obtient finalement :

$$\begin{aligned}(\mathrm{D}^2\varphi)(\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_{m+2}) &= (d\omega \wedge \varphi)(h\mathrm{X}_1, h\mathrm{X}_2, \ldots h\mathrm{X}_{m+2}) \\ &= (\Omega \wedge \varphi)(\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_{m+2}), \ \text{c.q.f.d.}\end{aligned}$$

On aurait pu aussi remarquer, comme on l'a fait ci-dessus, que $\omega \wedge \omega$ est la représentation linéaire de $[\omega, \omega]$, soit $\omega \wedge \omega \wedge \varphi = [\omega, \omega] \wedge \varphi$.

Corollaire IX.6.C : Première identité de Bianchi, vérifiée par la courbure d'une connexion sur le fibré tangent à une variété : la différentielle covariante $\mathrm{D}\Theta$ de la forme de torsion Θ (Exemple

IX.6.C n° 2), elle-même différentielle covariante de la forme identique θ, s'écrit :

$$\boxed{D\Theta = D^2\theta = \Omega \wedge \theta},$$

soit
$$\boxed{\begin{gathered} D\Theta(X,Y,Z) = \Omega(X,Y)\theta(Z) + \Omega(Y,Z)\theta(X) \\ +\Omega(E,X)\theta(Y) \end{gathered}}$$

Lorsque la connexion sur $T(\mathfrak{V})$ est symétrique ($\Theta = 0$), la première et la seconde identité de Bianchi s'expriment sur $\mathfrak{V}$ à l'aide du tenseur de courbure $\mathfrak{R} \in (P \times_G \mathfrak{g}) \otimes T(\mathfrak{V})^*$ par les formules suivantes :

$$\boxed{\begin{gathered} (\mathfrak{R}(X,Y)Z + \mathfrak{R}(Y,Z)X + \mathfrak{R}(Z,X)Y = 0 \\ (D_X\mathfrak{R})(Y,Z) + (D_Y\mathfrak{R})(Z,X) + (D_Z\mathfrak{R})(X,Y) = 0) \end{gathered}}.$$

Terminons ce paragraphe avec l'expression locale sur $\mathfrak{V}$ de la dérivation covariante Df d'une m-forme tensorielle $f \in E \otimes \wedge^m T(\mathfrak{V})^* = (P \times_G \mathfrak{g}) \otimes \wedge^m(T\mathfrak{V})^*$. Soient σ_U une section de P au-dessus de l'ouvert U de $\mathfrak{V}$, définissant un champ de « repères » sur U, $\gamma_U = \sigma_U^*\omega$ la forme locale de la connexion $f_U = \overset{-1}{\sigma}_U \cdot f$, l'expression de f dans ce champ de repères ; f_U est donc une m-forme extérieure sur U à valeur dans F, et si φ est le développement de f sur P, $f_U = \varphi \circ \sigma_U$.

La dérivée covariante $D\varphi$ de φ sur P est une $(m+1)$-forme tensorielle à valeurs dans F, développement de la $(m + 1$-forme $Df \in E \otimes \wedge^{m+1}T(\mathfrak{V})^*$ sur $\mathfrak{V}$, soit :

$$\begin{aligned} Df &= \sigma_U \cdot D_U f = \sigma_U\{D\varphi \circ \sigma_U\} \\ &= \sigma_U\{(d\varphi + \omega \wedge d\varphi) \circ \sigma_U\} \\ &= \sigma_U\{df_U + \gamma_U \wedge f_U\} \end{aligned}$$

d'où :
$$\boxed{Df_U + \gamma_U \wedge f_U},$$

qui a la forme standard de la dérivation covariante rencontrée dans tous les cas précédents.

Si Ω est la 2-forme de courbure sur P de la connexion, et $\mathfrak{R}_U = \Omega \circ \sigma_U$, on a $D^2 f_U = \mathfrak{R}_U \wedge f_U$.

IX.7 – Une application à la physique : les « théories de jauge »

En physique, les diverses interactions : électromagnétiques, faibles, fortes, présentent localement des symétries, dites « symétries internes ». Si G est un groupe de symétries internes, la description mathématique des phénomènes physiques correspondants se fait alors au moyen d'un espace fibré principal P ayant pour base l'espace-temps, et pour groupe G. Une section locale de P est l'équivalent de la fixation en chaque point d'un repère de référence. Si F est un espace vectoriel muni d'une représentation linéaire de G et E l'espace fibré vectoriel $\mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ associé à P, une section de E est la représentation mathématique d'un « champ de matière » : fonction d'onde de particules, $e, \mu, \tau, \nu_e, \nu_\mu, \nu_\tau$, etc..., ou de leurs antiparticules.

On représente le *potentiel d'interaction* par la forme α d'une connexion sur P, à valeurs dans l'algèbre de Lie $\mathfrak{g}$ du groupe de symétrie G. Le *champ* ou *force d'interaction est alors la forme de courbure* Ω également à valeurs dans $\mathfrak{g}$.

Les *quanta* du potentiel d'interaction α sont les particules qui portent l'interaction (photons, bosons intermédiaire, gluons). La contribution de l'interaction au lagrangien du système s'exprime par le demi-carré de la longueur de la courbure de la connexion : $\dfrac{1}{2}\,\|\Omega\|^2$.

Exemples :

1) Dans le cas du champ électromagnétique, la plus simple des théories de jauge, le groupe de symétrie est $\mathrm{U}(1)$. Il est commutatif et a pour algèbre de Lie $\mathbb{R}$. Une connexion du fibré principal est ainsi une forme différentielle ordinaire α, à valeurs dans $\mathbb{R}$: c'est le potentiel électromagnétique. La courbure $\omega = d\alpha$ est le champ électromagnétique. Le Lagrangien est $\mathrm{L} = \dfrac{1}{2}\displaystyle\int \|\omega\|^2$ les équations de Maxwell sont l'équation de Bianchi $d\omega = 0$ et l'équation de Lagrange $\delta\omega = 0$ qui exprime que L est stationnaire : $\delta\mathrm{L} = 0$.

2) Dans la théorie de Salam-Weinberg du couplage entre interactions faibles et électromagnétiques, le groupe de symétrie est $\mathrm{U}(2)$. Son centre $\mathrm{U}(1)$ représente les interactions électromagnétiques et le quotient $\mathrm{SU}(2)$ les interactions faibles. Les trois dimension de l'algèbre de Lie $\mathrm{SU}(2)$ correspond aux trois bosons intermédiaires échangés dans ces interactions.

3) Dans la chromodynamique quantique qui se veut une théorie des interactions fortes, le groupe de symétrie est $\mathrm{SU}(3)$.

Les applications à la physique conduisent ainsi à la classification de fibrés principaux munis d'une connexion telle que l'intégrale $\int \mathrm{L} = \frac{1}{2}\int \|\Omega\|^2$ du demi carré de la norme de la forme de courbure soit stationnaire. La forme de courbure Ω satisfait alors aux deux équations : $d\Omega = 0$ (la seconde identité de Bianchi) et $\delta\Omega = 0$ (équation de Lagrange) où δ est l'opérateur adjoint de d. On peut exprimer ces équations en disant que la forme de coubure est harmonique. Ce sont des modèles, au moins théoriques, de systèmes dynamiques selon les théories de jauge. Lorsque $\mathrm{G} = \mathrm{SU}(2)$ et la variété de base, la sphère S_4, ils ont été décrits et classés sous le nom d'instantons à l'aide d'un invariant topologique qui est le premier nombre de Pontrjagin.

IX.8 – Tenseurs de courbure

Soient P un espace fibré principal différentiable de groupe G, de base la variété $\mathfrak{V}$, F un espace vectoriel réel de dimension finie où G opère linéairement et $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$ le fibré vectoriel associé. On suppose donnée une connnexion sur P, de forme ω, à valeurs dans l'algèbre de Lie $\mathfrak{g}$ de G.

Si U est un ouvert de $\mathfrak{V}$ où sont définies des coordonnées locales $(x^1, x^2, \dots x^n)$ ainsi qu'une section σ de P (champ local de « repères » de E), la forme locale $\gamma = \sigma^*\omega$ de la connexion est une 1-forme différentielle sur U à valeurs dans $\mathfrak{g}$ qui opère sur F. On peut ainsi considérer γ comme une 1-forme à valeurs dans $\mathfrak{L}(\mathrm{F})$ qui peut s'écrire : $\gamma = \sum \Gamma_r dx^r$, $\Gamma_r \in \mathfrak{L}(\mathrm{F})$.

Relativement à une base fixée $(\varepsilon_1, \varepsilon_2, \dots \varepsilon_p)$ de F, et à la base duale $(\lambda^1, \lambda^2, \dots, \dots \lambda^p)$ de F^* (cf. §IX.4), les Γ_r sont des matrices et $\gamma = \sum \Gamma^k_{jr} \varepsilon_k \otimes \lambda^j \otimes dx^r$. La forme locale de courbure est $\Omega_{\mathrm{U}} = \sigma^*\Omega = d\gamma + [\gamma, \gamma]$ avec :

$$\begin{aligned} d\gamma &= \sum \varepsilon_k \otimes \lambda^j \otimes d\Gamma^k_{jr} \wedge dx^r \\ &= \sum \varepsilon_k \otimes \lambda^j \otimes \left(\frac{\partial \Gamma^k_{js}}{\partial x^r} - \frac{\partial \Gamma^k_{jr}}{\partial x^s} \right) dx^r \wedge ds^s \end{aligned}$$

$$\begin{aligned}[\gamma,\gamma] &= \sum(\Gamma_r\Gamma_s - \Gamma_s\Gamma_r)dx^r \wedge dx^s \\ &= \sum \varepsilon_k \otimes \lambda^j \otimes \sum \Gamma^k_{lr}\Gamma^l_{js} - \sum \Gamma^k_{ms}\Gamma^m_{jr})dx^r \wedge dx^s.\end{aligned}$$

On obtient ainsi les *composantes locales du tenseur de courbure :* $\mathfrak{R} = \sum \mathfrak{R}_{k,jrs}\varepsilon_k \otimes \lambda^j \otimes (dx^r \wedge dx^s)$ en fonction des coefficients Γ de la connexion :

$$\boxed{\mathfrak{R}^k_{jrs} = \left(\frac{\partial \Gamma^k_{js}}{\partial x^r} - \frac{\partial \Gamma^k_{jr}}{\partial x^s} + \sum \Gamma^k_{lr}\Gamma^l_{js} - \sum \Gamma^k_{ms}\Gamma^m_{jr}\right)}.$$

Sur $\mathfrak{V}$, la 2-forme tensorielle de courbure $\mathfrak{R}$, dont Ω est le développement sur P, est une section différentiable du fibré $(\mathrm{P} \times_{\mathrm{G}} \mathfrak{g}) \otimes \wedge^2 \mathrm{T}(\mathfrak{V})^*$, qui opère sur $\mathrm{E} = \mathrm{P} \times_{\mathrm{G}} \mathrm{F}$. A tout champ de vecteurs v de E correspond ainsi une 2-forme tensorielle $\mathfrak{R} \cdot v \in \mathrm{E} \otimes \wedge^2 \mathrm{T}(\mathfrak{V})^*$.

Supposons maintenant qu'il existe un produit scalaire $(|)$ sur le fibré vectoriel E. A tout couple de champs de vecteurs tangents X, Y sur $\mathfrak{V}$, et à tout couple de champs de vecteurs u, v de E, on peut associer la fonction à valeurs réelles :
$\mathfrak{K}(\mathrm{X}, \mathrm{Y}, u, v) = (\mathfrak{R}(\mathrm{X}, \mathrm{Y})u \mid v)$.

Lorsque la connexion sur E conserve le produit scalaire, l'opérateur de courbure est à valeurs dans l'algèbre de Lie orthogonale $\mathfrak{g}_0$. les opérateurs $\mathfrak{R}(\mathrm{X}, \mathrm{Y})$ sont antisymétriques relativement au produit scalaire (cf. Théorème IX.5). La fonction $\mathfrak{K}$ est donc alternée dans chaque couple d'arguments (X, Y) et (u, v).

Lorsque E est l'espace tangent $\mathrm{T}(\mathfrak{V})$ d'une variété *riemannienne ou pseudoriemannienne muni de sa connexion canonique* (théorème IX.4.C) on a déjà indiqué que des raisons géométriques (signe de la courbure scalaire) conduisent à utiliser l'opposé : $\mathrm{R} = -\mathfrak{R}$ de l'opérateur de courbure, soit en fonction de la dérivation covariante : $\mathrm{R}(\mathrm{X}, \mathrm{Y}) = \mathrm{D}_{[\mathrm{X},\mathrm{Y}]} - (\mathrm{D}_\mathrm{X}\mathrm{D}_\mathrm{Y} - \mathrm{D}_\mathrm{Y}\mathrm{D}_\mathrm{X})$
$\mathrm{R} \in \mathrm{T}(\mathfrak{V}) \otimes \mathrm{T}(\mathfrak{V})^* \otimes \wedge^2(\mathfrak{V})^*$ est donc un champ de tenseurs 1 fois contravariant, 3 fois covariant, antisymétrique relativement aux deux derniers indices covariants. Le tenseur 4 fois covariant associé, obtenu en contractant l'indice contravariant avec le tenseur métrique g, est couramment appelé le *tenseur de courbure de Riemann* ou mieux *tenseur covariant de Riemann.*

C'est donc la forme quadrilinéaire K définie par :

$$\mathrm{K}(\mathrm{X},\mathrm{Y},\mathrm{U},\mathrm{V}) = (\mathrm{R}(\mathrm{X},\mathrm{Y})\mathrm{U} \mid \mathrm{V})$$

K est alterné par rapport à (X,Y) et aussi par rapport à (U,V) puisque $\mathrm{R}(\mathrm{X},\mathrm{Y})$ est antisymétrique.

Si F est une forme quadrilinéaire, alternée par rapport à ses deux premiers et par rapport à ses deux derniers arguments, l'expression :

$$\mathrm{A_F}(\mathrm{X},\mathrm{Y},\mathrm{U},\mathrm{V}) = \mathrm{F}(\mathrm{X},\mathrm{Y},\mathrm{U},\mathrm{V}) + \mathrm{F}(\mathrm{Y},\mathrm{U},\mathrm{X},\mathrm{V}) + \mathrm{F}(\mathrm{U},\mathrm{X},\mathrm{Y},\mathrm{V})$$

compte tenu des antisymétries partielles, est l'antisymétrisée de la forme F (antisymétrisation « minimum ») : $\mathrm{A_F}$ est ainsi une forme quadrilinéaire alternée.

La première identité de Bianchi.

$$\mathrm{K}(\mathrm{X},\mathrm{Y},\mathrm{U},\mathrm{V}) + \mathrm{K}(\mathrm{Y},\mathrm{U},\mathrm{X},\mathrm{V}) + \mathrm{K}(\mathrm{U},\mathrm{X},\mathrm{Y},\mathrm{V}) = 0$$

exprime que l'antisymétrisée de K est nulle : le tenseur K appartient en chaque point x de $\mathfrak{V}$ au noyau de l'application d'antisymétrisation A :

$$\wedge^2(\mathrm{T}(\mathfrak{V})^* \otimes \wedge^2\mathrm{T}(\mathfrak{V})^* \to \wedge^4\mathrm{T}(\mathfrak{V})^*.$$

Cette identité a pour conséquence une propriété supplémentaire de symétrie pour K : symétrie relative aux deux couples d'arguments : $\mathrm{K}(\mathrm{X},\mathrm{Y},\mathrm{U},\mathrm{V} = \mathrm{K}(\mathrm{U},\mathrm{V},\mathrm{X},\mathrm{Y})$.

La démonstration se fait par une manipulation évidente. De :

(1) $\underline{\mathrm{K}(\mathrm{XYZU})} + (\mathrm{YZXU}) + \mathrm{K}(\mathrm{ZXYU}) = 0$

(2) $\mathrm{K}(\mathrm{YZUX}) + \underline{\mathrm{K}(\mathrm{ZUYX})} + \mathrm{K}(\mathrm{UYZX}) = 0$

(3) $\underline{\mathrm{K}(\mathrm{ZUXY})} + \mathrm{K}(\mathrm{UXZY}) + \mathrm{K}(\mathrm{XZUY}) = 0$

(4) $\mathrm{K}(\mathrm{UXYZ}) + \underline{\mathrm{K}(\mathrm{XYUZ})} + \mathrm{K}(\mathrm{YUXZ}) = 0$

résulte l'égalité : $(1) + (2) - (3) - (4) = 0$ qui, compte tenu du caractère alterné de K dans chaque couple d'arguments, se simplifie en : $2\{\mathrm{K}(\mathrm{XYZU}) - \mathrm{K}(\mathrm{ZUXY})\} = 0$.

Il en résulte qu'en un point donné x de $\mathfrak{V}$, K détermine canoniquement une forme bilinéaire symétrique sur l'espace vectoriel $\wedge^2\mathrm{T}(\mathfrak{V})_x$. Une telle forme est entièrement déterminée par la forme quadratique correspondante, donc par la fonction :

$$(\mathrm{X},\mathrm{Y}) \in \mathrm{T}(\mathfrak{V})_x \to \mathrm{K}(\mathrm{XY},\mathrm{X},\mathrm{Y}) = (\mathrm{R}(\mathrm{XY})\mathrm{X} \mid \mathrm{Y})$$

appelée *courbure sectionnelle* du bivecteur $X \wedge Y$. Plus précisément, on appelle courbure sectionnelle du plan $H \subset T(\mathfrak{V})_x$ le scalaire $K(H) = K(e_1, e_2, e_1, e_2)$ où (e_1, e_2) est une base orthonormée arbitraire de H (toutes les bases orthonormales de H définissent au signe près le même bivecteur $e_1 \wedge e_2$ de $\wedge^2 T(\mathfrak{V})_x$).

Les composantes du tenseur précédemment calculées s'écrivent dans le cas d'une variété pseudoriemannienne (avec $\mathbb{R} = -\mathfrak{R}$).

$$R^i_{jkl} = \frac{\partial \Gamma^i_{jk}}{\partial x^l} - \frac{\partial \Gamma^i_{jl}}{\partial x^k} + \sum \Gamma^m_{jk}\Gamma^i_{ml} - \sum \Gamma^m_{jl}\Gamma^i_{mk} = \mathfrak{R}\Big(\frac{\partial}{\partial x^k} \cdot \frac{\partial}{\partial x^l}\Big)^i_j.$$

L'ancienne géométrie différentielle les appelait les « symboles de Riemann de seconde espèce », exprimés en fonction des « symboles de Christoffel de seconde espèces » :

$$\Gamma^k_{ij} = \sum g^{kr}\Gamma_{ijr} = \frac{1}{2}\sum g^{kr}\Big(\frac{\partial}{\partial x^i}g_{jr} + \frac{\partial}{\partial x^j}g_{ir} - \frac{\partial}{\partial x^r}g_{ij}\Big)$$

(cf. théorème IX.4.C).

Les composantes du tenseur K (« symboles de Riemann de première espèce ») sont les composantes quatre fois covariantes du tenseur de courbure : $R_{ijkl} = \sum g_{im}R^m_{jkl}$ dont on peut donner des expressions variées en fonction de la métrique, par exemple :

$$R_{ijkl} = \frac{\partial}{\partial x^l}\Gamma_{ikj} - \frac{\partial}{\partial x^k}\Gamma_{ilj} + \sum \Gamma^m_{il}\Gamma_{jkm} - \sum \Gamma^m_{ik}\Gamma_{jlm}$$

Les propriétés de symétrie de K s'expriment sur les composantes :

$$R_{ijkl} = \Big(R\Big(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\Big)\frac{\partial}{\partial x^j} \;\Big|\; \frac{\partial}{\partial x^i}\Big),$$

par : $R_{ijkl} = -R_{jikl}$, $R_{ijkl} = -R_{ijkl}$, $R_{ijkl} = R_{klij}$

$$R_{ijkl} + R_{jkil} + R_{kijl} = 0$$

et entraînent qu'il n'y a que $\dfrac{n^2(n^2-1)}{12}$ composantes distinctes de K ($n = \dim \mathfrak{V}$), soient, si $n = 2$, une, si $n = 3$, six, $n = 4$, vingt.

Le tenseur R n'a, au signe près qu'un seul contracté non trivial. En effet, la contraction de deux premiers ou des deux derniers indices donne zéro à cause de l'antisymétrie et il ne reste que la contraction d'un indice du premier couple avec un indice du second, ces quatre contractés étant les mêmes au signe près.

Rappelons la correspondance entre arguments et indices du tenseur covariant de Riemann :

$$\begin{array}{rcccccc} \mathrm{K}(\mathrm{X},\mathrm{Y},\mathrm{Z},\mathrm{U}) = (\mathrm{R}(& \mathrm{X}, & \mathrm{Y}) & \mathrm{Z} & \mid & \mathrm{U}) \\ & \downarrow & \downarrow & \downarrow & & \downarrow \\ \textit{indices}: & 3^{\mathrm{e}} & 4^{\mathrm{e}} & 2^{\mathrm{e}} & & 1^{\mathrm{er}} \end{array}$$

Définition IX.8.A :

Le *tenseur de Ricci* S d'une variété riemannienne ou pseudo-riemannienne $\mathfrak{V}$ munie de sa connexion canonique est le contracté du tenseur de courbure sur le premier et le quatrième indice (ou le second et le troisième). Si $e = (e_1, e_2, \dots e_n)$ est une base locale des champs de vecteurs :

$$\begin{aligned} \mathrm{S}_{jk} &= \sum \mathrm{R}^i_{jki} = \sum g^{il}\mathrm{R}_{ijkl} = \sum g^{il}\mathrm{R}_{jilk} \\ &= \sum g^{il}\mathrm{R}_{klij} = \sum g^{li}\mathrm{R}_{klij} = \mathrm{S}_{kj} \end{aligned}$$

S est donc un champ de tenseurs covariants d'ordre deux symétriques. Intrinsèquement :

$$\mathrm{S}(\mathrm{X},\mathrm{Y}) = \mathrm{Trace}(\mathrm{H} \to \mathrm{R}(\mathrm{X},\mathrm{H})\mathrm{Y}) = \sum g^{ij}(\mathrm{R}, e_j)\mathrm{Y} \mid e_i)$$

Si $(e_1, e_2, \dots e_n)$ est une base orthogonale de l'espace tangent en un point x : $\mathrm{S}(\mathrm{X},\mathrm{Y})_x = \sum\limits_m g^{mm}\mathrm{K}(e_m, \mathrm{X}, \mathrm{Y}, e_m)$.

La *courbure scalaire* s de $\mathfrak{V}$ est la fonction numérique sur $\mathfrak{V}$ contractée du tenseur de Ricci : $s(x) = (\sum g^{jk}\mathrm{S}_{jk})_x$.

$\mathfrak{V}$ est dite *d'Einstein* si le tenseur de Ricci est porportionnel au tenseur métrique : $\mathrm{S} = \lambda g$, et λ est alors nécessairement constant (voir le Corollaire de la Proposition IX.8.A ci-dessous).

Proposition IX.8.A

La divergence div S du tenseur de Ricci est égale à la moitié de la différentielle de la courbure scalaire : $\mathrm{div}\ \mathrm{S} = 1/2ds$ (cf. Définition IX.4.C).

Preuve : Dans l'identité de Bianchi :

$$(\mathrm{D_X R(YZ)U} \mid \mathrm{V}) + (\mathrm{D_Y R(Z,X)U} \mid \mathrm{V}) + (\mathrm{D_Z R(X,Y)U} \mid \mathrm{V}) = 0$$

on effectue la contraction sur les deux indices représentés par X et V, c'est-à-dire qu'on remplace X par e_j, V par e_i qu'on multiplie par g^{ij} et qu'on somme sur i et j. Le second et le troisième terme donnent : $\sum g^{ij}(\mathrm{D_Y R}(\mathrm{Z}, e_j)\mathrm{U} \mid (e_j) = \mathrm{D_Y} \cdot \{\sum g^{ij}(\mathrm{R}(\mathrm{Z}, e_j)\mathrm{U} \mid e_i)\}$ puisque $\mathrm{D_Y}$ commute avec les contractions. De même :

$$\sum g^{ij}(\mathrm{D_Z R}(e_j, \mathrm{Y})\mathrm{U} \mid e_i) - \sum g^{ij}(\mathrm{D_Z R}(\mathrm{Y}, e_j)\mathrm{U} \mid e_i) = -\mathrm{D_Z S}(\mathrm{Y}, \mathrm{U}).$$

Le premier terme est la divergence du champ R :

$$\sum g^{ij}(\mathrm{D}_{e_j}\mathrm{R}(\mathrm{Y}, \mathrm{Z})\mathrm{U} \mid e_i) = (\mathrm{div\ R})(\mathrm{Y}, \mathrm{Z}, \mathrm{U}).$$

D'où : $\quad (\mathrm{div\ R})(\mathrm{Y}, \mathrm{Z}, \mathrm{U}) + \mathrm{D_Y S}(\mathrm{Z}, \mathrm{U}) - \mathrm{D_Z S}(\mathrm{Y}, \mathrm{U}) = 0$

On va maintenant contracter une nouvelle fois sur le couple d'indices représenté par Z et U, en remplaçant Z par e_k, U par e_l, en multipliant par g^{kl} et en sommant :

$$\sum g^{kl} g^{ij}(\mathrm{D}_{e_j}\mathrm{R}(\mathrm{Y}, e_k)e_l \mid e_j) = -\sum g^{ij}\mathrm{D}_{e_i}\mathrm{S}(\mathrm{Y}, e_j) = -\mathrm{div\ S}(\mathrm{Y})$$

$$\sum g^{kl}\mathrm{D_Y S}(e_k, e_l) = \mathrm{D_Y S} = ds(\mathrm{Y})$$

d'où $\quad 2\ \mathrm{div\ S} = ds(\mathrm{Y}) \quad \text{et} \quad \mathrm{div\ S} = \dfrac{1}{2}\, ds$

Définition IX.8.B :

On appelle *tenseur d'Einstein* sur une variété pseudoriemannienne $(\mathfrak{V}, g)$ le tenseur symétrique d'ordre deux : $\mathrm{G} = \mathrm{S} - \dfrac{1}{2}\, sg$ (G comme gravitation). Sa divergence est nulle (voir ci-dessous la preuve du Corollaire IX.8). *L'équation d'Einstein* : $\mathrm{G} = \chi\mathrm{T}$, consiste à égaler G au tenseur d'énergie $\chi\mathrm{T}$. T est un tenseur symétrique du second ordre de divergence nulle, cette dernière propriété exprimant la conservation de l'énergie.

Une *solution de l'équation d'Einstein* est un couple formé d'une métrique g et d'un champ de tenseurs T satisfaisant à l'équation d'Einstein.

Corollaire IX.8 : Si dans une variété pseudoriemannienne connexe de dimension $n > 2$, le tenseur de Ricci est proportionnel en chaque point au tenseur métrique : $\mathrm{S} = \lambda g$, la fonction numérique λ est constante.

Les solutions de l'équation d'Einstein dans le vide ($\mathrm{T} = 0$) sont donc les variétés d'Einstein (définition IX.8.A).

Preuve : La contraction sur les deux indices de $\mathrm{S} = \lambda g$ donne : $s = n\lambda$, et la divergence s'écrit $\operatorname{div} \mathrm{S} = (\operatorname{div}(\lambda g)$. Or : $\operatorname{div}(\lambda g)(\mathrm{Y}) = \sum g^{ij} d\lambda(e_i) g(e_j, \mathrm{Y}) = \sum g^{ij} d\lambda(e_i) g_{jk} \mathrm{Y}^k$, soit, puisque $\sum_j g^{ij} g_{jk} = \delta^i_k$, $\operatorname{div}(\lambda g)(\mathrm{Y}) = \sum d\lambda(e_i) \cdot \mathrm{Y}^i = d\lambda(\mathrm{Y})$ d'où : $\operatorname{div} \mathrm{S} = d\lambda = \frac{1}{n}\, ds = \frac{1}{2}\, ds$, ce qui entraîne $ds = 0$ si $n > 2$.

Quelques remarques sur la relativité générale :

L'équation d'Einstein a été découverte simultanément par Einstein et Hilbert en 1915 en suivant des chemins différents :

1) Raisonnement d'Einstein « par analogie » : la loi d'attraction de Newton $\mathrm{F} = \gamma\, \frac{mm'}{r^2}$ a pour conséquence que le champ de gravitation est le gradient d'un potentiel U déterminé par la répartition des masses. Elle a pour expression locale au sein d'une distribution de matière, de densité ρ, l'équation de Poisson : $\Delta \mathrm{U} = 4\pi\gamma\rho$ qui se réduit dans le vide à l'équation de Laplace $\Delta \mathrm{U} = 0$ (Δ est le laplacien scalaire). L'équation de la gravitation en relativité générale doit avoir une expression invariante par tout changement de coordonnées, c'est-à-dire tensorielle, et se réduire à la limite dans l'univers galiléen à l'équation de Poisson. Cela entraîne que le tenseur de gravitation doit être une fonction de la métrique g ne faisant intervenir que ses dérivées jusqu'au deuxième ordre, et être linéaire par rapport à ses dernières. On démontre (E. Cartan, Journal de Mathématiques 1922, pp. 141, 203) que les seuls tenseurs du second ordre satisfaisant à cette condition, et de divergence nulle, sont de la forme : $\mathrm{S} - \frac{1}{2}\, g(s + \wedge)$ où $\wedge$ est une constante dite *constante cosmologique*. Elle est de dimension L^{-2} et en relativité générale, estimée, inférieure à 10^{-54} cm^{-2}, d'où le choix initial d'Einstein $\wedge = 0$.

La constante χ de l'équation d'Einstein : $\mathrm{G}_{jk} = \mathrm{S}_{jk} - \frac{1}{2}\, s g_{jk} = \chi \mathrm{T}_{jk}$ est la *constante de gravitation d'Einstein*, égale à : $\chi = \frac{8\pi\gamma}{c^4}$ (en unités CGS : $2{,}073 \times 10^{-48}$ $\mathrm{sec}^2 \mathrm{cm}^{-1} \mathrm{gm}^{-1}$) ou en unité adaptées ($\gamma = 1$, $c = 1$) : $\chi = 8\pi$.

Dans le même ordre d'idées (analogie) la comparaison des formules exprimant la déformation gravitationnelle (marées) dans la dynamique newtonienne et la déformation des géodésiques en géométrie riemannienne (cf. n° suivant IX.9) conduit également au tenseur d'Einstein.

2) Raisonnement direct de Hilbert.

Il est naturel de rechercher, à l'exemple des théories classiques « de la moindre action », une formulation lagrangienne de la relativité générale. Hilbert a proposé, comme intégrale d'action, l'intégrale de la courbure scalaire s : $\mathrm{L}(g) = \int s\omega$, qui s'écrit dans des coordonnées locales $\int s|g|^{1/2}dx^1dx^2\dots dx^n$ (ici $n = 4$!). Si l'on prend une famille à un paramètre de métriques : $t \to g(t)$ issue de $g(0) = g$, on obtient :

$$\frac{d\mathrm{L}}{dt} = \int \operatorname{div} \mathrm{V} \cdot \omega + \int (\mathrm{S}_{jk} - \frac{1}{2}\, sg_{jk})\, \frac{dg^{jk}}{dt}\, |g|^{1/2}dx^1\dots dx^n$$

où V est un champ de vecteurs dépendant de la métrique. C'est le tenseur d'Einstein qui apparaît sous le signe somme. Pour obtenir l'équation de Einstein complète il suffit d'ajouter au lagrangien $s\omega$ un lagrangien représentant la matière et l'énergie.

Les propriétés de symétrie du tenseur covariant de Riemann et ses relations avec les contractés : tenseur de Ricci et courbure scalaire, conduisent à examiner de plus près les décompositions naturelles de certains espaces de tenseurs restreints sur un espace pseudoeuclidien E, de dimension n, et simultanément sur une variété pseudoriemannienne $\mathfrak{V}$. Une *forme double symétrique de degré p*, sur E ou sur $\mathfrak{V}$, est une fonction $\omega(\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_p; \mathrm{Y}_1, \mathrm{Y}_2, \dots \mathrm{Y}_p)$ de deux p-uples de vecteurs (de E, ou tangents sur $\mathfrak{V}$) satisfaisant aux trois propriétés :

1) ω est linéaire ($\mathbb{R}$-linéaire sur E, $\mathfrak{F}$-linéaire sur $\mathfrak{V}$) en chacun de ses arguments.

2) ω est antisymétrique en $\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_p$, et antisymétrique en $\mathrm{Y}_1, \mathrm{Y}_2, \dots \mathrm{Y}_p$.

3) ω est symétrique relativement à ses deux p-uples d'arguments : $\omega(\mathrm{Y}_1, \mathrm{Y}_2, \dots \mathrm{Y}_p; \mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_p) = \omega(\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_p; \mathrm{Y}_1, \mathrm{Y}_2, \dots \mathrm{Y}_p)$

Sur $\mathfrak{V}$, on suppose bien entendu ω différentiable.

Exemples : Les tenseurs symétriques covariants d'ordre deux, dont le tenseur métrique sont les formes doubles symétriques de degré 1; le tenseur covariant de Riemann est une forme double symétrique de degré 2; les formes de degré zéro sont les scalaires : $\mathbb{R}$ pour E, $\mathfrak{F}$ pour $\mathfrak{V}$.

Les formes doubles symétriques d'ordre p sur E forment l'espace vectoriel : $(\wedge^p E^*) \vee (\wedge^p E^*)$, produit symétrisé de $\wedge^p E^*$ par lui-même. Celles de $\mathfrak{V}$ forment les sections différentiables Γ de $(\wedge^p T^*(\mathfrak{V}) \vee \wedge^p T^*(\mathfrak{V}))$.

Si l'on permute l'ensemble des p premiers arguments et l'ensemble des p derniers d'une $2p$-forme extérieure ω, elle est multipliée par $(-1)^{p^2}$. Si p est *pair*, on peut donc considérer une $2p$-forme extérieure comme une forme double symétrique de degré p et on a ainsi une inclusion naturelle $\wedge^{2p} E^* \subset (\wedge^p E^*) \vee (\wedge^p E^*)$. De la même façon on a pour p pair une inclusion :

$$\Gamma(\wedge^{2p} T^*(\mathfrak{V})) \subsetneq \Gamma(\wedge^p T^*(\mathfrak{V}) \vee \wedge^p T^*(\mathfrak{V}))$$

La somme directe de $p = 0$ à n des formes doubles symétriques est appelée l'*algèbre de courbure* $\mathfrak{C}$. Elle possède une riche structure algébrique :

1) une multiplication (cf. [E] p. 48) simplement composée des multiplications extérieures partielles sur chaque groupe d'arguments. Ainsi, le produit de ω de degré p par Π, de degré q s'écrit (cf. §IV.8) :

$$\begin{aligned}(\omega \cdot \Pi)(X_1, \dots X_{p+q}; Y_1, \dots Y_{p+q}) \\ = \sum \varepsilon(\sigma)\varepsilon(\sigma')\omega(X_{\sigma 1}, \dots X_{\sigma p}; Y_{\sigma' 1}, \dots Y_{\sigma' p}) \\ \Pi(X_{\sigma(p+1)} \dots X_{\sigma(p+q)}; Y_{\sigma'(p+1)} \dots Y_{\sigma'(p+q)})\end{aligned}$$

où σ et σ' parcourent indépendamment l'ensemble $\mathfrak{S}_{p,q}$ des permutations de $\mathfrak{S}_{p+q}$ qui conservent l'ordre des p premiers indices : $\sigma 1 < \sigma 2 < \dots \sigma p$; $\sigma' 1 < \sigma' 2 < \dots \sigma' p$ et l'ordre des q derniers indices : $\sigma(p+1) < \dots \sigma(p+q)$; $\sigma'(p+1) < \dots \sigma'(p+q)$. Cette multiplication est commutative.

Exemple : si g et h sont deux tenseurs symétriques covariants (éléments de $\mathfrak{C}_1$), leur produit est la forme $(g \cdot h)$ de degré 2 :

$$\begin{aligned}(g \cdot h)(X_1, X_2; Y_1, Y_2) = g(X_1 Y_1) h(X_2 Y_2) - g(X_2 Y_1) h(X_1 Y_2) \\ - g(X_1 Y_2) h(X_2 Y_1) + g(X_2 Y_2) h(X_1 Y_1)\end{aligned}$$

qui possède la propriété de symétrie du tenseur covariant de Riemann. D'ailleurs, pour une variété riemannienne dont la courbure sectionnelle est constante et égale à K, le tenseur n de Rieman est $\frac{1}{2}\,\mathrm{K}(g \cdot g)$.

2) des opérateurs de contraction ou traces $t : \mathfrak{C}_p \to \mathfrak{C}_{p-1}$ pour chaque p, qui généralisent la contraction de Ricci ($p = 2$) : soient $\omega \in \mathfrak{C}_p$ et (e_i) une base de E (une base locale s'il s'agit de $\mathfrak{V}$) :

$$(t\omega)(\mathrm{X}_1, \ldots \mathrm{X}_{p-1}; \mathrm{Y}_1, \ldots \mathrm{Y}_{p-1}) = \sum g^{ij}\omega(e_i; \mathrm{X}_1, \ldots \mathrm{X}_{p-1}; e_j, \mathrm{Y}_1, \ldots \mathrm{Y}_{j-1})$$

3) un endomorphisme β de degré zéro, qui applique chaque $\mathfrak{C}_p$ en lui-même, consistant à antisymétriser complètement une forme double en tous ses arguments :

$$(\beta\omega)(\mathrm{X}_1, \ldots \mathrm{X}_p; \mathrm{X}_{p+1}, \ldots \mathrm{X}_{2p}) = \frac{1}{\mathrm{P}} \sum \varepsilon(\sigma)\omega(\mathrm{X}_{\sigma 1}, \ldots \mathrm{X}_{\sigma p}, \mathrm{X}_{\sigma(p+1)}, \ldots \mathrm{X}_{\sigma(2p)}$$

où σ parcourent les $\binom{2p}{p} = \mathrm{P}$ éléments de l'ensemble $\mathfrak{S}_{p,p}$ des permutations qui conservent l'ordre des p premiers et des p derniers indices. Si p est impair, $\beta = 0$, si p est pair β est un projecteur (idempotent) sur le sous-espace des $2p$ formes extérieures, sous-espace contenu dans le noyau de t.

4) la multiplication par $g \in \mathfrak{C}_1$ qui applique linéairement $\mathfrak{C}_p$ dans $\mathfrak{C}_{p+1} : \omega \to g \cdot \omega$.

Un calcul direct immédiat montre que :

$$t(g \cdot \omega) = g \cdot t\omega + (n - 2p)\omega.$$

Le calcul par récurrence de $t^k(g \cdot \omega)$ pour $k = 1, \ldots p+1$ permet alors de montrer que l'application $\omega \to g \cdot \omega$ de $\mathfrak{C}_p$ dans $\mathfrak{C}_{p+1}$ est injective si $p < \frac{n}{2}$.

Le groupe orthogonal G, groupe d'invariance de la métrique g de E opère sur chaque espace $(\wedge^p \mathrm{E}^*) \vee (\wedge^p \mathrm{E}^*)$ par une représentation r_p, produit symétrique des représentations de G dans chaque $\wedge^p \mathrm{E}^*$. Si $a \in \mathrm{G}$, $r_p(a)\omega(\mathrm{X}_1, \ldots; \ldots \mathrm{Y}_p) = \omega(a\mathrm{X}_1, \ldots; \ldots a\mathrm{Y}_p)$. Toutes les opérations : multiplications, contractions, β, multiplication par g- *commutent* avec les opérations du groupe G autrement dit, sont des G-morphismes. Il en résulte que la multiplication par g applique (injectivement si $p < \frac{n}{2}$) $\mathfrak{C}_p$ sur un sous-espace $g\cdot\mathfrak{C}_p$ de $\mathfrak{C}_{p+1}$, stable par G, et que le noyau de β ainsi que le noyau de la contraction t de $\mathfrak{C}_{p+1}$ dans $\mathfrak{C}_p$ sont, eux aussi, des sous-espaces stables par G.

Si $\omega \in \mathfrak{C}_{p+1}$, avec $n > 2$, on peut le « purifier » de ses traces successives, en définissant l'image de Weyl de ω, « reste » de ω après divisions successives par g :

$$\mathrm{W}(\omega) = \omega - \frac{g \cdot t\omega}{n-2p} + \frac{g^2 \cdot t^2\omega}{2(n-2p)(n-2p+1)} - \dots$$
$$+(-1)^{k+1} \frac{g^{k+1} \cdot t^{t+1}\omega}{(k+1)! \prod_{j=0}^{k} (n-2p+j)}.$$

W possède les propriétés suivantes :

1) si la trace de ω est nulle, $t\omega = 0$, $\mathrm{W}(\omega) = \omega$.

2) $t\cdot W(\omega) = 0$: W est donc un projecteur de chaque $\mathfrak{C}_p$ $(p < \frac{n}{2})$ sur le sous-espace des éléments de trace nulle.

3) $\mathrm{W}(\omega) = 0$ si et seulement si ω appartient à l'image de g : $\omega = g \cdot \pi$.

4) $\mathrm{W}(\omega)$ est, relativement à la métrique g, un « invariant conforme ». En effet, si $g' = fg$ où f est un scalaire dans le cas de E, une fonction numérique différentiable dans le cas de $\mathfrak{V}$, on a pour tout k : $g'^k t'^k \omega = g^k \cdot t^k \omega$.

5) Si $\omega = \mathrm{R}$ est le tenseur covariant de Riemann, W(R) est le *tenseur de Weyl* qui s'écrit en fonction de son contracté $\mathrm{S} = t\mathrm{R}$, tenseur de Ricci, et de $s = t^2\mathrm{R}$, courbure scalaire, pour $n \geqslant 3$:

$$\begin{aligned}\mathrm{W}(\mathrm{R}) &= \mathrm{R} - \frac{g \cdot t\mathrm{R}}{n-2} + \frac{g^2 t^2 \mathrm{R}}{2(n-2)(n-1)} \\ &= \mathrm{R} - \frac{1}{n-2}\, g \cdot \left(\mathrm{S} - \frac{s}{n}\, g\right) - \frac{s}{2n(n-1)}\, g \cdot g.\end{aligned}$$

Il résulte sans difficulté de ce qui précède que l'espace des formes doubles symétriques de degré deux $\mathfrak{C}_2(\mathrm{E}) = (\wedge^2\mathrm{E}^*) \vee$

$(\wedge^2 E^*)$ est décomposé de façon unique par l'action du groupe orthogonal G en la somme directe de quatre sous-espace stables par G, irréductibles (cf. §V.4) :

1) le sous-espace $\wedge^4 E^*$, image de β, qui *intervient pas* dans la courbure en raison de la seconde identité de Bianchi qui exprime précisément que le tenseur covariant de Riemann appartient au noyau de β.

2) le sous-espace de dimension un : $\mathbb{R}g \cdot g$.

3) le sous-espace, image bijective par g si $n > 4$, du sous-espace irréductible des éléments de trace nulle de $E^* \vee E^*$, formé des éléments $g \cdot \left(S - \dfrac{g \cdot tS}{n}\right)$.

4) le sous-espace des tenseurs de courbure de Weyl, intersection $\operatorname{Im} W \cap \operatorname{Ker} \beta$.

Un tenseur de courbure covariant R se décompose donc canoniquement en la somme de trois composantes :

$$R = s(g \cdot g) + g \cdot \left(S - \frac{g \cdot tS}{n}\right) + W(R)$$

Le champ des tenseurs de courbure covariants d'une variété pseudoriemannienne $(\mathfrak{V}, g)$ se décompose canoniquement par cette même formule en la somme de trois champs de tenseurs.

Remarque : le sous-espace W de $(\wedge^2 E^*) \vee (\wedge^2 E^*)$ est le sous-espace irréductible de « poids maximal » et n'est autre que le « produit de Cartan » $(\wedge^2 E^*) \square (\wedge^2 E^*)$ des représentations irréductibles de G dans $\wedge^2 E^*$.

IX. 9 – Géodésiques, coordonnées normales, déviation

La simple donnée d'une dérivation covariante sur une variété différentielle $\mathfrak{V}$ (définition IX.5.A) que nous supposerons $\mathfrak{C}^\infty$, permet d'y définir les géodésiques comme les chemins différentiables : $t \in [a, b] \subset \mathbb{R} \to c(t) \in \mathfrak{V}$, d'accélération nulle $\dfrac{D}{dt}\left(\dfrac{dc}{dt}\right) = 0$ (définition IX.4.B). Le paramétrage d'une géodésique est déterminé à une transformation affine près de la droite. En effet, si θ est un difféomorphisme : $u \in [c,d] \to \theta(u) \in [a,b]$, $\dfrac{dc}{du} = \dfrac{dc}{dt}\,\theta'$ et $\dfrac{\mathrm{D}}{du}\Big(\dfrac{dc}{du}\Big) = \theta'^2 \dfrac{\mathrm{D}}{dt}\Big(\dfrac{dc}{dt}\Big) + \theta'\theta''\dfrac{dc}{dt}$ n'est alors nul que si $\theta'' = 0$, soit $\theta(u) = \alpha u + \beta$.

Dans un système de coordonnées locales $\Big(x^1, x^2, \dots x^n, \dfrac{\partial}{\partial x^1}\dfrac{\partial}{\partial x^2}, \dots \dfrac{\partial}{\partial x^n}\Big)$ de $\mathrm{T}(\mathfrak{V})$ au voisinage d'un point $\mathrm{M}(x_0^1, x_0^2, \dots x_0^n)$, les équations des géodésiques :

$$\frac{d^2x^i}{dt^2} + \sum \Gamma^i_{jk}\,\frac{dx^j}{dt}\frac{dx^k}{dt} = 0, \quad i = 1, 2, \dots n$$

donnent un système d'équations différentielles linéaires du premier ordre :

$$\frac{dx^i}{dt} = v^i; \quad \frac{dv^i}{dt} = \sum \Gamma^i_{jk} v^j v^k$$

dont les seconds membres ne contiennent pas la variable t.

Le théorème d'existence et d'unicité entraîne qu'il existe un voisinage W de $(x_0, 0)$ dans $\mathrm{T}(\mathfrak{V})$:

$$\operatorname{Sup}|x^i - x_0^i| < \rho; \quad \operatorname{Sup}|\mathrm{X}^j| < r$$

tel que pour tous $(x_1, \mathrm{X}_1) \in \mathrm{W}$, le système possède une solution unique $x = \gamma(t; x_1, \mathrm{X}_1)$ satisfaisant aux conditions initiales :

$$\gamma(0; x_1, \mathrm{X}_1) = x_1 \quad \text{et} \quad \frac{d\gamma}{dt}\Big(0; x_1, \mathrm{X}_1) = \mathrm{X}_1.$$

Cette solution est définie pour $|t| \leqslant a$, est $\mathfrak{C}^\infty$ en t, et dépend également de façon $\mathfrak{C}^\infty$ des conditions initiales x_1 et X_1.

On peut utiliser l'homogénéité des équations pour avoir un intervalle de définition fixe $|t| \leqslant 1$, quelles que soient les conditions initiales X en modifiant le voisinage dans lequel celles-ci peuvent être choisies, celui-ci étant ramené à $\operatorname{Sup}|\mathrm{X}^i| < r_1$.

En effet $\gamma(\lambda t; x_1, \mathrm{X}_1) = \gamma(t; x_1, \lambda\mathrm{X}_1)$ en raison de l'unicité. En prenant $\lambda = |a|$, $\gamma([a]t; x_1, \mathrm{X}_1)$ est défini $|t| \leqslant 1$ et pour $\operatorname{Sup}|a\mathrm{X}_1^i| < r$ soit $\operatorname{Sup}|\mathrm{X}_1^i| < \dfrac{r}{|a|}$. L'application qui à chaque vecteur tangent X de ce voisinage de 0 dans l'espace tangent en

M fait correspondre le point $\gamma(1; x_0, X) = \gamma x(1)$ de la géodésique est appelée l'application exponentielle et notée exp. On a : $\gamma(t; x_0, u) = \gamma(1; x_0, tu) = \exp(tu)$, d'où :

$$\frac{d\gamma^i(t,x,u)}{dt} = \sum_j \frac{\partial\gamma^i(1,x;tu)}{\partial X^j}\, u^j$$

soit, en posant $tu = v$ et en multipliant par t :

$$t\, \frac{d\gamma^i(t,x,u)}{dt} = \sum_j \frac{\partial(\exp v)^i}{\partial v^j}\, v_j.$$

Puisque, lorsque t tend vers zéro, $\dfrac{d\gamma^i(t,x,u)}{dt}$ tend vers u^i :

$$t(u^i + \varepsilon(t)) = v^i + t\varepsilon(t) = \sum_j \frac{\partial(\exp v)^i}{\partial v^j}\, v^j$$

il en résulte que :

$$\boxed{\left(\frac{\partial(\exp v)^i}{\partial v^j}\right)_{v=0} = \delta^i_j}$$

L'application exponentielle a ainsi pour application tangente à l'origine l'identité. D'après le théorème des fonctions implicites, il existe un voisinage de l'origine dans $T(\mathfrak{V})_M$, sur lequel exp est un difféomorphisme. Son image est un voisinage U de M qui *possède la propriété que chacun de ses points peut être joint à* M *par une géodésique unique.*

Le choix d'une base quelconque $(e_1, e_2, \ldots e_n)$ dans $T(\mathfrak{V})_M$ détermine par l'application inverse $\overset{-1}{\exp}$ un système de coordonnées dans le voisinage U dit système de *coordonnées normales* : si $P \in U$, $P = \exp(\zeta^1 e_1 + \zeta^2 e_2 + \ldots \zeta^n e_n)$ a pour coordonnées normales $\zeta^1, \zeta^2, \ldots \zeta^n$. Dans ce système de coordonnées, les géodésiques issues de M ont pour équation $\zeta^i = a^i t$ d'où il résulte que les symboles de Christoffel Γ^i_{jk} correspondant vérifient $\sum \Gamma^i_{jk}(a^1 t, \ldots a^n t) a^j a^k = 0$ tout le long de la géodésique définie par le vecteur a de $T(\mathfrak{V})_M$. Mais pour $t = 0$, on a l'égalité : $\sum (\Gamma^i_{jk}) a^j a^k = 0$ quel que soit le vecteur (a^j) d'où : $(\Gamma^i_{jk})_0 + (\Gamma^i_{kj})_0 = 0, \forall i, j, k$. Si la connexion est *symétrique*, il en résulte que tous les symboles de Christoffel relatifs aux coordonnées (ζ^j) sont nuls à l'origine M de ces coordonnées. Les *dérivées covariantes des champs de vecteurs ou de tenseurs en* M *sont alors les dérivées ordinaires par rapport aux coordonnées normales* ζ^i. Cette propriété rend très précieuses les coordonnées normales dans les démonstrations. Si la connexion est celle d'une variété pseudoriemannienne, cela entraîne que toutes les dérivées $\left(\dfrac{\partial g_{ij}}{\partial \zeta^k}\right)_0$ des composantes du tenseur métrique sont nulles en M. On peut encore particulariser dans ce cas en prenant pour base de référence e dans $\mathrm{T}(\mathfrak{V})_{\mathrm{M}}$ une base orthonormale.

Une manifestation de la courbure, faisant apparaître celle-ci comme gravitationnelle en relativité générale, est la *déviation des géodésiques*.

Considérons une application deux fois continûment différentiable c d'un rectangle : $0 \leqslant t \leqslant 1$, $-a < u < a$ tel que chaque chemin $\gamma_u : t \to \gamma_u(t) = c(t, u)$ soit une géodésique.

Le champ de vecteurs $\mathrm{J} = \dfrac{\partial c}{\partial u}\,(t, 0)$ le long de la géodésique γ_0 *mesure l'écart* infinitésimal entre γ_0 et les géodésiques voisines. $\dfrac{\mathrm{D}}{\partial t}\left(\dfrac{\partial c}{\partial u}\right)$ mesure la vitesse d'écartement le long de γ_0 et $\dfrac{\mathrm{D}^2}{\partial t^2}\left(\dfrac{\partial c}{\partial u}\right)$ l'accélération de cet écartement, indiquant la *déviation* d$'$ avec un écartement uniforme.

Puisque D est symétrique (sans torsion) et que le crochet $\left[\dfrac{\partial c}{\partial u}, \dfrac{\partial c}{\partial v}\right]$ est nul :

$$\frac{\mathrm{D}}{\partial t}\left(\frac{\partial c}{\partial u}\right) = \mathrm{D}_{\frac{\partial c}{\partial t}}\left(\frac{\partial c}{\partial u}\right) = \mathrm{D}_{\frac{\partial c}{\partial u}}\left(\frac{\partial c}{\partial t}\right) = \frac{D}{\partial u}\left(\frac{\partial c}{\partial t}\right).$$

On peut donc écrire :

$$\frac{\mathrm{D}^2}{\partial t^2}\left(\frac{\partial c}{\partial u}\right) = \frac{\mathrm{D}}{\partial t}\frac{\mathrm{D}}{\partial u}\left(\frac{\partial c}{\partial t}\right) = \frac{\mathrm{D}}{\partial u}\frac{\mathrm{D}}{\partial t}\left(\frac{\partial c}{\partial t}\right) + \mathfrak{R}\left(\frac{\partial}{\partial t'}\frac{\partial c}{\partial u}\right)\frac{\partial c}{\partial t}$$

$\dfrac{\mathrm{D}}{\partial t}\left(\dfrac{\partial c}{\partial t}\right) = 0$ puisque γ_u est une géodésique quel que soit u, et on obtient finalement l'équation différentielle vérifiée par le champ des vitesses d'écartement $\mathrm{J} = \dfrac{\partial c}{\partial u}\,(t, 0)$ le long de γ_0 qui montre que la déviation des géodésiques d'une famille à un paramètre est une manifestation de la courbure. En posant $\mathrm{T} = \dfrac{d\gamma_0}{dt}$ elle s'écrit :

$$\boxed{\frac{\mathrm{D}^2\mathrm{J}}{dt^2} = \mathfrak{R}(\mathrm{T},\mathrm{J})\mathrm{T}}.$$

Un champ de vecteurs J le long d'une géodésique γ_0 satisfaisant à cette équation différentielle est appelé un champ de Jacobi. Ils forment une famille à $2n$ paramètres comprenant une famille à 2 paramètres de solutions triviales : $\mathrm{J} = (\alpha t + \beta)\mathrm{T}$. (On peut comparer cette équation à celle qui régit l'écart de deux masses ponctuelles en mouvement soumises à un champ gravitationnel newtonien dans $\mathbb{R}^3$ ou au mouvement d'une particule chargée dans un champ électromagnétique, cf. [R], [MTW]).

IX.10 – Intégration des formes extérieures

Soit $\mathfrak{V}$ une variété différentielle de dimension n de classe $r \geqslant 2$. Nous allons montrer que la *seule* structure de variété différentielle *orientée* paracompacte (cf. théorème IX.4.A) permet de définir l'intégrale des n-formes différentielles extérieures à support compact sur $\mathfrak{V}$, et l'intégrale des p-formes sur les « p-chaînes ».

Considérons d'abord une n-forme ω ayant un *support compact contenu dans un ouvert* U *de* $\mathfrak{V}$*, où deux cartes* φ *et* ψ *de même orientations sont définies*. Si $x \in \mathrm{U}$, soient $x^1, x^2, \ldots x^n$, les coordonnées de $\varphi(x)$ et $y^1, y^2, \ldots y^n$ les coordonnées de $\psi(x)$. Des deux expressions ci-dessous de ω, où le déterminant fonctionnel $\dfrac{\mathrm{D}(x^1,\ldots x^n)}{\mathrm{D}(y^1,\ldots y^n)}$ est $\geqslant 0$:

$$\begin{aligned}\omega &= a(x^1,\ldots x^n)dx^1dx^2\ldots dx^n\\ &= a(x^1(y^1,\ldots y^n),\ldots)(\mathrm{D}(x^1,\ldots x^n)/\mathrm{D}(y^1,\ldots y^n)dy^1dy^2\ldots dy^n\end{aligned}$$

la seconde s'obtient à partir de la première par le changement de variables déterminé par le changement de cartes :

$$(y^1y^2\ldots y^n) \xrightarrow{\varphi\psi^{-1}} (x_1x^2\ldots x^n).$$

Les n-formes apparaissant dans les seconds membres ci-dessus sont des n-formes de $\mathbb{R}^n$ à support compact, et d'après la formule de changement de variables dans les intégrales multiples :

$$\int_{\mathbb{R}^n} a(x^1, \dots x^n) dx^1 \dots dx^n$$
$$= \int_{\mathbb{R}^n} a(x^1(y^1, \dots y^n), \dots)(D(x^1, \dots x^n)/D(y^1, \dots y^n)) dy^1 \dots dy^n.$$

D'après le principe énoncé au §IX.1, on peut donc définir l'intégrale sur $\mathfrak{V}$ de ω comme étant l'intégrale sur $\mathbb{R}^n$ de son image par une carte quelconque :

$$\int_{\mathfrak{V}} \omega = \int_{\mathbb{R}^n} a(x^1, \dots x^n) dx^1 \dots dx^n.$$

Supposons maintenant *fixés* un atlas *orienté* $\mathfrak{A} = \{\varphi_i; \mathrm{U}_i\}$ de $\mathfrak{V}$ dont les ouverts U_i forment un recouvrement localement fini de $\mathfrak{V}$ et une partition $\mathfrak{C}^r$ de l'unité : $1 = \sum \alpha_i$ subordonnée à ce recouvrement : support $\alpha_i \subset \mathrm{U}_i$ (cf. Th. IX.4.A). Toute n-forme à support compact ω peut s'écrire $\omega = \sum \alpha_i \omega$ et est somme d'un nombre fini de formes $\alpha_i \omega$ de même classe que ω, le support compact de $\alpha_i \omega$ étant contenu dans U_k.

Les intégrales sur $\mathfrak{V}$ des $\alpha_i \omega$ viennent d'être définies et il est naturel d'écrire : $\displaystyle\int \omega = \sum \int_{\mathfrak{V}} \alpha_i \omega$. Cette intégrale dépend linéairement de ω. Elle ne dépend pas de l'atlas et de la partition de l'unité choisie car si $\sum \beta_j = 1$ est une autre partition $\mathfrak{C}^r$ de l'unité on a, en vertu de la linéarité :

$$\int \omega = \int \sum_i \alpha_i \omega = \int \sum_{i \cdot j} \alpha_i \beta_j \omega$$
$$= \sum_j \int (\sum_i \alpha_i) \beta_j \omega = \sum_j \int \beta_j \omega.$$

Si maintenant $\mathfrak{V}$ est une variété différentielle avec un élément de volume ω, ce dernier, jouant le rôle de la mesure de Lebesgue $dx^1 dx^2 \dots dx^n$ sur $\mathbb{R}^n$, permet de définir l'intégrale des fonctions numériques continues à support compact sur $\mathfrak{V}$ par :
$\displaystyle\int_{\mathfrak{V}} f = \int f\omega.$

L'intégration des p-formes dans une variété différentielle généralise les intégrales curvilignes et de surface dans $\mathbb{R}^3$.

Une p-forme ω dans $\mathfrak{V}$ est une mesure p-dimensionnelle sur $\mathfrak{V}$, c'est-à-dire qu'à tout p-vecteur tangent $\mathrm{X}_1 \wedge \mathrm{X}_2 \ldots \wedge \mathrm{X}_p$, ω attribue une mesure : $\omega(\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_p)$. On va donc intégrer les p-formes sur des sous-espaces de $\mathfrak{V}$ « à p-dimensions ».

Une chemin différentiable par morceaux est un chemin continu $t \in \mathbb{R} \to c(t) \in \mathfrak{V}$ qui est différentiable à l'exception de points isolés $\ldots < t_i < T_{i+1} < \ldots$ où il est seulement différentiable à droite et à gauche.

Il peut être considéré comme la réunion des images différentiables des segments $[t_i, t_{i+1}]$ de $\mathbb{R}$, la différentiabilité incluant la différentiabilité à droite en t_i et à gauche en t_{i+1} pour chaque segment.

Comme analogues p-dimensionnels des segments de $\mathbb{R}$ pour $p > 1$, on peut choisir les simplexes : le simplexe type à p-dimensions est l'enveloppe convexe dans $\mathbb{R}^p$ de l'origine et des vecteurs de base : ensemble des $(\zeta^1, \zeta^2, \ldots \zeta^p)$ tels que : $\zeta^j \geqslant 0$, $\forall j$ et $\sum \zeta^j \leqslant 1$. Mais on peut aussi choisir les cubes : $\{0 \leqslant \zeta^k \leqslant 1, j = 1, 2, \ldots p\}$ = cube unité de $\mathbb{R}^p$. Les faces des simplexes sont des simplexes, les faces des cubes sont des cubes. Les simplexes et les cubes de dimension p ne sont, pas plus que les segments segments de $\mathbb{R}$, des variétés, à cause de leurs bords. Ce sont des exemples de variétés « polyédrales » dont la structure est définie par un atlas $\mathfrak{C}^r$ de cartes prenant leurs images non pas dans $\mathbb{R}^p$ mais dans l'angle polyèdre $\mathrm{P}_p = \mathrm{P}(\mathbb{R}^p)$ formé de l'ensemble des points de $\mathbb{R}^p$ à coordonnées toutes positives : $\zeta^1 \geqslant 0, \ldots \zeta^p \geqslant 0$. La différentiabilité en un point de P_p dont une ou plusieurs coordonnées $\zeta^{j_1}, \zeta^{j_2} \ldots \zeta^{j_k}$ sont nulles, fait intervenir uniquement les vecteurs X issus de ce point dont les composantes correspondantes $\mathrm{X}^{j_1}, \mathrm{X}^{j_2} \ldots \mathrm{X}^{j_k}$ sont positives c'est-à-dire qui sont « intérieure » à P_p.

On n'a donc en un point d'une variété polyédrale dont l'image par une carte dans P_p est un tel point qu'un « morceau » (angle polyèdre) d'espace tangent, (analogue aux demi-droites tangentes aux extrêmités de l'image différentiable d'un segment).

Cependant, une fonction des vecteurs d'un angle polyèdre, qui est linéaire ou multilinéaire pour les seuls vecteurs qui lui appartiennent, est évidemment la restriction à cet angle d'une fonction linéaire ou multilinéaire de l'espace vectoriel qu'il engendre.

Si $j_1, j_2, \ldots j_k$ sont k indices pris parmi $(1, 2, \ldots n)$, le sous-espace $\mathrm{F}_{j_1 j_2 \ldots j_k} = \mathrm{P}_n \cap \{\zeta^{j_1} = \zeta^{j_2} = \zeta^{j_k} = 0\}$ est appelé une k-face ou face de codimension k de P_n. Si l'on désigne par $\mathrm{E}_{j_1 j_2 \ldots j_k}$ l'espace vectoriel de dimension $(n-k)$ engendré dans $\mathbb{R}^n$ par les vecteurs de base dont l'indice i est $\neq j_1, j_2 \ldots$ ou j_k, $\mathrm{F}_{j_1 j_2 \ldots j_k} = \mathrm{P}(\mathrm{E}_{j_1 j_2 \ldots j_k})$ est l'ensemble des points de coordonnées positives de $\mathrm{E}_{j_1 j_2 \ldots j_k}$.

L'orientation canonique de $\mathbb{R}^n$ détermine une orientation sur chaque 1-face F_i de P_n de la façon suivante : une base $(u_2 u_3 \ldots u_n)$ de F_i est dite positive si, en la complétant par un vecteur unitaire normal u_1 *dirigé vers l'extérieur de* P_n on obtient une base positive de $\mathbb{R}^n$: $(u_1; u_2, \ldots u_{n-1}, u_n)$.

Puisque $e_1 \wedge e_2 \wedge \ldots e_n = (-1)^{i-1} e_i \wedge (e_1 \wedge e_2 \wedge \ldots \wedge \widehat{e}_i \wedge \ldots e_n)$ et que e_i est dirigé vers l'intérieur de P_n, $(-1)^i e_1 \wedge e_2 \ldots \wedge \widehat{e}_i \wedge \ldots e_n$ est un $(n-1)$-vecteur positif de F_i. L'élément de volume orienté positivement correspondant est :

$$\boxed{(-1)^i d\zeta^1 \wedge \ldots \wedge \widehat{d\zeta^i} \wedge \ldots \wedge d\zeta^n},$$

F_{ij} peut être considérée comme 1-face de F_i et comme 1-face de F_j. *Les orientations induites par* F_i *et par* F_j *sur* F_{ij} *sont opposées.* Supposons en effet $i < j$:

L'orientation induite par F_i : $(-1)^i (-1)^{j-1} d\zeta^1 \wedge \ldots \widehat{d\zeta^i} \wedge \ldots \wedge \widehat{d\zeta^j} \wedge \ldots \wedge d\zeta^n$ et l'orientation induite par F_j : $(-1)^j (-1)^i d\zeta^1 \wedge \ldots \widehat{d\zeta^i} \wedge \ldots \wedge d\zeta^n$ sont bien de signes contraires. (il est conseillé de faire une figure pour le cas $n = 3$)

Un polyèdre S de $\mathbb{R}^p$ est une variété polyèdrale orientée de dimension p : on peut choisir un atlas fini orienté de cartes dans P_p. Les k-faces du polyèdre S sont des polyèdres de dimension $(p - k)$.

Par un changement de variables, c'est-à-dire par un difféomorphisme d'un domaine d'intégration D (supposé borné connexe) dans $\mathbb{R}^p$ sur son image $\varphi\mathrm{D}$, également dans $\mathbb{R}^p$, l'intégrale d'une p-forme se trouve conservée, si φ conserve l'orientation : $\text{dét}(d\varphi) > 0$ ou changée en son opposée si φ change l'orientation :

$$\int_{\mathrm{D}} \omega = \operatorname{sgn}(\text{dét } d\varphi) \int_{\varphi \mathrm{D}} \overset{-1}{\varphi}{}^* \omega.$$

Un *domaine d'intégration doit donc être orienté.* Si D est un domaine d'intégration borné de $\mathbb{R}^p$ avec une certaine orientation, il est naturel de noter $(-\mathrm{D})$, relativement à l'intégrale, le même domaine avec l'orientation opposés.

Cela revient à considérer D comme une forme linéaire sur l'espace vectoriel des p-formes de classe $\mathfrak{C}^1$ de $\mathbb{R}^p$ par le moyen de l'intégrale.

Avec cette convention, si l'on désigne par $\partial\mathrm{S}$, ou *bord de* S, la somme formelle des 1-faces d'un polyèdre orienté S, munies de l'orientation induite :

$$\partial\mathrm{S} = \mathrm{F}_1 + \mathrm{F}_2 + \ldots \mathrm{F}_k.$$

On a alors $\partial^2\mathrm{S} = \partial\partial\mathrm{S} = \partial\mathrm{F}_1 + \partial\mathrm{F}_2 + \ldots \partial\mathrm{F}_k = 0$

Par exemple, le bord du segment orienté $[a, b]$ de $\mathbb{R}$ est $\partial[a, b]$ $= b - a$, et on a : $\displaystyle\int_{[a,b]} df = \int_{\partial[a,b]} f = f(b) - f(a)$.

Le domaine d'intégration élémentaire naturel pour une p-forme différentielle extérieure ω sur la variété $\mathfrak{V}$ est constitué par le couple $\mathfrak{S} = (\mathrm{S}, s)$ formé d'un polyèdre S de $\mathbb{R}^p$, *orienté*, et d'une application différentiable s de S dans $\mathfrak{V}$. On appelle $\mathfrak{S}$ un *polyèdre singulier* de $\mathfrak{V}$. On pose alors $\displaystyle\int_{\mathfrak{S}} \cdot\omega = \int_{\mathrm{S}} s^*\omega$.

Si l'on a un nombre fini de couples $\mathfrak{S}_i = (\mathrm{S}_i, s_i)$ de cette nature, leur combinaison linéaire formelle à coefficients réels r_i : $\sum r_i\mathfrak{S}_i$ est appelée une *chaîne polyèdrale* de dimension p de $\mathfrak{V}$ avec la convention que : $-\mathfrak{S} = (-\mathrm{S}, s)$ est obtenu à partir de $\mathfrak{S}$ par le changement d'orientation de S. L'intégration d'une p-forme sur une chaîne polyèdrale est immédiatement définie par : $\displaystyle\int_{\sum r_i\mathfrak{S}_i} \omega = \sum r_i \int_{\mathfrak{S}_i} \omega$, et on a bien : $\displaystyle\int_{-\mathfrak{S}} .\omega = -\int_{\mathfrak{S}} .\omega$.

Ici encore, une chaîne polyédrale de dimension p sur $\mathfrak{V}$ apparaît comme une forme linéaire sur l'espace vectoriel des p-formes de classe $\mathfrak{C}^1$ de $\mathfrak{V}$ par le moyen de l'intégrale.

Les chaînes les plus couramment utilisées sont les *chaînes simpliciales* et les *chaînes cubiques*, où tous les S_i sont des simplexes ou des cubes.

Appelons bord de $\mathfrak{S} = (\mathrm{S}, s)$ la chaîne de dimension $(p-1)$ que forme la somme des faces S_j de S munies de l'orientation induite et de la restriction de s : $\partial s = \sum(\mathrm{S}_j, s|_{\mathrm{S}_j})$.

Le bord d'une chaîne : $c = \sum r_i\mathfrak{S}_i$ s'écrit donc $\partial c = \sum r_i\partial\mathfrak{S}_i$, et on a $\partial^2 c = \partial\partial c = 0$. Le bord d'une chaîne simpliciale est une chaîne simpliciale, le bord d'une chaîne cubique est une chaîne cubique.

Si l'on désire intégrer une p-forme sur une sous-variété polyédrale orientée $\mathfrak{W}$ de dimension p de $\mathfrak{V}$, il suffit de décomposer $\mathfrak{W}$ en simplexes singuliers, cubes singuliers, ou polyèdres singuliers, et de la remplacer par une chaîne simpliciale, ou cubique, ou polyédrale pour effectuer l'intégration.

Théorème de Stokes IX.10.A : Soient ω une $(p-1)$-forme différentielle extérieure de classe $\mathfrak{C}^1$ sur la variété polyédrale de dimension p de $\mathfrak{V}$. On a :

$$\boxed{\int_{\mathrm{C}} d\omega = \int_{\partial \mathrm{C}} \omega}.$$

Preuve : Il suffit par linéarité de considérer un polyèdre singulier $\mathfrak{S} = (\mathrm{S}, s)$. On peut trouver un nombre fini d'ouverts de $\mathbb{R}^n$ qui recouvrent le compact S, et tels que leurs intersections avec S possèdent une carte dans P_n (voir ci-dessus). Si $\{\alpha_k\}$ est une partition $\mathfrak{C}^1$ de l'unité subordonné à ce recouvrement, l'égalité à démontrer s'écrit :

$$\int_{\mathrm{S}} d(\sum_k \alpha_k \mathrm{s}^*\omega) = \int_{\partial \mathrm{S}} (\sum_k \alpha_k \mathrm{s}^*\omega) \text{ puisque } \sum_k \alpha_k = 1.$$

Il suffit donc de démontrer :

$$\int_{\mathrm{S}} d(\alpha_k s^*\omega) = \int_{\partial \mathrm{S}} \alpha_k s^*\omega$$

c'est-à-dire le théorème de Stokes pour une forme sur S : $\alpha_k s^*\omega$ dont le support est contenu dans un ouvert possédant une carte dans P_p.

On peut donc se placer dans P_p sur lequel on se donne une $(p-1)$-forme *à support compact* :

$$\sum_j a_j dx^1 \wedge dx^2 \ldots \wedge \widehat{dx^j} \ldots \wedge dx^p$$

et il suffit de démontrer l'égalité pour une $(p-1)$-forme décomposable ω, où a est une fonction différentiable à support compact sur P_p :

$$\omega = a dx^1 \wedge dx^2 \wedge \cdot \widehat{dx^j} \ldots \wedge dx^p$$

$$d\omega = (-1)^{j-1} \frac{\partial a}{\partial x^j}\, dx^1 \wedge dx^2 \dots dx^p$$
$$\int_{\partial \mathrm{P}_p} \omega = \sum_i \int_{\mathrm{F}_i} \omega = \int_{\mathrm{F}_j} \omega$$

car ω est nulle sur toute face $x^i = 0$ pour $i \neq j$.

F_j est orientée par $(-1)^j dx^1 \wedge \dots \widehat{dx^j} \dots \wedge dx^p$ et l'intégrale de ω sur F_j (orientée) est donc :

$$\int_{\mathrm{F}_j} \omega = (-1)^j \int a(x^1, \dots, 0^j, \dots x^p) dx^1 dx^2 \dots \widehat{dx^j} \dots dx^p$$

(intégrale multiple ordinaire)

$$\begin{aligned} \int_{\mathrm{P}_p} d\omega &= (-1)^{j-1} \int_{\mathrm{P}_p} \frac{\partial a}{\partial x^j}\, dx^1 dx^2 \dots dx^p \\ &= (-1)^{j-1} \int [a]_{x^j=\infty, x^j=0}\, dx^1 dx^2 \dots \widehat{dx^j} \dots dx^p \\ &= -(-1)^{j-1} \int a(x^1, \dots, \mathrm{O}^j, \dots x^p) dx^1 dx^2 \dots \widehat{dx^j} \dots dx^p \end{aligned}$$

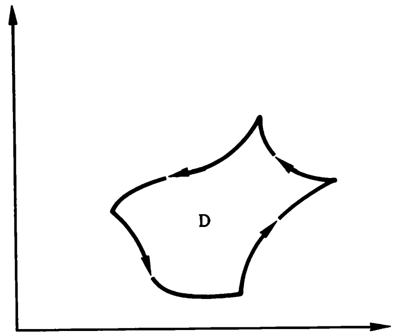
car a étant à support compact est nulle pour x^j assez grand.

On a bien : $\displaystyle\int_{\mathrm{P}_p} d\omega = \int_{\partial \mathrm{P}_p} \omega$ c.q.f.d.

Exemples : Formules de Stokes dans $\mathbb{R}^2$ et $\mathbb{R}^3$:

1) Dans le plan, si D est un domaine borné ayant pour frontière un nombre fini d'arcs différentiables, orientés par l'orientation induite de celle de $\mathbb{R}^2$, la *formule de Green* s'écrit : si $\omega = \mathrm{P}\, dx + \mathrm{Q}\, dy$
$d\omega = \left(\frac{\partial \mathrm{Q}}{\partial x} - \frac{\partial \mathrm{P}}{\partial y}\right) dx\, dy$

$$\boxed{\int_{\partial \mathrm{D}} \mathrm{P}\, dx + \mathrm{Q}\, dy = \int_{\mathrm{D}} \left(\frac{\partial \mathrm{Q}}{\partial x} - \frac{\partial \mathrm{P}}{\partial y}\right) dx\, dy}$$



2) Dans $\mathbb{R}^3$, considérons d'abord un morceau de surface différentiable S borné par un nombre fini d'arcs différentiables orientés de façon cohérente.

Si $\omega = \mathrm{P}\,dx + \mathrm{Q}\,dy + \mathrm{R}\,dz$ est une 1-forme dans $\mathbb{R}^3$:

$$\boxed{\begin{aligned}\int_{\partial \mathrm{D}} \mathrm{P}\,dx + \mathrm{Q}\,dy + \mathrm{R}\,dz = \int_{\mathrm{S}} \Big(\frac{\partial \mathrm{Q}}{\partial x} - \frac{\partial \mathrm{P}}{\partial y}\Big) dx\,dy \\ + \Big(\frac{\partial \mathrm{R}}{\partial y} - \frac{\partial \mathrm{Q}}{\partial z}\Big) dy\,dz + \Big(\frac{\partial \mathrm{P}}{\partial z} - \frac{\partial \mathrm{R}}{\partial x}\Big) dz\,dx\end{aligned}}\,.$$

Si D est un domaine borné limité par un nombre fini de morceaux de surfaces différentiables et $\omega = \mathrm{A}\,dy\,dz + \mathrm{B}\,dz\,dx + \mathrm{C}\,dx\,dy$ une 2-forme dans $\mathbb{R}^3$

$$\boxed{\int_{\partial \mathrm{D}} \mathrm{A}\,dy\,dz + \mathrm{B}\,dz\,dx + \mathrm{C}\,dx\,dy = \int_{\mathrm{D}} \Big(\frac{\partial \mathrm{A}}{\partial x} + \frac{\partial \mathrm{B}}{\partial y} + \frac{\partial \mathrm{C}}{\partial z}\Big) dx\,dy\,dz}\,,$$

(formule de Gauss ou d'Ostrogradsky).

Nous allons étendre aux variétés pseudoriemanniennes cette dernière formule et prouver une forme particulière du théorème de Stokes.

Théorème IX.10 : « Théorème de Gauss » : Soient X un champ de vecteurs différentiable sur la variété pseudoriemannienne de dimension n : $(\mathfrak{V}, g)$, ζ la 1-forme associée à X. Alors, la $(n-1)$-forme $*\zeta$ représente l'« élément de flux » du champ de vecteurs X. Si D est un domaine compact connexe limité par un nombre fini d'hypersurfaces régulières, et si $\overrightarrow{\nu}$ est le champ de vecteurs unitaires normaux orientés vers l'extérieur défini sur le bord ∂D (à l'exception des « arêtes ») le flux sortant de X à travers ∂D est égal à l'intégrale dans D de la divergence de X.

$$\boxed{\int_{\partial \mathrm{D}} *\zeta = \int_{\partial \mathrm{D}} (\mathrm{X} \mid \overrightarrow{\nu}) d\sigma = \int \operatorname{div} \mathrm{X} \cdot \omega = \int d(*\zeta)}.$$

Preuve : Prenons des coordonnées locales $x^1, x^2, \ldots x^n$ au voisinage d'un point x et soit $\overline{g} = \text{dét}|g_{ij}|$.

Si $\mathrm{X} = \sum \mathrm{X}^i \dfrac{\partial}{\partial x^i}$, $\zeta = \sum \mathrm{X}_j dx^j$, on a (fin du §IX.3) :

$$*\zeta = |\overline{g}|^{1/2} \sum (-1)^{i-1} \mathrm{X}^i dx^1 \ldots \widehat{dx^i} \ldots dx^n.$$

Si $v_1, v_2, \ldots v_{n-1}$ sont $(n-1)$ vecteurs issus de x, la valeur de cette $(n-1)$-forme pour le $(n-1)$-vecteur $v_1 \wedge v_2 \wedge \ldots v_{n-1}$ est le produit par $|\overline{g}|^{1/2}$ du développement par rapport à la première colonne du déterminant :

$$\det \begin{vmatrix} \mathrm{X}^1 v_1^1 \ldots v_{n-1}^1 \\ \mathrm{X}^2 v_1^2 \ldots v_{n-1}^2 \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ \mathrm{X}^n v_1^n \ldots v_{n-1}^n \end{vmatrix}.$$

C'est donc (§VII.1 et VIII.1) le volume algébrique des n vecteurs : $\mathrm{vol}(\mathrm{X}, v_1, v_2, v_{n-1})$. Or, le $(n-1)$-vecteur $v_1 \wedge v_2 \wedge \ldots v_{n-1}$ étant régulier il possède une droite normale régulière N sur laquelle on peut choisir un vecteur unitaire $\overrightarrow{\nu}$ tel que $(\overrightarrow{\nu}, v_1, v_2, \ldots v_{n-1})$ soit orienté positivement. Décomposons X en la somme de sa composante normale sur N : $\mathrm{X_N} = (\mathrm{X} \mid \overrightarrow{\nu})\overrightarrow{\nu}$ et de sa composante dans l'hyperplan $(v_1, v_2, \ldots v_{n-1})$. Le déterminant ne change pas si l'on remplace X par $\mathrm{X_N}$, et est égal à $\mathrm{vol}(\mathrm{X_N}, v_1, v_2, \ldots v_{n-1})$ $= (\mathrm{X} \mid \overrightarrow{\nu})\mathrm{vol}(\overrightarrow{\nu}, v_1, v_2, \ldots v_{n-1})$.

Convenons de désigner par $d\sigma$ la forme élément de volume orienté de l'hypersurface S définie par (cf. §VIII.1) $d\sigma(v_1 \wedge v_2 \wedge \ldots v_{n-1}) =$ volume algébrique $(\overrightarrow{\nu}, v_1, v_2, \ldots v_{n-1})$ (c'est l'élément de volume standard si $\mathfrak{V}$ est riemannienne).

On a alors : $*\zeta = (\mathrm{X} \mid \overrightarrow{\nu})d\sigma$.

On a déjà vu que $\operatorname{div} \mathrm{X} \cdot \omega = d(*\zeta)$ mais on peut retrouver cette égalité en différentiant (§IX.3).

$$\begin{aligned} d(*\zeta) &= \sum \frac{\partial}{\partial x^i} \, (|\overline{g}|^{1/2}\mathrm{X}^i) dx^1 dx^2 \dots dx^n \\ &= |\overline{g}|^{-1/2} \sum \frac{\partial}{\partial x^i} \, (|\overline{g}|^{1/2}\mathrm{X}^i)\omega \\ &= \operatorname{div} \mathrm{X} \cdot \omega, \ \text{c.q.f.d.} \end{aligned}$$

Exemple : l'équation de continuité d'Euler.

Si un corps de densité ρ se déforme en fonction du temps t, la vitesse d'augmentation de la masse dans le domaine D est égale à la vitesse de la masse pénétrant dans D par sa frontière ∂D soit :

$$\int_{\mathrm{D}} \frac{\partial \rho}{\partial t} \, \omega = - \int_{\partial \mathrm{D}} (\rho \overrightarrow{v} \mid \overrightarrow{\nu}) d\sigma$$

où v désigne la vitesse d'un point matériel. Ceci devant avoir lieu pour un domaine arbitraire, le théorème de Gauss entraîne : $\dfrac{\partial \rho}{\partial t} + \operatorname{div}(\rho \overrightarrow{\nu}) = 0$ (équation d'Euler)

IX.11 – Calcul des variations sur une variété. Applications

Introduction : Le calcul des variations sur une variété peut présenter des aspects très différents. En effet, tout d'abord la variété peut être : soit l'espace des paramètres définissant un système qui évolue en fonction du temps, ce dernier pouvant être ou non considéré comme un paramètre privilégié (temps absolu de la mécanique ou de la physique classiques)

– soit l'espace, ou l'espace-temps où se tiennent les phénomènes : mouvements, interactions et évolution des champs (électromagnétique, de gravitation) etc... (Le cas exceptionnel du mouvement dans l'espace d'une particule, repérée par ses coordonnées, appartient à chacun des deux aspects ci-dessus)

Les propriétés d'un système sont concentrées dans l'intégrale d'une fonction L des paramètres appelée *action* dont on étudie le comportement lorsqu'on fait « varier » certaines conditions, afin de tester si l'on est ou non dans un état d'équilibre, dit « de moindre action ».

Une première catégorie de problèmes correspond aux systèmes ayant un nombre *fini* de degrés de liberté ressortissant au calcul des variations classiques « à une variable ». L'action est une intégrale « curviligne » dont on fait varier le chemin d'intégration. Parmi les chemins possibles sont distingués ceux qui réalisent « la moindre action » ou chemin extrêmaux.

Une seconde catégorie de problèmes correspond aux systèmes ayant une infinité de degrés de liberté. De façon précise, ils dépendent d'une ou plusieurs fonctions, et non plus d'un nombre fini de paramètres numériques. L'action est cette fois une intégrale multiple, généralement une intégrale sur un domaine borné d'une variété. Du domaine, on ne pourrait faire éventuellement varier que la frontière, ce qui est insuffisant. C'est pourquoi les variations utilisées sont cette fois celles des fonctions qui servent de paramètres. Exemples : en mécanique : cordes ou plaques vibrantes ; en physique : théories des champs.

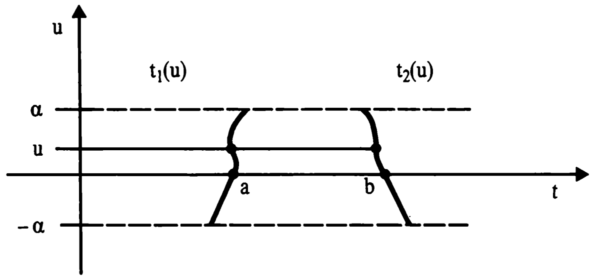
Ces deux catégories de problèmes conduisent au même type d'équation : les *équations de Lagrange* qui expriment simplement que la différentielle de l'action (prise dans un espace fonctionnel) est nulle.

Nous allons calculer la « variation » ou différentielle d'une intégrale d'action J sur la variété différentielle $\mathfrak{V}$ de dimension n, définie par une fonction numérique L de classe $\mathfrak{C}^r$ sur $\mathbb{R} \times \mathrm{T}(\mathfrak{V})$ appelée *lagrangien*. Elle s'exprime localement par une fonction de $(2n+1)$ variables : $\mathrm{L}(t; x; \mathrm{X})$.

Sur un chemin $t \in [a, b] \to x(t) \in \mathfrak{V}$, l'action J s'écrit :

$$\mathrm{J} = \int_a^b \mathrm{L}(t; x(t); x'(t))\, dt.$$

Nous allons faire varier le chemin d'intégration, y compris ses limites en fonction d'un paramètre u. Soient $t_1(u)$ et $t_2(u)$ deux fonctions $\mathfrak{C}^2$ telles que $t_1(0) = a$, $t_2(0) = b$, définies pour $-\alpha < u < \alpha$ et φ une application $\mathfrak{C}^2$ dans $\mathfrak{V}$ d'un domaine contenant les points (t, u) tels que : $-\alpha < u < \alpha$; $t_1(u) \leqslant t \leqslant t_2(u)$ et telle que $\varphi(t; 0) = x(t)$ pour $a \leqslant t \leqslant b$.



Le chemin : $x_u(t) = \varphi(t;u)$ pour $t_1(u) \leqslant t \leqslant t_2(u)$ est considéré comme une variation du chemin $x(t)$. On obtient ainsi une fonction $\mathfrak{C}^2$ de u :

$$\mathrm{F}(u) = \int_{t_1(u)}^{t_2(u)} \mathrm{L}(t;\varphi(t,u);\frac{\partial\varphi}{\partial t}\ (t,u))dt$$

dont on va calculer la dérivée en utilisant pour cela des coordonnées locales $(x^1, x^2, \dots x^n)$ dans $\mathfrak{V}$:

$$\mathrm{F}'(u) = \Big[\mathrm{L}(t;\varphi(t;u);\frac{\partial\varphi}{\partial t}\ (t,u))\ \frac{dt}{du}\Big]_{t=t_1(u)}^{t=t_2(u)} \\ + \int_{t_1(u)}^{t_2(u)} \Big(\sum \frac{\partial \mathrm{L}}{\partial x^j}\ \frac{\partial \varphi^j}{\partial u} + \sum \frac{\partial \mathrm{L}}{\partial x'^j}\ \frac{\partial^2\varphi^j}{\partial t\partial u}\Big)dt.$$

Puisque $\dfrac{d}{dt}\Big(\sum \dfrac{\partial \mathrm{L}}{\partial x'^j}\ \dfrac{\partial\varphi^j}{\partial u}\Big) = \sum \dfrac{\partial \mathrm{L}}{\partial x'^j}\ \dfrac{\partial^2\varphi^j}{\partial t\partial u} + \sum \dfrac{d}{dt}\Big(\dfrac{\partial \mathrm{L}}{\partial x'^j}\Big)\dfrac{\partial\varphi^j}{\partial u}$ une intégration par parties donne avec des notations évidentes :

$$\mathrm{F}'(u) = \Big[\mathrm{L}\ \frac{dt}{du} + \sum \frac{\partial \mathrm{L}}{\partial x'^j}\ \frac{\partial\varphi^j}{\partial u}\Big]_{t_1(u)}^{t_2(u)} \\ + \int_{t_1(u)}^{t_2(u)} \sum \Big\{\frac{\partial \mathrm{L}}{\partial x^j} - \frac{d}{dt}\ \frac{\partial \mathrm{L}}{\partial x'^j}\Big\}\ \frac{\partial\varphi^j}{\partial u}\ dt.$$

Les extrêmités du chemin x_u sont :

$$x_1(u) = x_u(t_1(u)) = \varphi(t_1(u);u) \text{ et } x_2(u) = \varphi(t_2(u);u)$$

d'où : $\dfrac{dx^1}{du} = \Big(\dfrac{\partial\varphi}{\partial t}\Big)_1 \dfrac{dt_1}{du} + \Big(\dfrac{\partial\varphi}{\partial u}\Big)_1$ et $\dfrac{dx^2}{du} = \Big(\dfrac{\partial\varphi}{\partial t}\Big)_2 \dfrac{dt^2}{du} + \Big(\dfrac{\partial\varphi}{\partial u}\Big)_2$

soit, puisque $\dfrac{\partial\varphi}{\partial t}\,(t;u)$ est le vecteur tangent au chemin x_u :

$$\frac{dx^1}{du}=x_1'\,\frac{dt_1}{du}+\Big(\frac{\partial\varphi}{\partial u}\Big)_1 \text{ et } \frac{dx_2}{du}=x_2'\,\frac{dt_2}{du}+\Big(\frac{\partial\varphi}{\partial u}\Big)_2$$

ce qui permet de transformer l'expression de $\mathrm{F}'(u)$ en y remplaçant $\Big(\dfrac{\partial\varphi^j}{\partial u}\Big)_1$ et $\Big(\dfrac{\partial\varphi^j}{\partial u}\Big)_2$ par leurs expressions tirées des égalités précédentes :

$$\begin{aligned}\mathrm{F}'(u)=&\Big\{\Big[\mathrm{L}_2-\sum\frac{\partial\mathrm{L}}{\partial x'^j}\,x_2'^j\Big]\frac{dt_2}{du}+\sum\frac{\partial\mathrm{L}_2}{\partial x'^j}\,\frac{dx_2^j}{du}\Big\}\\&-\Big\{\Big[\mathrm{L}_1-\sum\frac{\partial\mathrm{L}}{\partial x'^j}x_1'^j\Big]\frac{dt_1}{du}+\sum\frac{\partial\mathrm{L}_1}{\partial x'^j}\,\frac{dx_1^j}{du}\Big\}\\&+\sum\int_{t_1(u)}^{t_2(u)}\Big\{\frac{\partial\mathrm{L}}{\partial x^j}-\frac{d}{dt}\Big(\frac{\partial\mathrm{L}}{\partial x'^j}\Big)\Big\}\frac{\partial\varphi^j}{\partial t}\,dt\end{aligned}$$

$\Big(\dfrac{\partial\varphi}{\partial u}\Big)_{u=0}$ est un champ de vecteurs contravariants le long du chemin $x(t)$ de composantes $\Big(\dfrac{\partial\varphi^j}{\partial u}\Big)_{u=0}$ définissant au premier ordre la variation φ. Les $\mathrm{E}_j=\dfrac{d}{dt}\Big(\dfrac{\partial\mathrm{L}}{\partial x'^j}\Big)-\dfrac{\partial\mathrm{L}}{\partial x^j}$ sont les composantes d'un champ de vecteurs covariants E le long du chemin $x(t)$ et la partie intégrale de $\mathrm{F}'(0)$ s'écrit comme intégrale le long du chemin de leur crochet de dualité :

$$-\int_a^b\langle\mathrm{E},\Big(\frac{\partial\varphi}{\partial u}\Big)_{u=0}\rangle dt.$$

La variation φ pouvant être choisie arbitrairement, pour que le chemin $x(t)$ rende stationnaire l'intégrale d'action J ($\mathrm{J}'(0)=0$) il est nécessaire qu'il satisfasse aux équations différentielles de Lagrange $\mathrm{E}=0$, soit : $\dfrac{d}{dt}\Big(\dfrac{\partial\mathrm{L}}{\partial x'^j}\Big)-\dfrac{\partial\mathrm{L}}{\partial x^j}=0$ pour $j=1,2,\dots n$.

Exemples :

1) Prenons pour action *l'énergie* d'un chemin $x(t)$ d'une variété riemannienne $(\mathfrak{V}, g)$:

$$\mathrm{J}(x) = \int_a^b \mathrm{T}\, dt = \frac{1}{2}\int_a^b \|x'(t)\|^2 dt = \frac{1}{2}\int_a^b (x'(t) \mid x'(t))dt$$
$$= \frac{1}{2}\int_a^b \sum g_{kl} x'^k x'^l dt.$$

Faisons varier le chemin en fonction d'un paramètre u en gardant ses extrêmités fixes.

$$\mathrm{J}'(0) = \int_a^b \Big(\frac{\mathrm{D}}{du}\,\frac{\mathrm{D}x}{dt} \;\Big|\; \frac{\mathrm{D}x}{dt}\Big)_{u=0} dt = -\int_a^b \Big(\frac{\mathrm{D}^2 x}{dt^2} \;\Big|\; \frac{dx}{du}\Big)_{u=0} dt$$

ce qui montre que le vecteur covariant E n'est autre que le vecteur accélération du chemin en composantes covariantes. Les extrêmales sont les géodésiques de $(\mathfrak{V}, g)$.

Un calcul direct des expressions E_j fait apparaître les symboles de Christoffel Γ_{ijk} et Γ^i_{jk}. Le moment $p_i = \dfrac{\partial \mathrm{T}}{\partial x'^i}$ est la i-ème composante covariante du vecteur vitesse : $p_i = x'_i$.

2) Notons $l(t)$ la longueur $\|x'(t)\|$ du vecteur-vitesse du chemin $x(t)$. On peut comparer les équations de Lagrange de l'intégrale d'énergie : $\dfrac{1}{2}\displaystyle\int_a^b l^2 dt$ à celles de l'intégrale de longueur $\displaystyle\int_q^b l dt$. Soit s une abscisse curviligne sur le chemin, d'où : $\dfrac{ds}{dt} = l$. Puisque $l^2 = \sum g_{ij} x'^i x'^j$, on a $l\,\dfrac{\partial l}{\partial x'^i} = (x')_i$ soit : $p_i = \dfrac{\partial l}{\partial x'^i} = \dfrac{1}{l}\Big(\dfrac{dx}{dt}\Big)_i = \Big(\dfrac{dx}{ds}\Big)_i$, ce que l'on aurait pu aussi déduire de l'identité d'Euler : $\sum p_i x'^i = l$. Or

$$\Big(\frac{\mathrm{D}^2 x}{dt^2}\Big) = \frac{\mathrm{D}}{dt}\Big(l\,\frac{dx}{ds}\Big)_i = \frac{dl}{dt}\Big(\frac{dx}{ds}\Big)_i + l^2\Big(\frac{\mathrm{D}^2 x}{ds^2}\Big)_i$$
$$= \frac{1}{2}\Big\{\frac{d}{dt}\Big(\frac{\partial l^2}{\partial x'^i}\Big) - \frac{\partial l^2}{\partial x^i}\Big\}$$
$$= \frac{dl}{dt}\,\frac{\partial l}{\partial x'^i} + l\Big\{\frac{d}{dt}\Big(\frac{\partial l}{\partial x'^i}\Big) - \frac{\partial l}{\partial x^i}\Big\}$$

d'où il résulte que les premiers membres des équations de Lagrange de l'intégrale de longueur sont les composantes du vecteur covariant :

$$\frac{d}{dt}\Big(\frac{\partial l}{\partial x'^i}\Big) - \frac{\partial l}{\partial x^i} = l\Big(\frac{\mathrm{D}^2 x}{ds^2}\Big)_i.$$

Les extrêmales sont évidemment les géodésiques de la variété riemannienne.

3) le lagrangien le plus simple est linéaire homogène en les x'^j, un paramètre privilégié t pouvant être inclus parmi les x^j. J est alors l'intégrale d'une forme différentielle ω de degré un, que nous supposerons de classe $\mathfrak{C}^2$: $\mathrm{J} = \displaystyle\int \omega = \int \sum a_j(x)dx^j = \int (\sum a_j(x)x'^j)du$ et les équations de Lagrange s'écrivent : $\mathrm{E}_j = \dfrac{d}{dt}\Big(\dfrac{\partial \mathrm{L}}{\partial x'^j}\Big) - \dfrac{\partial \mathrm{L}}{\partial x^j} = \dfrac{da_j}{du} - \sum \dfrac{\partial a_k}{\partial x^j}\, x'^k$ soit : $\mathrm{E}_j = \sum \Big(\dfrac{\partial a_j}{\partial x^k} - \dfrac{\partial a_k}{\partial x^j}\Big)\, x'^k$ pour $j = 1, 2, \dots n$.

Le vecteur covariant E est le contracté du tenseur antisymétrique deux fois covariant de matrice $\left|\dfrac{\partial a_j}{\partial x^k} - \dfrac{\partial a_k}{\partial x^j}\right|$ par le vecteur vitesse x'. Le vecteur E est identiquement nul si $d\omega = 0$. Cette condition est localement équivalente au fait que ω est la différentielle exacte d'une fonction numérique : $\omega = df$. *Localement*, la valeur de l'intégrale d'action $\mathrm{J} = \displaystyle\int_{\mathrm{A}}^{\mathrm{B}} \omega = \int_{\mathrm{A}}^{\mathrm{B}} df = f(\mathrm{B}) - f(\mathrm{A})$ ne dépend pas du chemin suivi : tous les chemins entre A et B rendent l'intégrale J stationnaire.

Les équations de Lagrange associées à un Lagrangien quelconque L étant linéaires en L restent donc inchangées si on ajoute à L la différentielle df d'une fonction numérique.

Le lagrangien $\mathrm{L}(t; x^1, x^2, \dots x^n; x'^1, x'^2, \dots x'^n)$ est une fonction numérique définie sur l'espace tangent $\mathrm{T}(\mathfrak{V})$ dépendant du paramètre t. La différentielle partielle de sa restriction à la fibre T_x est une forme linéaire sur T_x donc un élément de T^*_x. C'est le covecteur *moment* p dont les composantes dans le système de coordonnées choisi sont les $p_j = \dfrac{\partial \mathrm{T}}{\partial x'^j}$. On obtient ainsi une application de $\mathrm{T}(\mathfrak{V})$ dans $\mathrm{T}^*(\mathfrak{V})$ qui à chaque élément $x' \in \mathrm{T}(\mathfrak{V})$ fait correspondre le moment $p(x')$, différentielle de L sur la fibre au point x'.

Les équations de Lagrange associés à l'intégrale $\int \mathrm{L}(t; x; x')dt$ forment un système de n équations différentielles linéaires par rapport aux dérivées secondes x''^j. Si le déterminant des coefficients de ces dernières : dét$\left|\dfrac{\partial^2 \mathrm{L}}{\partial x'^i \partial x'^j}\right|$ est différent de zéro au voisinage d'un ensemble de valeurs $(\tau, \zeta^i, \zeta'^j)$ on peut alors résoudre par rapport aux x''^j et affirmer l'existence locale d'une solution unique satisfaisant à ces données initiales. Cette condition est aussi celle qui permet de résoudre par rapport aux x'^j le système d'équations : $p_i = \dfrac{\partial \mathrm{L}}{\partial x'^i}$ et d'exprimer les dérivées x'^j en fonction des moments p_i. Nous allons montrer qu'à la variable t correspond également un moment, noté p_0.

En effet, le problème de calcul des variations posé par l'intégrale d'action $\int \mathrm{L}(t; x; x')dt$ peut toujours être mis sous forme homogène en considérant ce qui était initialement le paramètre t comme une variable supplémentaire (idée de Lagrange 1788). Les x^j et t sont maintenant considérés comme fonction d'un paramètre u. Cela revient à effectuer un changement de variable dans l'intégrale.

$$\begin{aligned}\int \mathrm{L}\Big(t; x; \frac{dx}{dt}\Big)dt &= \int \mathrm{L}\left(t(u); x(t(u)); \frac{\dfrac{dx}{du}}{\dfrac{dt}{du}}\right)\frac{dt}{du}\,du \\ &= \int \mathrm{L}(t; x; \overset{-1}{t}{}'x')t'\,du = \int \overline{\mathrm{L}}du.\end{aligned}$$

Calculons le moment de L relatif à la nouvelle variable t que l'on peut aussi noter x_0 : $p_0 = \dfrac{\partial \overline{\mathrm{L}}}{\partial t'} = \mathrm{L} - \sum \dfrac{\partial \mathrm{L}}{\partial x'^j}\,\dfrac{dx^j}{dt} = \mathrm{L} - \sum p_j \dfrac{dx^j}{dt}$.

Dans les problèmes de mécanique, la variable t est le temps et l'opposé $(-p_0)$ du moment par rapport à t est *l'énergie* du système. C'est la raison pour laquelle le vecteur covariant $(p_0, p_1, \dots p_n)$ sur $\mathbb{R} \times \mathfrak{V}$ est appelé vecteur moment-énergie.

Les dérivées x'^j étant exprimées en fonction des moments p_j en résolvant le système d'équations $p_j = \dfrac{\partial \mathrm{L}}{\partial x'^j}$ (sous la condition de régularité : dét$\left|\dfrac{\partial^2 \mathrm{L}}{\partial x'^i \partial x'^j}\right| \neq 0$), on appelle *hamiltonien* H l'opposé $(-p_0)$ du moment par rapport à t soit : $\mathrm{H} = \sum p_j x'^j - \mathrm{L}$, *considéré comme fonction des* $(2n+1)$ *variables* : $t; x^j; p_j$.

Sa différentielle après simplification s'écrit :

$$d\mathrm{H} = -\frac{\partial \mathrm{L}}{\partial t}\,dt - \sum \frac{\partial \mathrm{L}}{\partial x^j}\,dx^j + \sum x'^j dp_j$$

$$\text{Soit :}\qquad d\mathrm{H} = -\frac{\partial \mathrm{L}}{\partial t} - \sum \frac{dp_j}{dt}\,dx^j + \sum \frac{dx^j}{dt}\,dp_j.$$

(en supposant vérifiées les équations de Lagrange)

On peut y lire les dérivées partielles de H par rapport aux $(2n+1)$ variables indépendantes $t; x^j, p_j$ et en déduire un système de $2n$ équations différentielles du premier ordre équivalent aux n équations de Lagrange du second ordre. Ce système a une forme symétrique simple dite canonique :

$$\frac{dx^j}{dt} = \frac{\partial \mathrm{H}}{\partial p_j} \qquad \frac{dp_j}{dt} = -\frac{\partial \mathrm{H}}{\partial x^j}.$$

Ce sont les équations d'Hamilton, découvertes par Lagrange (1808). Les seconds membres sont les dérivées partielles d'une seule et même fonction H cette fois *définie sur l'espace cotangent* $\mathrm{T}^*(\mathfrak{V})$ qui, en la circonstance, prend le nom d'*espace des phases*. Les équations d'Hamilton peuvent d'ailleurs être considérées comme les équations de Lagrange d'une intégrale d'action portant sur une fonction M de $2n$ variables : $x^1, x^2, \ldots x^n; p^1, p^2, \ldots p^n$ indépendantes (au lieu des n initiales $x^1, x^2, \ldots x^n$) :

$$\begin{aligned}\int \mathrm{M}\,dt &= \int (\textstyle\sum p_j x'^j - \mathrm{H}(t, x^1, \ldots x^n; p^1, \ldots p^n))dt \\ &= \int (\textstyle\sum p_j dx^j - \mathrm{H}\,dt).\end{aligned}$$

En effet les équations de Lagrange associées à cette intégrale s'écrivent $\dfrac{d}{dt}\Big(\dfrac{\partial \mathrm{M}}{\partial x'^j}\Big) - \dfrac{\partial \mathrm{M}}{\partial x^j} = \dfrac{dp_i}{dt} + \dfrac{\partial \mathrm{H}}{\partial x^j} = 0$

et $\dfrac{d}{dt}\Big(\dfrac{\partial \mathrm{M}}{\partial p'_j}\Big) - \dfrac{\partial \mathrm{M}}{\partial p_j} = 0 = -\dfrac{dx^j}{dt} + \dfrac{\partial \mathrm{H}}{\partial p_j} = 0.$

Le nouveau lagrangien M s'interprète en mécanique comme la différence entre l'énergie cinétique $\sum p_j x'^j$, (d'origine purement cinématique) et l'énergie potentielle H.

Une propriété remarquable des équations de Lagrange est qu'elles ont la même forme dans tout système de coordonnées. Leur covariance apparaît dans l'expression de la variation de l'intégrale d'action mais il n'est pas inutile d'en faire la vérification directe :

Théorème IX.11 : Par un changement de coordonnées locales : $x^1 = x^1(y^1, y^2, \dots y^n) = x^n(y^1, y^2, \dots y^n)$ les premiers membres des équations de Lagrange associées à une intégrale $\displaystyle\int \mathrm{L}(t; x^1, \dots x^n; x'^1, x'^2, \dots x'^n)dt$ se transforment comme les composantes d'un vecteur covariant :

si $\mathrm{E}_j(x) = \dfrac{d}{dt}\Big(\dfrac{\partial \mathrm{L}}{\partial x'^j}\Big) - \dfrac{\partial \mathrm{L}}{\partial x^j}$, on a $\mathrm{E}_k(y) = \sum \mathrm{E}_j(x)\Big(\dfrac{\partial x^j}{\partial y^k}\Big)$ soit : $\sum \mathrm{E}_j(x)dx^j = \sum \mathrm{E}_k(y)dy^k$.

Preuve : Puisque $\dfrac{dx^j}{dt} = \sum \dfrac{\partial x^j}{\partial y^k}\,\dfrac{dy^k}{dt}$, on a en remplaçant dans L les x'^j par leurs expressions :

$$\frac{\partial \mathrm{L}}{\partial y'^k} = \sum \frac{\partial \mathrm{L}}{\partial x'^j}\,\frac{\partial x^j}{\partial y^k} \text{ tandis que } \frac{\partial \mathrm{L}}{\partial y^k} = \sum \frac{\partial \mathrm{L}}{\partial x^j}\,\frac{\partial x^j}{\partial y^k} + \sum \frac{\partial \mathrm{L}}{\partial x'^j}\,\frac{\partial x'^j}{\partial y^k}$$

$$\mathrm{E}_k(y) = \sum\Big(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial x'^j} - \frac{\partial \mathrm{L}}{\partial x^j}\Big)\frac{\partial x^j}{\partial y^k} + \left(\sum \frac{\partial \mathrm{L}}{\partial x'^j}\,\frac{d}{dt}\,\frac{\partial x^j}{\partial y^k} - \sum \frac{\partial \mathrm{L}}{\partial x'^j}\,\frac{\partial x'^j}{\partial y^k}\right)$$

et les termes de la dernière parenthèse s'annulent, c.q.f.d.

Les équations d'Hamilton gardent la même forme par une famille beaucoup plus vaste de changements de variables, appelés *transformations canoniques* :

$$\begin{cases} x^j = x^j(y^1, \dots y^n;\ \mathrm{P}_1, \dots \mathrm{P}_n; t) \\ p_j = p_j(y^1, \dots y^n;\ \mathrm{P}_1, \dots \mathrm{P}_n; t) \end{cases}$$

sous la seule condition de satisfaire une égalité de la forme :

$$\sum p_j dx^j - \mathrm{H}\, dt = \sum \mathrm{P}_k dy^k - \widetilde{\mathrm{H}}\, dt + d\mathrm{S}(t; x^1 \ldots x^n; p_1, \ldots p_n)$$

qui garantit, d'après l'exemple 3 ci-dessus, l'équivalence des équations de Lagrange de $\int \sum p_j dx^j - \mathrm{H}\, dt$ et de $\int \sum \mathrm{P}_k dy^k - \widetilde{\mathrm{H}}\, dt$.

L'intérêt de la formulation hamiltonienne des équations de variation n'est pas de nature pratique, comme instrument de calcul dans la recherche des solutions des équations de Lagrange. Son intérêt est théorique. Elle permet de montrer que la structure intime des principes de moindre action s'exprime par la notion de variété symplectique (cf. ci-dessous Définition IX.11.B). Un caractère important de la formulation hamiltonienne est aussi que les paramètres et les moments associés n'apparaissent plus comme étant de nature différente, mais au contraire comme ayant exactement le même statut, ce qui permet une grande liberté dans la description des systèmes physiques. Enfin la description mathématique d'un système physique classique par une variété symplectique, éventuellement munie d'un groupe d'automorphismes, est aussi bien adaptée à la mécanique classique qu'à la mécanique quantique, rendant naturel le passage de l'une à l'autre, et permettant l'élaboration de processus de quantification.

Nous allons voir comment s'étendent aux variétés les structures symplectiques des espaces vectoriels vues au chapitre VIII.

Soient $\mathfrak{V}$ une variété différentielle de classe $\mathfrak{C}^\infty$ *de dimension paire* $2n$, $\mathfrak{F}$ l'algèbre des fonctions numériques $\mathfrak{C}^\infty$ sur $\mathfrak{V}$. Un champ de tenseurs P deux fois contravariants, antisymétriques, non-dégénérés, de classe $\mathfrak{C}^\infty$ sur $\mathfrak{V}$ définit en chaque point x une forme bilinéaire antisymétrique non-dégénérée P_x sur l'espace cotangent T^*_x. La forme bilinéaire inverse $\overset{-1}{\mathrm{P}}_x$ est une forme bilinéaire antisymétrique, non-dégénérée, sur T_x à laquelle correspond canoniquement une forme extérieure ω_x de degré deux sur T_x, de rang maximum égal à $2n$. Les $\overset{-1}{\mathrm{P}}_x$ forment un champ $\overset{-1}{\mathrm{P}}$ de tenseurs, deux fois covariants, antisymétriques, non-dégénérés, de classe $\mathfrak{C}^\infty$, les ω_x une forme différentielle extérieure ω de degré deux, de rang $2n$, sur $\mathfrak{V}$. Réciproquement $\overset{-1}{\mathrm{P}}$, ou ω, détermine P.

La donnée d'un tel champ de tenseurs P, ou $\overset{-1}{\mathrm{P}}$, ou ω, détermine :

1) En raison des identifications canoniques de $T \otimes T$ avec $\mathfrak{L}(T^*, T)$, de $T^* \otimes T^*$ avec $\mathfrak{L}(T; T^*)$, des isomorphismes entre les espaces tangents et cotangents de $\mathfrak{V}$; Cependant, ayant affaire à des tenseurs antisymétriques, on a deux possibilités d'identification différant par le signe. Il n'y a pas dans la littérature d'accord sur un « meilleur » choix, d'autant qu'on peut aussi remplacer ω par $(-\omega)$, et disposer d'un signe dans la définition du crochet de Poisson ci-dessous.

Nous choisirons (cf. chap. II, V, VIII) de faire correspondre, au champ de vecteurs X, la forme ζ image de X par l'application ρ *associée à droite* de la forme bilinéaire ω, soit $X \in \mathfrak{X} \to \zeta = \rho X \in \mathfrak{A}$ définie par $\omega(Y, X) = \langle Y, \zeta \rangle = \zeta(Y)$ soit : $\boxed{\zeta = -i(X)\omega = -X \lrcorner\, \omega}$. A une 1-forme α correspond le champ de vecteurs $A = \overset{-1}{\rho}\,\alpha$. Si γ est l'application associée à gauche de ω : $\gamma X = -\rho X = -\zeta$ et $\overset{-1}{\gamma}\,\alpha = -A$. On a :

$$\omega(X, Y) = \overset{-1}{P}(X, Y) = P(\gamma X, \rho Y) = -P(\zeta, \eta)$$

Si la 1-forme α est la différentielle df d'une fonction numérique, le champ de vecteurs G_f qui lui est associé, défini par : $df(Y) = Y \cdot f = \omega(Y, G_f)$, soit par $df = -i(G_g)\omega$ est le *gradient symplectique de f*.

2) Une multiplication antisymétrique sur $\mathfrak{F}$, appelée *crochet de Poisson* : $\{f, g\} = -P(df, dg) = \omega(G_f, G_g) = G_f \cdot g = -G_g \cdot f$.

Calculons le crochet des champs de vecteurs gradients G_f et G_g :

$$\begin{aligned}[G_f, G_g] \cdot h &= G_f(G_g \cdot h) - G_g \cdot (G_f \cdot h) = G_f \cdot \{g, h\} - G_g\{f, h\} \\ &= \{f\{g, h\}\} + \{g, \{h, f\}\} + \{h\{f, g\}\} + G_{\{f,g\}} \cdot h.\end{aligned}$$

Un calcul élémentaire montre que :

$$d\omega(G_f, G_g, G_h) = \{f\{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\}.$$

Il en résulte que la condition pour ω d'être fermée équivaut à la condition pour $\mathfrak{F}$ munie du crochet de Poisson, d'être une algèbre de Lie. Si elle est réalisée l'application $f \to G_f$ est un homomorphisme d'algèbres de Lie de $\mathfrak{F}$ dans l'algèbre de Lie des champs de vecteurs, dont le noyau est formé des constantes.

Définition IX.11.B :

On appelle *variété symplectique* $\mathfrak{V}$ une variété $\mathfrak{C}^\infty$ de dimension paire $2n$ munie d'une forme 2-forme différentielle extérieure ω *fermée et de rang maximum* $2n$ en *tout point*. Une variété symplectique possède une orientation naturelle et un élément de volume canonique $(-1)^{n-1}\dfrac{\omega^n}{n!}$ (cf. VIII.1). L'algèbre $\mathfrak{F}$ des fonctions numériques $\mathfrak{C}^\infty$ sur $\mathfrak{V}$ munie du crochet de Poisson est alors une algèbre de Lie. Elle prend le nom *d'algèbre de Poisson de* $(\mathfrak{V}; \omega)$.

Les fonctions numériques $\mathfrak{C}^\infty$ sur $(\mathfrak{V}, \omega)$ prennent le nom *d'hamiltoniens*.

Les champs de vecteurs tangents qui sont les gradients symplectiques des fonctions numériques $\mathfrak{C}^\infty$ sur $\mathfrak{V}$ sont appelés *champs hamiltoniens de* $(\mathfrak{V}, \omega)$. Ils forment une sous-algèbre de Lie de l'algèbre de Lie des champs de vecteurs de $\mathfrak{V}$.

Exemples :

1) les surfaces orientables (ω est l'élément d'aire),

2) les variétés kähériennes et pseudokählériennes : ce sont des variétés complexes $\mathfrak{V}$ dont les espaces tangents sont munis d'un produit scalaire hermitien ou pseudohermitien tel que la forme différentielle extérieures ω sur $\mathfrak{V}$ représentant la partie imaginaire de ce produit scalaire (cf. §VIII.1 Exemple 3) soit fermée $(d\omega = 0)$

3) les orbites coadjointes des groupes de Lie (theorie de Kirillov).

4) en physique, une variété symplectique $(\mathfrak{V}, a)$ représente l'espace des phases d'un système, tandis que les fonctions $\mathfrak{C}^\infty$ sur $\mathfrak{V}$ en sont les observables.

5) les espaces cotangents en vertu du lemme suivant :

Lemme XI : L'espace cotangent d'une variété différentielle de classe $\mathfrak{C}^\infty$ possède une structure naturelle de variété symplectique.

Preuve : Soit $(x^1, x^2, \dots x^n)$ un système de coordonnées dans un ouvert V de la variété $\mathfrak{V}$. Les $(dx^1, dx^2, \dots dx^n)$ forment une base des espaces cotangents T^*_x, $x \in \mathrm{V}$, et un covecteur de T^*_x s'écrit $\sum p_j dx^j$. Soient Π l'application de projection de l'espace cotangent $\mathrm{T}^*(\mathfrak{V})$ sur sa base $\mathfrak{V}$ et $q^j = x^j \circ \Pi$. Les $2n$ fonctions $(q^1, \dots q^n; p_1, \dots p_n)$ forment un système de coordonnées sur l'ouvert T^*_V de $T^*(\mathfrak{V})$. La 2-forme différentielle extérieure $\omega_V = \sum dp_j \wedge dq^j$ de rang $2n$ sur T^*_V est la différentielle extérieure de la 1-forme $\Omega_V = \sum p_j dq^j$, donc fermée. La position des indices montre qu'elle est invariante par changement de coordonnées et qu'il existe ainsi une forme ω, sur $\mathfrak{V}$ tout entière, dont elle est l'expression locale.

On peut d'ailleurs donner une définition directe globale de ω. Soit Ω la 1-forme sur $T^*(\mathfrak{V})$ définie pour un vecteur tangent Y à $T^*(\mathfrak{V})$ en $\alpha \in T^*(\mathfrak{V})$ par $\Omega(Y) = \alpha(\Pi_* Y)$. Dans le système de coordonnées précédents, $\Omega = \sum p_j dq^j$. On définit alors globalement ω par : $\omega = d\Omega$.

Le théorème de Darboux (§VIII.5) affirme qu'il existe dans tout ouvert V assez petit d'une variété symplectique quelconque, des *coordonnées normales locales* : $(q^1, q^2, q^n \ldots; p_1, p_2, \ldots p_n)$, caractérisées par le fait que dans V, ω s'écrive : $\omega = \sum dp_j \wedge dq^j$.

Cela signifie que toutes les variétés symplectiques de même dimension ont la même structure symplectique locale, entièrement exprimée par n'importe quel système de coordonnées normales. Globalement, on a vu que sur un espace cotangent, la forme symplectique est une différentielle exacte : $\omega = d\Omega$. Cela n'est pas vrai en général, et en particulier impossible lorsque la variété est compacte. En effet le volume de la variété : $\int \omega^n = \int d(\Omega \wedge \omega^{n-1})$ serait alors nul d'après le théorème de Stokes.

Rappelons que sur l'algèbre des formes différentielles extérieures, la dérivée de Lie L_X relative a un champ de vecteur X peut s'écrire $L_X = i(X)d + di(X)$. Appliquée à ω, qui est fermée, on obtient $L_X\omega = di(X)\omega$. Les difféomorphismes locaux associés au champ X laissent donc ω invariante ($L_X\omega = 0$) si et seulement si $i(X)\omega$ est fermée c'est-à-dire si et seulement si localement $i(X)\omega = df$ ce que l'on exprime en disant que X est *localement hamiltonien*, le crochet de deux champs X et Y localement hamiltoniens s'écrit $[X, Y] = G_{\omega(X,Y)}$ et est donc toujours globalement hamiltonien. Un changement de coordonnées faisant passer de coordonnées normales $(q^1, \ldots q^n, p_1, \ldots, p_n)$ à d'autres coordonnées normales $(Q^1, \ldots Q^n, P_1, \ldots P_n)$ est une transformation canonique quel que soit l'hamiltonien H, au sens défini ci-dessus. En effet, de $\omega = \sum dp_j \wedge dq^j = \sum dP_j \wedge dQ^j$, on déduit : $d(\sum p_j dq^j - \sum P_j dQ^j) = 0$ soit localement : $\sum p_j dq^j = \sum P_j dQ^j + dS$.

Nous arrivons au but de ce paragraphe avec le calcul *en coordonnées normales* :

1) du gradient symplectique G_H d'un hamiltonien H (fonction numérique $\mathfrak{C}^\infty$ sur $(\mathfrak{V}, \omega)$:

$$\begin{aligned}\omega(Y, G_H) &= (\textstyle\sum dp_j \wedge dq^j)(Y, G_H)\\ &= \textstyle\sum\langle dp_j, Y\rangle\langle dq^j, G_H\rangle - \sum\langle dp_j, G_H\rangle\langle dq^j, Y\rangle\\ &= dH(Y) = \textstyle\sum \dfrac{\partial H}{\partial p_j}\langle dp_j, Y\rangle + \sum \dfrac{\partial H}{\partial q^j}\langle dq^j, Y\rangle\end{aligned}$$

d'où : $$G_H = \sum \frac{\partial H}{\partial p_j}\frac{\partial}{\partial q^j} - \sum \frac{\partial H}{\partial q^j}\frac{\partial}{\partial p_j}.$$

Les trajectoires du champ de vecteurs hamiltonien G_H sur $\mathfrak{V}$ sont donc les solutions du système d'équations :

$$\frac{dq^j}{dt} = \frac{\partial H}{\partial p_j}; \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q^j}.$$

On a ainsi démontré l'origine symplectique des équations d'Hamilton.

2) du crochet de Poisson :

$$\begin{aligned}\{f, g\} = \omega(G_f, G_g) &= (\textstyle\sum dp_j \wedge dq^j)(G_f, G_g)\\ &= \textstyle\sum\langle dp_j, G_f\rangle\langle dq^j, G_g\rangle - \sum\langle dp_j, G_g\rangle\langle dq^j, G_f\rangle\\ &= \textstyle\sum\Big(\dfrac{\partial f}{\partial p_j}\dfrac{\partial g}{\partial q^j} - \dfrac{\partial f}{\partial q^j}\dfrac{\partial g}{\partial p_j}\Big)\end{aligned}$$

d'où : $\{p_j, p_k\} = 0$; $\{q^j, q^k\} = 0$; $\{p_j, q^k\} = \delta_j^k$; $\{H, q^j\} = \dfrac{\partial H}{\partial p_j}$; $\{H; p_j\} = -\dfrac{\partial H}{\partial q^j}$ ce qui permet d'écrire les équations des trajectoires du champ hamiltonien associé à H sous la forme :

$$\frac{dq^j}{dt} = \{H, q^j\}; \quad \frac{dp_i}{dt} = \{H, p_i\}$$

et plus généralement la dérivée d'une fonction différentiable F ne contenant pas la variable t le long de la trajectoire du champ G_H :

$$\frac{dF}{dt} = \sum \frac{\partial F}{\partial q^j} \frac{dq^j}{dt} + \sum \frac{\partial F}{\partial p_j} \frac{dp_j}{dt} = \sum \frac{\partial F}{\partial q^j} \frac{\partial H}{\partial p_j} - \sum \frac{\partial F}{\partial p_j} \frac{\partial F}{\partial q_j}$$
$$= \{H, F\}.$$

Cette écriture fait apparaître l'hamiltonien H comme un *opérateur* sur les fonctions numériques différentiables de $\mathfrak{V}$. La fonction F reste constante le long de chaque trajectoire de G_H (« constante du mouvement ») si et seulement si son crochet de Poisson avec H est nul. L'identité de Jacobi montre alors que si F_1 et F_2 sont deux « constantes du mouvement », il en est de même de leur crochet $\{F_1, F_2\}$. En particulier, si H ne contient pas la variable t, H est une « constante du mouvement » (conservation de l'énergie). Lorsque la fonction F dépend de t :

$$\frac{dF}{dt} = \{H, F\} + \frac{\partial F}{\partial t}. \text{ en particulier } \frac{dH}{dt} = \frac{\partial H}{\partial t}.$$

EXERCICES SUR LE CHAPITRE IX

EX. IX.1 :

Soient U_+ l'ouvert $\{x^{n+1} \neq 1\}$ et U_- l'ouvert $\{x^{n+1} \neq -1\}$ de la sphère unité S_n de l'espace euclidien E_{n+1}. Montrer que les projections stéréographiques des pôles nord et sud de S_n sur $\{x^{n+1} = 0\}$ définissent deux cartes formant un atlas $\mathfrak{C}_\infty$ sur S_n.

EX. IX.2 :

Soit $\mathbb{R}P_n$ l'espace projectif réel de dimension n, quotient de $\mathbb{R}^{n+1} - \{0\}$ par la relation d'équivalence $x \sim y \iff \exists\lambda \neq 0$ tel que $y = \lambda x$. Montrer que si $\mathfrak{U}_i$ est l'ouvert de $\mathbb{R}^{n+1}$ formé des vecteurs x dont la i-ème coordonnée x^i est non nulle et U_i son image dans $\mathbb{R}P_n$, les n nombres $\left(\frac{x^1}{x^i}, \dots \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \dots \frac{x^{n+1}}{x^i}\right)$ définissent des « coordonnées » sur U_i et que les $(n+1)$ U_i forment alors un atlas définissant une structure de variété $\mathfrak{C}^\infty$ sur $\mathbb{R}P^n$.

Montrer en utilisant cet atlas ou la projection de la sphère unité S_n sur $\mathbb{R}P_n$ que $\mathbb{R}P_n$ est orientable pour n impair, non orientable pour n pair.

EX. IX.3 :

Montrer qu'une variété analytique complexe, c'est-à-dire possédant un atlas dans les ouverts de $\mathbb{C}^n$ dont les changements de cartes sont analytiques complexes, est, en tant que variété différentielle réelle, toujours orientable. Exemples : les espaces projectifs complexes $\mathbb{C}P_n$.

EX. IX.4 :

Soit D la dérivation covariante d'une connexion affine symétrique sur la variété différentielle $\mathfrak{V}$ de dimension n. Soient X_1, X_2, X_n n champs de vecteurs tangents linéairement indépendants sur un ouvert U de $\mathfrak{V}$, $\Theta_1, \Theta_2, \dots, \Theta_n$ les formes différentielles de degré un sur U définies par $\Theta_j(X_i) = \delta_{ji}$. Montrer que sur U on peut exprimer la différentielle d des formes différentielles extérieures par :

$$d\omega = \sum_j \Theta_j, \wedge D_{X_j}\omega.$$

EX. IX.5 :

Montrer qu'une isométrie d'une variété riemannienne applique une géodésique sur une géodésique. En déduire que si la variété est connexe, une isométrie qui laisse fixes un point et les vecteurs tangents en ce point est l'identité.

EX. IX.6 :

Soit C une application continûment différentiable du carré unité C : $0 \leqslant s, t \leqslant 1$ de $\mathbb{R}^2$, dans la variété $\mathfrak{V}$ telle que $c(0,t) = x$ et $c(1,t) = y$ quel que soit t. c peut être considéré comme une déformation du chemin $c_0(s) = c(s,0)$ en le chemin $c_1(s) = c(s,1)$. On suppose définie sur $\mathfrak{V}$ une connexion symétrique D de courbure $\mathbb{R}$. Soient $X(s,t)$ le champ de vecteurs, $\alpha(s,t)$ le champ de covecteurs au-dessus de l'« élément de surface » $c(C)$ obtenus par transport parallèle pour chaque t le long du chemin $c_t(s) = c(s,t)$ de $X = X(0,0)$, tangent en x, et de $\alpha = \alpha(1,1)$ cotangent en y. On se propose de mesurer par α la différence entre le vecteur $X_0 = X(1,0)$ transporté de X le long de c_0 et $X_1 = X(1,1)$ le long de c_1. Montrer que l'on a :

$$\langle \alpha; X_1 - X_0 \rangle = \int\int_C \langle \alpha(s,t); R\Big(\frac{\partial c}{\partial s}, \frac{\partial c}{\partial t}\Big) X(s,t) \rangle ds\, dt.$$

En déduire que la condition nécessaire et suffisante pour que le transport parallèle ne dépende pas du chemin suivi est que la courbure $\mathbb{R}$ soit nulle (utiliser la formule de Stokes).

INDEX

Imprimé en France
Imprimerie des Presses Universitaires de France
73, avenue Ronsard, 41100 Vendôme
Avril 1993 — N° 38 646

COLLECTION *MATHÉMATIQUES*

Jean-Pierre AUBIN. *Analyse fonctionnelle appliquée. 1 et 2*

Marcel BERGER et Bernard GOSTIAUX. *Géométrie différentielle, variétés, courbes et surfaces* (2e éd. corrigée)

Josette CALAIS. *Eléments de théorie des groupes*

Bernard CHARLES et Denis ALLOUCH. *Algèbre générale*

Jean COMBES. *Suites et séries*

Paul DEHEUVELS. *L'intégrale*

René DEHEUVELS. *Formes quadratiques et groupes classiques*

Roger DESCOMBES. *Eléments de théorie des nombres*

Jacques DIXMIER. *Topologie générale*

Jean GEFFROY. *Equations différentielles*

Michel HERVÉ. *Les fonctions analytiques*

Michel HERVÉ. *Transformation de Fourier et distributions*

Daniel LEBORGNE. *Calcul différentiel et géométrie*

Daniel LEHMANN et Carlos SACRÉ. *Géométrie et topologie des surfaces*

Daniel LEHMANN et Rudolphe BKOUCHE. *Initiation à la géométrie*

Jacqueline LELONG-FERRAND. *Les fondements de la géométrie*

Henri MASCART et Marius STOKA. *Algèbre linéaire et applications. 1 et 2 : Exercices et corrigés*

Henri MASCART et Marius STOKA. *Fonctions d'une variable réelle. 1, 2, 3, 4 et 5 : Exercices et corrigés*

Maurice MIGNOTTE. *Algèbre appliquée à l'informatique*

Maurice MIGNOTTE. *Mathématiques pour le calcul formel*

Georges PUPION. *Exercices d'algèbre avec solutions et cours résumé*

Georges PUPION. *Exercices d'analyse avec solutions et cours résumé*

Pierre SAMUEL. *Géométrie projective*

Paul VER EECKE. *Fondements du calcul différentiel*

Paul VER EECKE. *Applications du calcul différentiel*

Les nombres et les vecteurs sont, à eux seuls, incapables de représenter toutes les grandeurs géométriques ou physiques et par là-même, d'exprimer certaines lois naturelles. Le calcul tensoriel, qui les généralise, est le langage dans lequel s'écrivent la géométrie différentielle, la relativité, l'électromagnétisme, la mécanique et, depuis peu, les interactions entre particules élémentaires. Ce livre détaille et justifie la structure algébrique des tenseurs, des spineurs, des connexions et de la dérivation covariante, en illustrant leur maniement par de très nombreux exemples : calcul extérieur, géométries riemannienne et symplectique, tenseurs de courbure, relativité restreinte et générale, équations de Maxwell et d'Einstein, invariants des algèbres de Lie, etc.

Suivant les principes de cette collection, on part de notions élémentaires pour construire peu à peu la théorie en l'approfondissant.

Couverture : © René Parola

395 FF

22408766 / 4 / 9