

Ecole Maths & Cerveau, I.H.P. juin 2005
5 cours sur la géométrie différentielle

Alain Chenciner
Université Paris 7 & IMCCE
chenciner@imcce.fr

5 juillet 2005

Résumé

Introduire à des techniques en si peu de temps serait une illusion ; on a cherché plutôt à mettre en évidence le caractère “naturel” de certaines notions de base de la géométrie différentielle. Seuls les 5 premiers chapitres ont été traités dans le cours oral.

1 D’un point à un autre : métriques, géodésiques, fronts d’onde

1.1 La longueur d’une courbe dans un espace euclidien

On peut définir la longueur d’une courbe dans le plan ou dans l’espace comme la limite des longueurs (euclidiennes) des polygones inscrits dans la courbe :

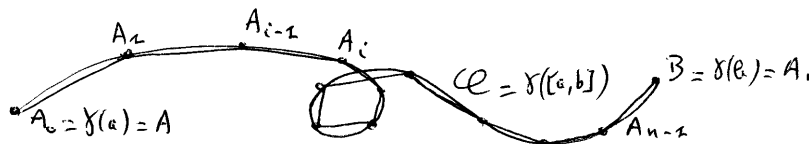


Figure 1 : $\mathcal{L}(C = \sup) \sum_{i=1}^k \|A_i - A_{i-1}\|$

Remarque. Aussi naturelle que soit cette définition, elle ne se généralise pas en une définition de l’aire d’une surface. Le contre-exemple classique est celui de la lanterne vénitienne, représentée dans la figure 6.5.4 de [BG]).

On déduit de la définition que, quel que soit le paramétrage monotone (i.e. sans retour en arrière) $\gamma : [a, b] \rightarrow \mathbb{R}^3$ de \mathcal{C} ,

$$\mathcal{L}(C) = \int_a^b \|\dot{\gamma}(t)\| dt,$$

où, comme les mécaniciens, on a noté avec un point la dérivée par rapport au paramètre t (t est pensé comme étant un temps et $\dot{\gamma}(t)$ comme étant une vitesse). On parlera indifféremment de longueur de \mathcal{C} ou de longueur de γ .

Cette formule vaut en particulier si γ est *absolument continue*, c'est-à-dire si sa dérivée $\dot{\gamma}$ au sens des distributions est une fonction (définie presque partout) intégrable et si $\gamma(t) = \gamma(a) + \int_a^t \dot{\gamma}(s) ds$.

Remarque. Attention, la longueur n'est pas une fonctionnelle continue sur l'ensemble des $\gamma : [a, b] \rightarrow \mathbb{R}$ absolument continues, muni de la topologie de la convergence uniforme : elle n'est que semi-continue inférieurement : une perturbation arbitrairement petite de la courbe peut faire augmenter brusquement sa longueur. La figure 2, classique, montre une famille de courbes de longueur $\sqrt{2}$ qui converge uniformément vers un segment de longueur 1.

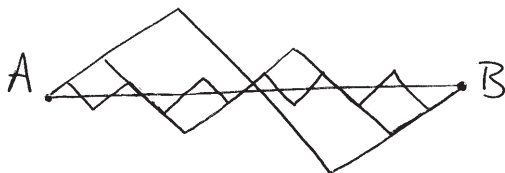


figure 2 : semi-continuité inférieure de la longueur

1.2 le plus court chemin et l'optique géométrique

Bien entendu, le plus court chemin entre deux points A et B est un segment de droite : l'inégalité triangulaire implique en effet qu'un segment de droite minimise la longueur parmi toutes les courbes polygonales (récurrence sur le nombre de morceaux), donc a fortiori parmi toutes les courbes absolument continues. Mais savez-vous le démontrer directement sur la formule intégrale ?

Première démonstration. Un déplacement ne changeant pas la longueur, on peut supposer que $A = (0, 0, 0)$ et $B = (\ell, 0, 0)$. Soit $\gamma(t) = (x(t), x^\perp(t))$ une courbe $\gamma : [a, b] \rightarrow \mathbb{R} \times \mathbb{R}^2 \equiv \mathbb{R}^3$ telle que $\gamma(a) = A$ et $\gamma(b) = B$ (figure 3). Puisque, pour tout t , $\dot{x}(t)$ et $\dot{x}^\perp(t)$ sont orthogonaux, on déduit du théorème de Pythagore que la longueur de γ vérifie

$$\mathcal{L}(\gamma) = \int_a^b \|\dot{x}(t)\|^2 dt + \int_a^b \|\dot{x}^\perp(t)\|^2 dt \geq \int_a^b \|\dot{x}(t)\|^2 dt \geq \ell.$$

L'assertion est donc démontrée par réduction à la comparaison avec le segment $[AB]$ de la projection orthogonale de la courbe γ sur la droite qui porte ce segment.

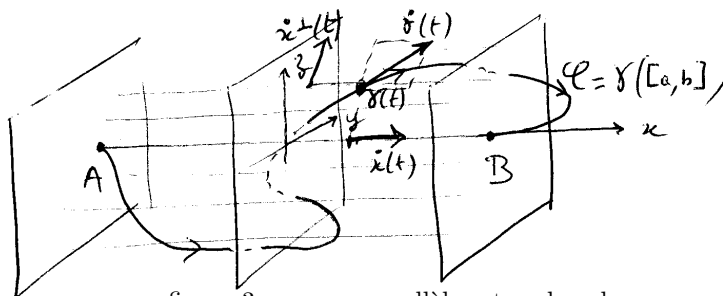


figure 3 : rayons parallèles et ondes planes

Deuxième démonstration. C'est la même en coordonnées polaires

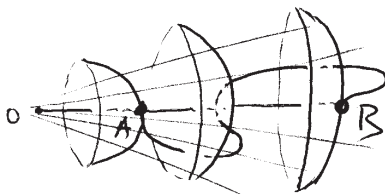


figure 4 : source ponctuelle et ondes sphériques

Troisième démonstration. On choisit cette fois des coordonnées *ondes-rayons* très générales par le procédé suivant : on part d'un morceau de surface (resp. de courbe) S orthogonal à la droite portant le segment $[AB]$ et on considère les normales à S (figure 4).

Assertion fondamentale : en portant sur ces normales des distances égales à partir de S , on définit, au moins localement, une famille de surfaces parallèles qui sont encore orthogonales à ces droites. En optique géométrique, les normales à S sont les rayons et les surfaces parallèles à S les fronts d'onde.

L'orthogonalité des ondes aux rayons permet à nouveau de calculer la longueur de γ en utilisant le théorème de Pythagore et donc de conclure.

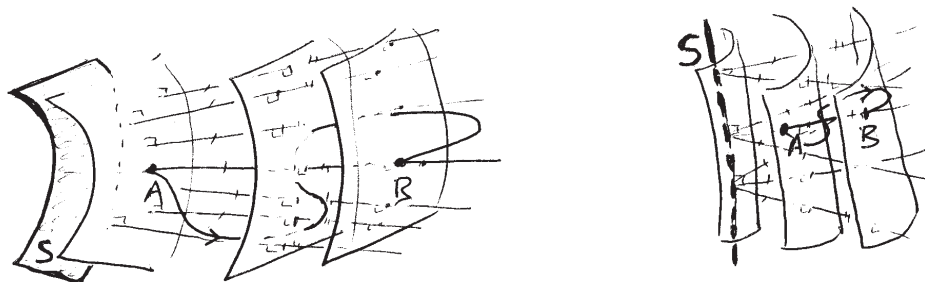


figure 5 : coordonnées "ondes-rayons" générales

Démonstration de l'assertion fondamentale ou les premiers pas dans le calcul des variations. Un inconvénient de la fonctionnelle de longueur est son insensibilité au paramétrage de C . Considérée sur l'ensemble des courbes paramétrées, elle apparaît comme dégénérée du fait même de sa nature *géométrique*. Dans beaucoup de questions, il est intéressant de la remplacer par la fonctionnelle d'énergie, de nature plus *dynamique*

$$\mathcal{E}(\gamma) = \frac{1}{2} \int_a^b \|\dot{\gamma}(t)\|^2 dt$$

dont les *extrémales* (i.e. les courbes γ telles que " $d\mathcal{E}$ "(γ) = 0) sont maintenant les droites paramétrées à vitesse constante, c'est-à-dire les courbes γ qui vérifient $\ddot{\gamma} \equiv 0$.

Considérons en effet une petite variation $\gamma + \epsilon X$ de γ , représentée par le champ de vecteurs X le long de γ s'annulant aux extrémités a et b (figure 6).

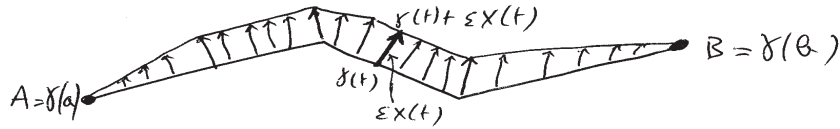


figure 6 : une petite variation de γ à extrémités fixées

La dérivée de \mathcal{E} en γ dans la direction de X est par définition

$$d\mathcal{E}(\gamma) \cdot X = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\mathcal{E}(\gamma + \epsilon X) - \mathcal{E}(\gamma)] = \int_a^b \langle \dot{\gamma}(t), \dot{X}(t) \rangle,$$

où une intégration par parties (l'une des mamelles de l'analyse, aurait dit Bobby Lapointe) transforme en la *formule de la variation première*

$$d\mathcal{E}(\gamma) \cdot X = - \int_a^b \langle \ddot{\gamma}(t), X(t) \rangle - \sum_{t_i} \langle \Delta \dot{\gamma}(t_i), X(t_i) \rangle,$$

où les t_i sont les instants de discontinuité de $\dot{\gamma}(t)$, instants qu'on a supposés en nombre fini pour simplifier. Ceci vient de ce que, si γ n'est dérivable que par morceaux, la différence $[\langle \dot{\gamma}, X \rangle]_a^b = \langle \dot{\gamma}(b), X(b) \rangle - \langle \dot{\gamma}(a), X(a) \rangle$ est la somme de l'intégrale sur $[a, b]$ de la dérivée de la fonction continue par morceaux $\langle \dot{\gamma}(t), X(t) \rangle$ et de ses sauts aux points t_i de discontinuité. Une extrémale γ devant, par définition, annuler $d\mathcal{E}(\gamma) \cdot X$ pour tout champ X le long de γ nul aux extrémités, on en déduit que γ est régulière (la vitesse n'a pas de sauts) et que $\ddot{\gamma}$ est identiquement nulle. Notons que fixer la longueur \mathcal{L} ou l'énergie \mathcal{E} d'une extrémale revient au même puisque, si $\ddot{\gamma} \equiv 0$, on a évidemment

$$\mathcal{E}(\gamma) = \frac{1}{2(b-a)} \mathcal{L}^2(\gamma).$$

L'assertion fondamentale découle alors du même calcul, à ceci près que l'on considère maintenant une famille à un paramètre u d'extrémales (i.e. de droites) γ_u et que l'on n'exige plus que les extrémités soient fixées (figure 7).

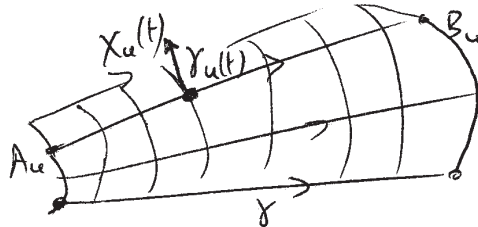


figure 7 : une variation d'extrémales

Il vient pour tout u ,

$$\frac{d}{du}\mathcal{E}(\gamma_u) = d\mathcal{E}(\gamma_u) \cdot X_u = \langle \dot{\gamma}_u(b), X_u(b) \rangle - \langle \dot{\gamma}_u(a), X_u(a) \rangle,$$

où la variation X_u est définie par $X_u(t) = \frac{d\gamma_u}{du}(t)$. Si maintenant la longueur et donc l'énergie de chaque segment $\gamma_u([a, b])$ est la même, on en déduit que

$$\langle \dot{\gamma}_u(b), X_u(b) \rangle = \langle \dot{\gamma}_u(a), X_u(a) \rangle.$$

En particulier, la fonction $\langle \dot{\gamma}_u(t), X_u(t) \rangle$ s'annule en b dès qu'elle s'annule en a , ce qui n'est autre que l'assertion fondamentale.

Remarque. L'assertion d'orthogonalité peut être lue comme une assertion – a priori étonnante – d'intégrabilité du champ des plans orthogonaux aux rayons (figure 8-1). Alors que l'intégrabilité d'un champ de droites dans \mathbb{R}^2 découle du théorème d'existence des solutions des équations différentielles ordinaires, la règle pour un champ de plans dans \mathbb{R}^3 est la non-intégrabilité, comme dans l'exemple suivant, paradigmatique.

Exemple : la structure de contact naturelle de \mathbb{R}^3 . Considérons le champ de plans dans \mathbb{R}^3 qui associe au point (x, y, z) le plan vertical $P_{(x,y,z)}$ d'équation $Y - zX = 0$, où (X, Y, Z) sont les coordonnées dans $T_{(x,y,z)}\mathbb{R}^3 \cong \mathbb{R}^3$ qui se déduisent des coordonnées de \mathbb{R}^3 par translation de l'origine en (x, y, z) (figure 8-2).

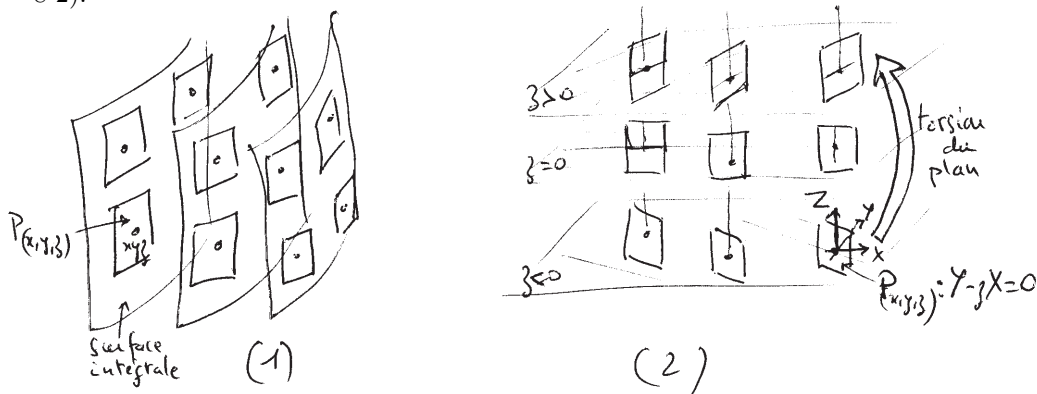


figure 8 : champs de plans intégrable (1) et non intégrable (2)

Première démonstration de la non-intégrabilité : s'il existe une surface intégrale Σ telle que $T_{(x,y,z)}\Sigma = P_{(x,y,z)}$ en chacun de ses points, l'intersection de Σ avec chaque plan horizontal $z = z_0$ doit être une courbe intégrale d'un champ de directions parallèle de pente z_0 , donc une droite, et une équation de Σ doit être de la forme $y = zx + f(z)$, ce qui interdit à ses plans tangents d'être tous verticaux.

Deuxième démonstration de la non-intégrabilité : on remarque que deux points quelconques (x, y, z) et (x', y', z') peuvent être joints par une courbe intégrale γ (ou courbe de Legendre), c'est-à-dire une courbe tangente en chacun de ses points

$\gamma(t)$ au plan $P_{\gamma(t)}$. L'existence de surfaces intégrales contredirait cette propriété car une courbe intégrale habiterait nécessairement dans une surface intégrale. On trouve dans le beau livre de Max Born *On the nature of cause and chance*, une description très claire de la caractérisation locale du deuxième principe de la thermodynamique donnée en ces termes par Caratheodory : l'existence d'une entropie S et d'une température absolue T qui vérifient $\frac{1}{T}dQ = dS$ équivaut à l'intégrabilité du champ des noyaux de la 1-forme dQ (les surfaces intégrales ont pour équation $S = \text{constante}$) et est donc conséquence de l'existence au voisinage d'un état donné d'un état arbitrairement proche que l'on ne peut atteindre adiabatiquement (i.e. avec $dQ = 0$).

Lecture. Le premier paragraphe du bel article de J. Lützen *Interactions between Mechanics and Differential Geometry in the 19th century*, *Archives for History of Exact Sciences*, vol. 49, Number 1, pp. 1 à 72, 1995, commence ainsi : “In 1828 Carl Friedrich Gauss (1777-1855) published his *Disquisitiones generales circa superficies curvas* and William Rowan Hamilton (1805-1865) published his *Theory of Systems of Rays*. There are good reasons for considering this double event the beginning of our history. Indeed Gauss's paper marked the start of the development of intrinsic differential geometry and Hamilton's paper became the basis for his subsequent works on mechanics. Moreover Gauss's and Hamilton's papers contained ideas that strike a modern reader as being highly similar. In a sense one might consider the present paper as the story of how intrinsic differential geometry and Hamiltonian mechanics after a start on common grounds first went in different directions, and only gradually came back together around 1870 on a higher (dimensional) level.”

Ce qui précède est une parfaite illustration de ce texte : l'orthogonalité des fronts d'onde aux rayons est connue sous le nom de *lemme de Gauss* en géométrie riemannienne et sous celui de *théorie de Hamilton-Jacobi* ou encore *préservation de la forme de Liouville* $\sum_i p_i dq_i$ en mécanique (voir [A]).

A suivre... par le principe de Huyghens, les caustiques (développées en géométrie, fig. 9), les singularités génériques des fronts d'onde,... et aussi par l'étude de l'optique géométrique dans un milieu homogène mais non isotrope comme premier exemple de métrique riemannienne et prélude à ce qui suit.



figure 9 : une caustique et la fonction “distance” multiforme associée

2 Des métriques induites aux structures riemanniennes

Considérons une surface paramétrée M dans \mathbb{R}^3 (ou, plus généralement une *sous-variété* paramétrée de dimension n de \mathbb{R}^N). Le paramétrage fournit des *coordonnées curvilignes* sur M . C'est une application régulière – disons indéfiniment dérivable, définie sur un domaine Ω du plan \mathbb{R}^2 ,

$$f : \Omega \rightarrow \mathbb{R}^3,$$

qui est une *immersion* (la dérivée en chaque point est de rang maximum, i.e. de rang 2) et même un *plongement* (elle est injective et propre). Rappelons que la dérivée $df(q)$ de f au point $q = (q_1, q_2) \in \Omega$ est l'application qui, au vecteur vitesse d'un mobile $t \mapsto c(t)$ passant par q à l'instant t fait correspondre le vecteur vitesse du mobile $t \mapsto f \circ c(t)$ au même instant. L'image de cette application est par définition le plan tangent $T_{f(q)}M$ à M au point $f(q)$ (figure 10). Une excellente référence pour ces notions est le livre [P]

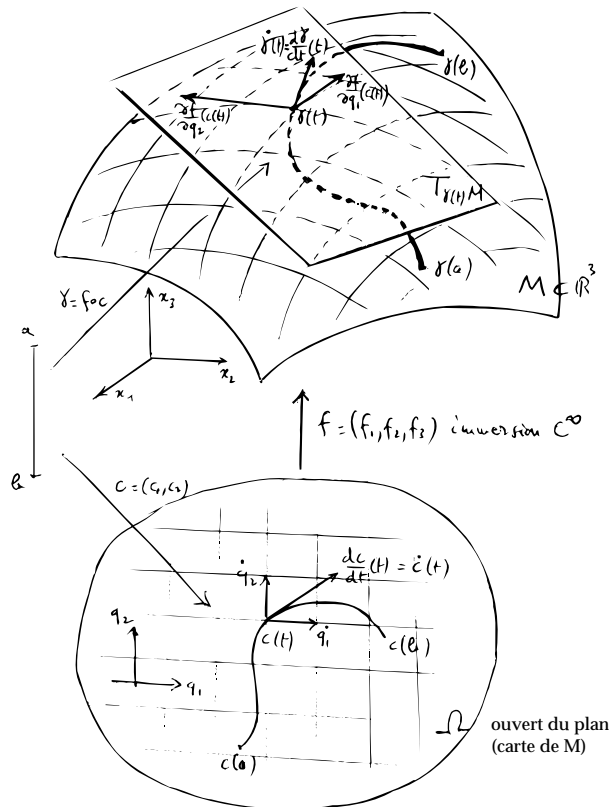


Figure 10 : une surface paramétrée

Une courbe \mathcal{C} , image de $\gamma = f \circ c$, tracée sur M peut être considérée comme

habitant dans \mathbb{R}^3 ; sa longueur (euclidienne) est donc définie et vaut

$$\mathcal{L}(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt = \int_a^b \|df(c(t))\dot{c}(t)\| dt.$$

La formule s'écrit encore

$$\mathcal{L}(\gamma) = \int_a^b \|\dot{c}(t)\|_{g(c(t))} dt = \mathcal{L}_g(c),$$

où \mathcal{L}_g désigne la longueur de la courbe c dans Ω muni de la métrique riemannienne g définie par

$$\|\dot{q}\|_{g(q)} = \sqrt{g(q)(\dot{q}, \dot{q})} = \sqrt{\langle df(q)\dot{q}, df(q)\dot{q} \rangle} = \|df(q)\dot{q}\|,$$

où le produit scalaire $\langle \cdot, \cdot \rangle$ et la norme $\|\cdot\|$ qui s'en déduit sont ceux de l'espace euclidien \mathbb{R}^3 . En chaque point q de Ω , $g(q)$ est une norme (et même un produit scalaire) sur les vecteurs \dot{q} tangents en q à Ω (i.e. sur les vecteurs d'origine q).

Définition. Une métrique riemannienne sur un ouvert Ω de \mathbb{R}^n est la donnée en chaque point $q \in \Omega$ d'un produit scalaire défini positif (i.e. d'une matrice symétrique $n \times n$) dont les coefficients $g_{ij}(q)$ dépendent de manière C^∞ de q :

$$\|\dot{q}\|_{g(q)}^2 = \langle \dot{q}, \dot{q} \rangle_{g(q)} = g(q)(\dot{q}, \dot{q}) = \sum_{i,j=1}^n g_{ij}(q) \dot{q}_i \dot{q}_j.$$

Une géodésique est une extrémale de l'intégrale d'énergie $\mathcal{E}_g(\gamma) = \frac{1}{2} \int_a^b \|\dot{\gamma}\|_g^2 dt$.

Dans notre cas, $g_{ij}(q) = \left\langle \frac{\partial f}{\partial q_i}(q), \frac{\partial f}{\partial q_j}(q) \right\rangle$. Une métrique euclidienne sur \mathbb{R}^n est caractérisée par le fait que ses coefficients sont indépendants de q ; la métrique canonique est définie par la matrice identité : $g_{ij}(q) = \delta_{ij}$ (symbole de Kronecker).

Ceci conduit directement à la notion générale de métrique riemannienne et de géodésique sur une variété, la seule difficulté étant de donner un sens à la différentiabilité du produit scalaire $g(x)$ sur $T_x M$ comme fonction du point x sur la variété (le problème est que l'espace vectoriel $T_x M$ sur lequel est défini $g(x)$ dépend lui-même de x). Notons cependant qu'un théorème célèbre (existence d'un plongement isométrique : Cartan, Nash, Gromov) dit que le gain en généralité n'est qu'illusoire.

Remarque. L'exemple de la sphère, dont les géodésiques sont les grands cercles, montre qu'une géodésique ne minimise pas nécessairement la longueur ou l'énergie entre deux de ses points. Elle ne le fait que si les deux points sont suffisamment proches l'un de l'autre.

2.1 Calcul des variations

Les intégrales de longueur et d'énergie d'une courbe dans un espace muni d'une métrique riemannienne sont des cas particuliers d'intégrales d'action qui,

à une courbe $t \mapsto c(t) \in \Omega$, associent l'intégrale

$$\mathcal{S}(c) = \int_a^b L(c(t), \frac{dc}{dt}(t)) dt$$

d'un *Lagrangien* $L(q, \dot{q})$. L est une fonction définie sur l'ensemble des vecteurs tangents à Ω (penser à (q, \dot{q}) comme définissant un vecteur \dot{q} d'origine q). On note brièvement l'action $\mathcal{S} = \int_a^b L(q, \dot{q}) dt$.

Reprenons le calcul ayant conduit en 1.2 à la démonstration de l'assertion fondamentale : $c_u(t) = q(u, t)$ est une famille de courbes (supposées régulières pour simplifier mais on pourrait traiter les sauts de la dérivée comme précédemment), la famille des vitesses est $\dot{c}_u(t) = \frac{\partial q}{\partial t}(u, t)$, et finalement celles des variations est $X_u(t) = \frac{\partial q}{\partial u}(u, t)$. Un calcul en tous points identiques à celui effectué dans le cas euclidien, donc basé sur l'intégration par parties, donne

$$\frac{d\mathcal{S}}{du} = \int_a^b \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \cdot X_u dt + \left[\frac{\partial L}{\partial \dot{q}} \cdot X_u \right]_a^b,$$

où la notation $[Y]_a^b$ signifie $Y(b) - Y(a)$. Pour une famille d'extrémales de $\int L dt$, le premier terme de la somme s'annule identiquement et le deuxième membre se réduit aux termes qui proviennent de la variation des extrémités :

$$\frac{d\mathcal{S}}{du} = \left[p \cdot \frac{\partial q}{\partial u} \right]_a^b, \text{ ou encore } d\mathcal{S} = [p \cdot dq]_a^b,$$

où p désigne la *variable conjuguée* de q , définie par $p(q, \dot{q}, t) = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t)$.

Pour le Lagrangien $L(q, \dot{q}) = E_g(q, \dot{q}) = \frac{1}{2} \|\dot{q}\|_{g(q)}^2$ d'une métrique riemannienne (celui qui correspond à l'intégrale d'énergie), cela devient

$$d\mathcal{S}(c)X = [g(\dot{q}, X)]_a^b,$$

ce qui implique l'*orthogonalité des rayons (= géodésiques) aux fronts d'onde pour une métrique arbitraire (lemme de Gauss en géométrie, théorie de Hamilton-Jacobi en mécanique)*. Dans le cas général, ce simple calcul est à la base de la théorie des *invariants intégraux* (Poincaré, Cartan) et de la structure *symplectique* (Lagrange, Hamilton) des équations de la mécanique conservative (i.e. sans frottement).

2.2 L'équation des géodésiques

Prenant $L = E_g$, on obtient les équations des géodésiques $q = c(t)$ sous la forme

$$\frac{d}{dt} \left(\sum_{j=1}^n g_{ij}(c(t)) \dot{c}_j(t) \right) = \frac{1}{2} \sum_{j,k=1}^n \frac{\partial g_{jk}}{\partial q_i}(c(t)) \dot{c}_j(t) \dot{c}_k(t), \quad i = 1, 2 \dots n,$$

ou encore, avec une notation raccourcie mais non ambiguë,

$$\sum_{j=1}^n g_{ij} \ddot{q}_j = \frac{1}{2} \sum_{j,k=1}^n \left[\frac{\partial g_{jk}}{\partial q_i} - 2 \frac{\partial g_{ij}}{\partial q_k} \right] \dot{q}_j \dot{q}_k, \quad i = 1, 2 \dots n.$$

On écrit classiquement ces équations sous une forme plus symétrique en égalant les coefficients de $\dot{q}_j \dot{q}_k$ et $\dot{q}_k \dot{q}_j$:

$$\sum_{j=1}^n g_{ij} \ddot{q}_j = \frac{1}{2} \sum_{j,k=1}^n \left[\frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k} - \frac{\partial g_{ik}}{\partial q_j} \right] \dot{q}_j \dot{q}_k, \quad i = 1, 2 \dots n.$$

La matrice $G(q)$ des $g_{ij}(q)$ étant inversible pour tout q , on obtient une équation différentielle du second ordre sur Ω qui généralise l'équation $\ddot{q}_i = 0$, $i = 1, \dots, n$:

$$\ddot{q}_i + \sum_{j,k=1}^n \Gamma_{jk}^i(q) \dot{q}_j \dot{q}_k = 0, \quad i = 1, 2 \dots n,$$

dans laquelle les *symboles de Christoffel* $\Gamma_{jk}^i(q) = \Gamma_{kj}^i(q)$ sont définis par

$$\forall j, k, \quad \begin{pmatrix} \Gamma_{jk}^1 \\ \Gamma_{jk}^2 \\ \vdots \\ \Gamma_{jk}^n \end{pmatrix} = G^{-1} \begin{pmatrix} \frac{1}{2} \left(\frac{\partial g_{1k}}{\partial q_j} + \frac{\partial g_{1j}}{\partial q_k} - \frac{\partial g_{jk}}{\partial q_1} \right) \\ \frac{1}{2} \left(\frac{\partial g_{2k}}{\partial q_j} + \frac{\partial g_{2j}}{\partial q_k} - \frac{\partial g_{jk}}{\partial q_2} \right) \\ \vdots \\ \frac{1}{2} \left(\frac{\partial g_{nk}}{\partial q_j} + \frac{\partial g_{nj}}{\partial q_k} - \frac{\partial g_{jk}}{\partial q_n} \right) \end{pmatrix},$$

c'est-à-dire

$$\Gamma_{jk}^i = \frac{1}{2} \sum_{l=1}^n g^{il} \left(\frac{\partial g_{kl}}{\partial q_j} + \frac{\partial g_{lj}}{\partial q_k} - \frac{\partial g_{jk}}{\partial q_l} \right),$$

où la matrice des $(g^{ij}(q))$ est l'inverse $G^{-1}(q)$ de $G(q)$.

Remarques.

1 - Si $f : \Omega \rightarrow \mathbb{R}^N$ est un paramétrage d'une sous-variété de \mathbb{R}^N , les symboles de Christoffel s'écrivent

$$\forall j, k, \quad \begin{pmatrix} \Gamma_{jk}^1 \\ \Gamma_{jk}^2 \\ \vdots \\ \Gamma_{jk}^n \end{pmatrix} = G^{-1} \begin{pmatrix} \frac{\partial^2 f}{\partial q_j \partial q_k} \cdot \frac{\partial f}{\partial q_1} \\ \frac{\partial^2 f}{\partial q_j \partial q_k} \cdot \frac{\partial f}{\partial q_2} \\ \vdots \\ \frac{\partial^2 f}{\partial q_j \partial q_k} \cdot \frac{\partial f}{\partial q_n} \end{pmatrix}.$$

2 - Chaque minimum de l'intégrale de longueur \mathcal{L}_g (à extrémités fixées) est une courbe régulière. Comme dans le cas euclidien, cette courbe, reparamétrisée à vitesse constante, devient un minimum de l'intégrale d'énergie \mathcal{E}_g . La démonstration est basée sur la généralisation à une métrique riemannienne quelconque du lemme de Gauss (ou de Hamilton-Jacobi), et du fait que 2 points assez proches

sont joints par un unique segment de géodésique (extrémum de \mathcal{E}_g) qui est l'unique courbe réalisant le min de la longueur entre ces deux points, ce qui montre qu'un min de \mathcal{L}_g est un segment de géodésique. Réciproquement, un min de \mathcal{E}_g n'a pas de *point conjugué* et appartient donc à un *champ*, c'est-à-dire à un système de coordonnées ondes-rayons (c'est le *théorème de Weierstrass*). Le raisonnement de 1.2 montre alors qu'il minimise la longueur \mathcal{L}_g .

3 - Dans des “coordonnées normales de Riemann” en un point q_0 , c'est-à-dire dans des coordonnées ondes sphériques-rayons comme sur la figure 4, centrées en q_0 , les symboles de Christoffel $\Gamma_{jk}^i(q_0)$ sont tous nuls! Si on remplace la métrique riemannienne par une “métrique lorentzienne” (i.e. un produit scalaire défini positif en tout point par un produit scalaire de signature $(-+++)$ dans l'espace-temps), les coordonnées normales de Riemann sont la réalisation d'un “repère localement inertiel”, tel celui qu'on obtient dans un ascenseur en chute libre (voir [MTW], par. 11.6)

2.3 Des sous-variétés aux variétés

Définition bootstrap : une variété est un espace qui ressemble localement à une variété. Plus on avance et plus on a de modèles. La seule différence avec ce qui précède est la non-existence de coordonnées globales, donc la nécessité de comprendre comment se transforment les calculs lorsqu'on passe d'une carte à l'autre. De plus un théorème de Whitney affirme qu'une variété peut toujours être considérée (non canoniquement!) comme sous-variété de \mathbb{R}^N , N grand.

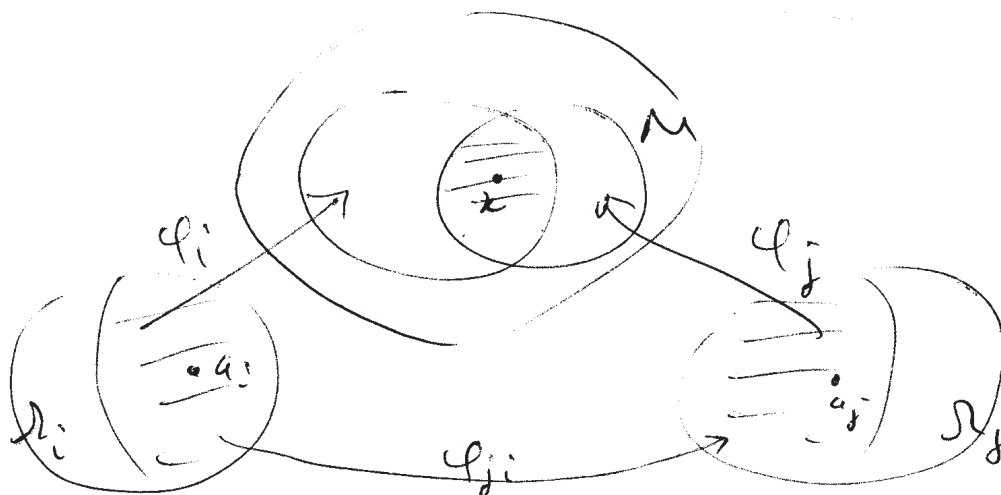


Figure 11 : cartes et changements de carte

Chemins et fonctions, vecteurs et covecteurs, champs de vecteurs et 1-formes différentielles, métriques riemanniennes

Chemins et fonctions sont les outils premiers de l'analyse des variétés ; leurs versions infinitésimales sont les *vecteurs* et *covecteurs* tangents (vitesse, dérivée), d'où les notions de *champs de vecteurs* et *1-formes différentielles*.

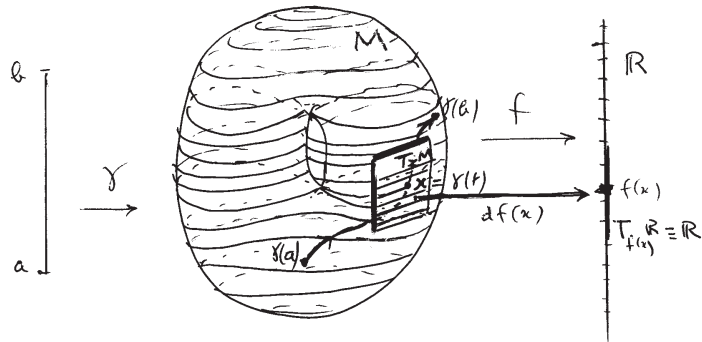


Figure 12 : chemin et fonction

Un champ de vecteurs X (resp. une 1-forme différentielle ω) sur une variété M est la donnée en chaque point $x \in M$ d'un vecteur tangent $X(x) \in T_x M$ (resp. d'un covecteur tangent $\omega(x) \in T_x^* M$, c'est-à-dire d'une forme linéaire $\omega(x) : T_x M \rightarrow \mathbb{R}$). La seule difficulté est de définir la régularité (C^1 , C^k , C^∞ , analytique) d'un tel objet, à valeurs dans un espace vectoriel variable. Dans une carte locale (coordonnées q_1, \dots, q_n), c'est facile car les translations permettent d'identifier canoniquement entre eux les espaces tangents ou cotangents en tous les points ; autrement dit, X et ω s'écrivent respectivement

$$X(q) = \sum_{i=1}^n X_i(q) \frac{\partial}{\partial q_i} = (X_1(q), \dots, X_n(q)), \quad \omega(q) = \sum_{i=1}^n \omega_i(q) dq_i.$$

Enfin, la règle de dérivation des applications composées implique l'indépendance de la notion de régularité ainsi définie du choix de la carte.

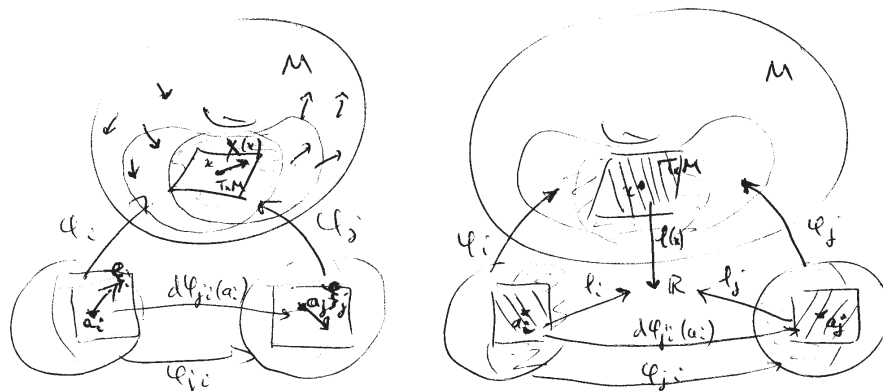


Figure 13 : champ de vecteurs et 1-forme différentielle

Enfin, une *métrie riemannienne* (resp. *lorentzienne*) sur une variété M est la donnée dans chaque espace tangent $T_x M$ d'un produit scalaire défini positif (resp. de signature $-, + + \dots$) qui varie "de façon régulière" avec le point x . Ici encore, la seule difficulté est de définir la "régularité", par exemple en le faisant dans une carte (c'est une propriété locale) puis en montrant l'indépendance du choix de la carte.

Nous avons déjà mentionné une différence fondamentale entre champs et 1-formes, à savoir l'intégrabilité locale de tout champ assez régulier (lipschitzien suffit) et la non-intégrabilité en général du champ d'hyperplans tangents à M défini par les noyaux d'une 1-forme différentielle : tout champ régulier est le champ des vitesses d'une famille de courbes paramétrées qui remplissent M alors qu'il n'existe en général pas de famille de surfaces p -dimensionnelles tangente à un champ de p -plans donné sur M , même très régulier (et même si $M = \mathbb{R}^3$). Nous reviendrons sur ce fait, à l'origine de l'existence de la courbure. Cependant, la donnée d'une métrie riemannienne fournit une identification entre champs et 1-formes ; en particulier le *gradient d'une fonction* f est le champ de vecteurs défini à partir de la 1-forme différentielle df , dérivée de f , par la formule

$$df \cdot X = g(\text{grad} f, X).$$

Une métrie riemannienne sur M fournit donc un isomorphisme du fibré tangent sur le fibré cotangent, cas particulier du *difféomorphisme de Legendre* en mécanique hamiltonienne et des *isomorphismes musicaux* en géométrie riemannienne.

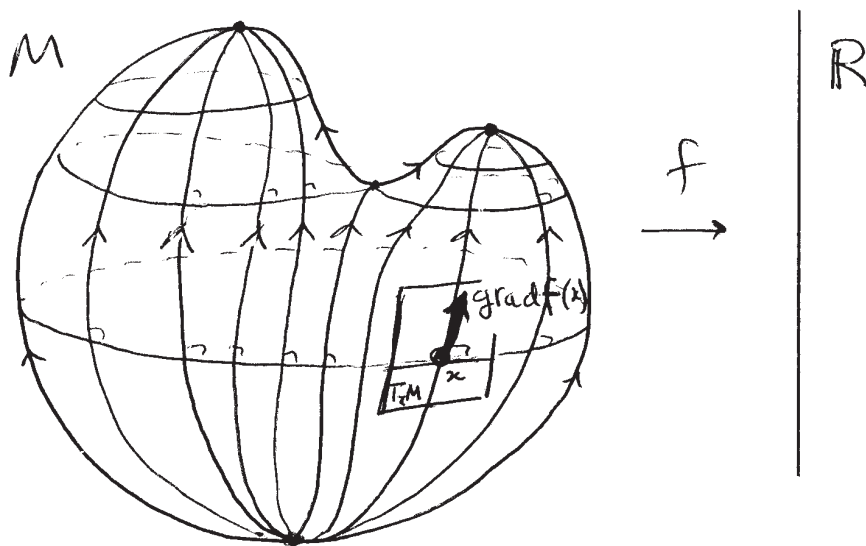


Figure 14 : courbes intégrales du gradient de f = "lignes de plus grande pente"

3 Du voisinage d'un point au voisinage d'un autre : transport parallèle et dérivation covariante

Venant en amont de la notion de métrique, la notion de *connexion* (Levi-Civita, Cartan, Schouten, Ehresmann, ...) est l'une des plus fondamentales de la physique et des mathématiques. C'est elle qui permet aux différentes *monades* que sont les voisinages infinitésimaux des divers points d'une variété de se parler. Une nouvelle définition des géodésiques d'une surface va nous conduire à la notion de développement, puis à celle de transport parallèle, de dérivée covariante et enfin à celle de connexion.

Dans ce qui suit, géodésique signifie extrémale de l'intégrale d'énergie \mathcal{E}_g ; en particulier, une géodésique est paramétrée de façon à ce que sa vitesse soit constante.

3.1 Géodésiques d'une surface et mouvements sans gravité

Les géodésiques d'une surface M pour la métrique induite par la métrique euclidienne de \mathbb{R}^3 , sont les courbes paramétrées $\gamma : [a, b] \rightarrow M$ dont l'accélération est normale à la surface. En effet, si M est définie par une équation régulière $F = 0$, une courbe γ tracée sur M est une géodésique pour la métrique induite si elle rend extrémale l'intégrale d'énergie $\int_a^b \|\dot{\gamma}(t)\|^2 dt$ parmi les courbes γ soumises à la contrainte (la *liaison*) $F \circ \gamma = 0$. Autrement dit,

$$d\mathcal{E}_g(\gamma)X = \int_a^b \langle \dot{\gamma}(t), \dot{X}(t) \rangle dt = - \int_a^b \langle \ddot{\gamma}(t), X(t) \rangle dt$$

doit être nul pour toutes les variations infinitésimales X de γ tangentes à M , c'est-à-dire telles que $d\mathcal{F}(\gamma)X = 0$, où l'équation des contraintes \mathcal{F} est définie par $\mathcal{F}(\gamma) = F \circ \gamma$. Il faut donc que $\ddot{\gamma}(t)$ soit proportionnel à la normale en $\gamma(t)$ à la surface, le facteur de proportionnalité (qui dépend de t) jouant le rôle d'une infinité de *multiplicateurs de Lagrange*. C'est le *Principe des travaux virtuels* de d'Alembert. L'interprétation dynamique est qu'*en l'absence de gravitation, une particule matérielle astreinte à se déplacer sur la surface suit une géodésique* : la seule force s'exerçant sur une telle particule est en effet la contrainte, normale à la surface. Dans ce qui suit, on donne au contraire une interprétation purement cinématique des géodésiques, dont sortiront les notions que nous avons en vue :

3.2 Roulements sans glissement ni pivotement

Soient M et M' deux surfaces régulières de l'espace euclidien \mathbb{R}^3 et soit $\gamma' : [a, b] \rightarrow M'$ une courbe paramétrée de M' . Appelons *roulement* de M sur M' le long de γ' , un chemin $t \mapsto F_t$ d'isométries (déplacements) de \mathbb{R}^3 tel qu'à chaque instant $t \in [a, b]$, l'image M_t de M par F_t soit tangente à M' en $\gamma'(t)$. Soit $\gamma : [a, b] \rightarrow M$ la *trace* de γ' sur M : $F_t(\gamma(t)) = \gamma'(t)$.

Le roulement est dit *sans glissement* en t_0 si la dérivée de l'application $t \mapsto F_t(\gamma(t_0))$ s'annule en t_0 (figure 15). Il est dit *sans pivotement* en t_0 si la *rotation instantanée* se fait autour d'un axe $\omega(t_0)$ parallèle au plan tangent à M' en $\gamma'(t_0)$. On peut montrer qu'il existe un unique roulement sans glissement ni pivotement de M sur M' le long de γ' , donc une unique trace γ laissée par γ' sur M (on doit bien entendu admettre que M et M' puissent se traverser l'une l'autre). De plus, on déduit de la caractérisation ci-dessus des géodésiques, le

Théorème. Soient M et M' deux surfaces de \mathbb{R}^3 et $\gamma : [a, b] \rightarrow M'$ une géodésique de M' pour la métrique induite par la métrique euclidienne de \mathbb{R}^3 . Si $\gamma : [a, b] \rightarrow M'$ est la trace laissée sur M' par γ' lorsque M roule sans glissement ni pivotement sur M' le long de $\gamma'([a, b])$, γ est une géodésique de M pour la métrique induite par la métrique euclidienne de \mathbb{R}^3 . En effet, si $Y(t) = \dot{\gamma}(t)$ est le champ des vitesses de γ et $Y'(t) = \dot{\gamma}'(t)$ celui de γ' , la condition de roulement sans glissement ni pivotement implique l'équivalence de l'orthogonalité de $\frac{dY}{dt}(t)$ à M avec celle de $\frac{dY'}{dt}(t)$ à M' .

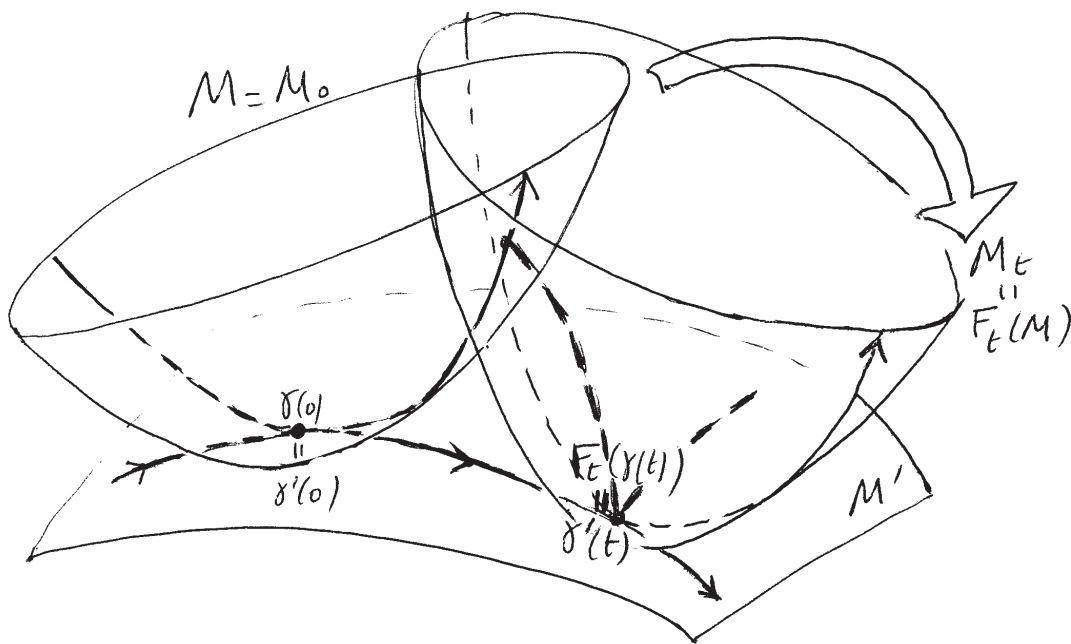


Figure 15 : roulement sans glissement ni pivotement

En particulier, prenant pour M' un plan, disons le plan horizontal $\mathbb{R}^2 \times \{0\}$ dans $\mathbb{R}^2 \times \mathbb{R} \equiv \mathbb{R}^3$, on obtient une géodésique de M en faisant rouler M sur le plan horizontal sans glissement ni pivotement le long d'une droite D' (figure 17). Un tel mouvement est parfaitement déterminé par la donnée à un instant fixé du point de contact de M avec le plan et du vecteur tangent à M en ce point porté

par la droite D' .

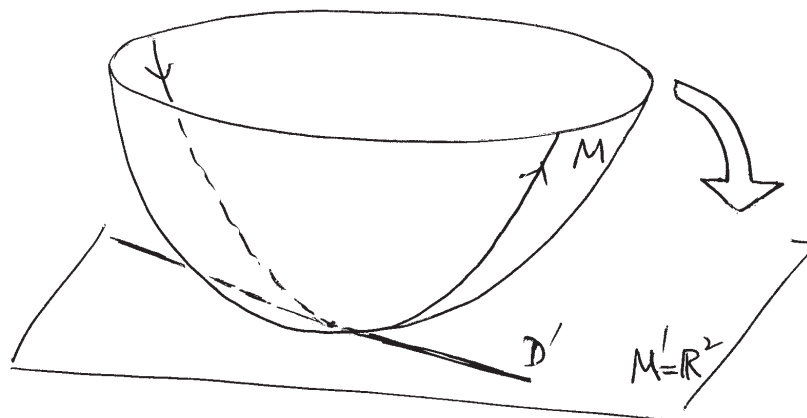


Figure 16 : engendrement d'une géodésique à partir d'une droite

3.3 Transport parallèle.

Pour définir la dérivée $\frac{dX}{dt}$ d'un champ de vecteurs le long d'une courbe γ dans \mathbb{R}^n , on a jusqu'ici implicitement identifié par translation le champ à une application de $[a, b]$ dans \mathbb{R}^n ; autrement dit, on transporte parallèlement chaque vecteur $X(t)$ de façon à ce que tous aient la même origine. C'est ainsi que l'on peut définir l'accélération d'un mobile et cet exemple montre qu'il est naturel de considérer que $\frac{dX}{dt}(t)$ est un vecteur d'origine $\gamma(t)$, donc que $\frac{dX}{dt}$ est, comme X , un champ de vecteurs le long de γ . L'analogue des translations existe dans une variété riemannienne, mais seulement le long d'une courbe. Il permet de dériver un champ de vecteurs (ou de tenseurs) le long d'une courbe (c'est la dérivée covariante) mais pas en général le long d'une surface à cause de l'existence de courbure.

C'est en 1917 que Levi-Civita découvre que la notion de roulement sans glissement ni pivotement (le *développement*) permet de définir la notion de champ de vecteurs parallèle le long d'une courbe intrinsèquement, c'est-à-dire uniquement en termes de la métrique g sur la surface : un champ de vecteurs $t \mapsto X(t)$ le long d'une courbe quelconque $t \mapsto \gamma(t)$ du plan euclidien \mathbb{R}^2 est dit *parallèle* si les vecteurs $X(t_1)$ et $X(t_2)$ sont parallèles quels que soient t_1 et t_2 , c'est-à-dire si $\frac{dX}{dt}$ est identiquement nulle.

On définit un champ de vecteurs parallèle Y tangent à M le long de γ comme la trace d'un champ de vecteurs parallèle Y' du plan par un roulement sans glissement ni pivotement de M sur le plan le long d'une courbe γ' (figure 18). Notons que $\frac{dY'}{dt} = 0$ étant orthogonal au plan, il suit du paragraphe précédent qu'un champ parallèle Y tangent à M le long d'une courbe γ est caractérisé par l'orthogonalité pour tout t de $\frac{dY}{dt}(t)$ (calculée dans \mathbb{R}^3) au plan tangent en M au

point $\gamma(t)$. Il suit du calcul des symboles de Christoffel donné dans la remarque 5 que, quels que soient j, k ,

$$\frac{\partial^2 f}{\partial q_j \partial q_k} = \sum_{i=1}^n \Gamma_{jk}^i \frac{\partial f}{\partial q_i} + \vec{N}_{jk},$$

où N_{jk} est normal à M . On en déduit que, si $Y(t) = df(c(t)) \cdot (y_1(c(t)), \dots, y_n(c(t)))$,

$$\frac{dY}{dt}(t) = \sum_{i=1}^n \left(\dot{y}_i + \sum_{j,k=1}^n \Gamma_{jk}^i(c(t)) \dot{c}_j(t) y_k(t) \right) \frac{\partial f}{\partial q_i}(c(t)) + Z(t),$$

où $Z(t)$ est normal à M . Le champ Y le long de c est donc parallèle si et seulement s'il vérifie les équations suivantes :

$$\dot{y}_i + \sum_{j,k=1}^n \Gamma_{jk}^i(c(t)) \dot{c}_j(t) y_k(t) = 0.$$

L'équation des champs parallèles apparaît donc comme une "polarisation" de l'équation, quadratique en les vitesses, des géodésiques et les géodésiques sont les courbes dont le champ des vitesses est parallèle.

C'est d'ailleurs par cette définition que commencent aujourd'hui les cours de géométrie riemannienne.

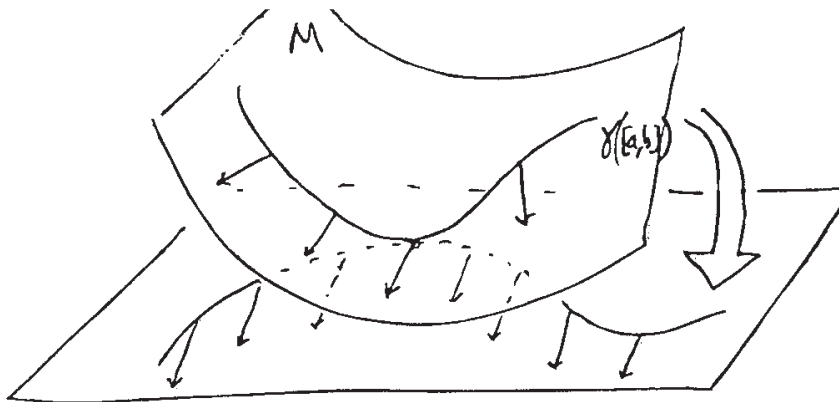


Figure 17 : engendrement d'un champ parallèle par roulement

L'opération ci-dessus définit de façon non ambiguë le *transport parallèle* d'un vecteur le long d'une courbe de M : on développe la courbe sur un plan, on transporte parallèlement le vecteur dans le plan et on revient en arrière. Par construction, le produit scalaire de deux vecteurs tangents en un même point, donc aussi la norme d'un vecteur, se conserve par transport parallèle. On peut ainsi construire des repères orthonormés parallèles le long d'une courbe.

Une méthode commode de construction du transport parallèle le long d'une courbe est de remarquer que l'on peut remplacer la surface M par le cône \mathcal{C} obtenu en prenant l'enveloppe des plans tangents à M le long de C (figure 18).

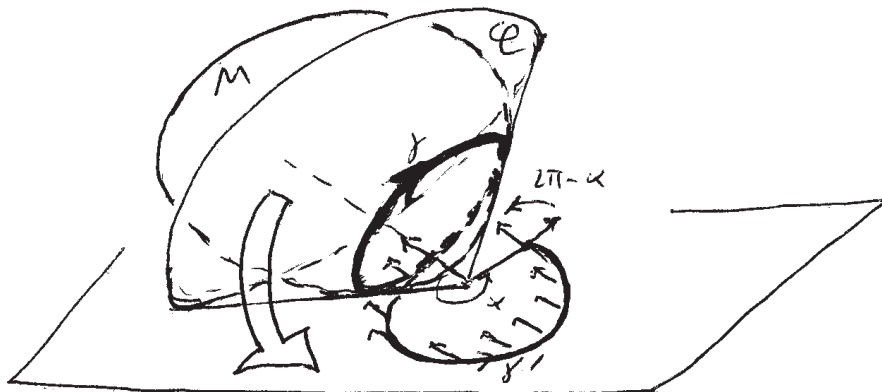


Figure 18 : développement via un cône

Remarque. Les découpages que l'on offre aux enfants sont un exemple de développement. On ne sait d'ailleurs toujours pas (sic!!! voir [Mi]) si le développement sans recouvrement d'un polyèdre convexe est possible en coupant celui-ci suivant ses arêtes.

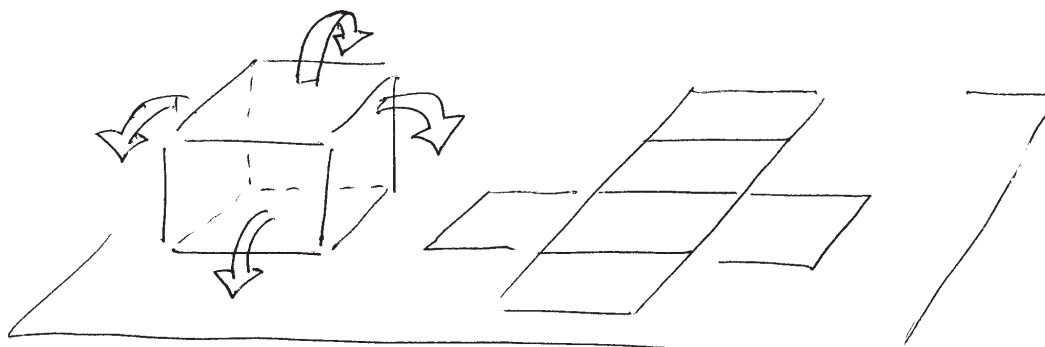


Figure 19 : développement d'un cube

3.4 Dérivation covariante.

C'est une mesure du non-parallélisme d'un champ de vecteurs tangent à $M \subset \mathbb{R}^3$ le long d'une courbe γ . Ce qui précède conduit à la

Définition. La dérivée covariante $\frac{DY}{dt}$ du champ de vecteurs Y le long de γ est le champ de vecteurs tangent à M le long de γ obtenu en projetant orthogonalement, en chaque point t , la dérivée $\frac{dY}{dt}(t)$ sur l'espace tangent $T_{\gamma(t)}M$.

Suivant l'expression de Do Carmo, $\frac{DY}{dt}$ est la dérivée $\frac{dY}{dt}$ vue de M (figure 20). La définition est manifestement faite pour qu'un champ de vecteurs tangent à une surface M le long d'une courbe γ soit parallèle si et seulement si sa dérivée covariante est le champ identiquement nul le long de γ .

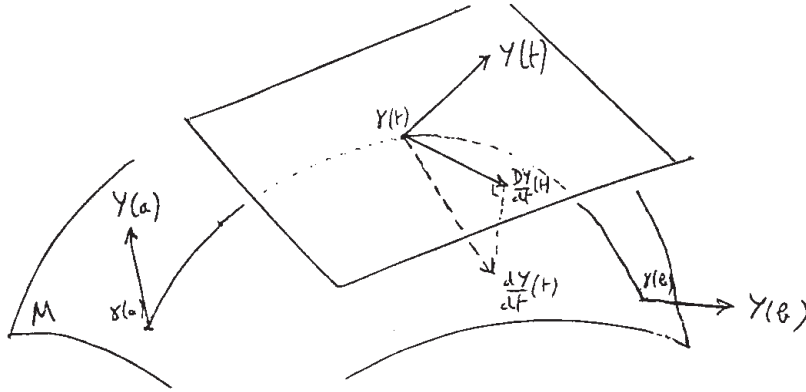


Figure 20 : la dérivée covariante comme projection (1)

Supposons que, dans une carte locale (un paramétrage) f de M , la courbe γ soit représentée par la courbe c et le champ de vecteurs Y le long de γ par le champ $y(t) = (y_1(t), y_2(t))$ le long de c ; autrement dit, supposons que

$$\gamma(t) = f \circ c(t), \quad Y(t) = \sum_{i=1}^2 \frac{\partial f}{\partial q_i}(c(t)) y_i(t).$$

On déduit du calcul de $\frac{dY}{dt}$ dans le paragraphe précédent que la dérivée covariante $\frac{DY}{dt}$ de Y est représentée par le champ $\frac{Dy}{dt}$ le long de c qui s'écrit

$$\frac{Dy}{dt}(t) = \left(\dot{y}_1(t) + \sum_{j,k=1}^2 \Gamma_{jk}^1(c(t)) \dot{c}_j(t) y_k(t), \dot{y}_2(t) + \sum_{j,k=1}^2 \Gamma_{jk}^2(c(t)) \dot{c}_j(t) y_k(t) \right).$$

On retrouve en particulier l'équation des géodésiques sous la forme $\frac{D\dot{c}}{dt} = 0$, c'est-à-dire $\frac{D\dot{\gamma}}{dt} = 0$.

Les géodésiques sont donc les courbes dont l'accélération covariante est identiquement nulle.

Exercice. Montrer que la dérivation covariante possède les propriétés formelles d'une dérivation : si Y et Z sont deux champs de vecteurs tangents à M le long de la même courbe γ , et si $\alpha : [a, b] \mapsto \mathbb{R}$ est une fonction de classe C^1 , on a

$$(i) \quad \frac{D}{dt}(Y + Z) = \frac{DY}{dt} + \frac{DZ}{dt}, \quad (ii) \quad \frac{D}{dt}(\alpha Y) = \frac{d\alpha}{dt} Y + \alpha \frac{DY}{dt}.$$

4 Connexion

4.1 Aspect algébrique (Koszul, Nomizu)

On peut utiliser la formule donnée dans le paragraphe précédent pour définir dérivée covariante et parallélisme d'un champ de vecteurs le long d'une courbe tracée sur une variété riemannienne quelconque : si, dans une carte, la courbe γ est représentée par $c = (c_1, \dots, c_n)$ et le champ Y le long de γ par $y = (y_1, \dots, y_n)$, $\frac{DY}{dt}$ est le champ le long de γ représenté dans la carte par

$$\left(\frac{Dy}{dt}(t) \right) = \left(\dots, \dot{y}_i(t) + \sum_{j,k=1}^n \Gamma_{jk}^i(c(t)) \dot{c}_j(t) y_k(t), \dots \right).$$

Les géodésiques sont toujours les courbes dont le champ des vitesses est parallèle, c'est-à-dire de dérivée covariante identiquement nulle. Il resterait à vérifier (exercice pour les courageux) que cette définition de la dérivée covariante ne dépend pas du choix de la carte mais cette façon de procéder manque d'élégance et le besoin naît d'une définition un peu plus intrinsèque.

La formule ci-dessus montre que, comme pour la dérivée usuelle, $\frac{DY}{dt}(t_0)$, s'il dépend de $Y(t)$ au voisinage de $t = t_0$, ne dépend que de la valeur en t_0 de $\dot{\gamma}(t)$. Nous le noterons $\nabla_{\dot{\gamma}(t_0)} Y$. A deux champs de vecteurs X et Y de classe C^∞ sur M , on peut alors associer un nouveau champ de vecteurs $\nabla_X Y$, également de classe C^∞ , par la formule

$$(\nabla_X Y)(m) = (\nabla_{X(m)} Y)(m).$$

Notons que $\nabla_X Y$ est $C^\infty(M, R)$ -linéaire en X (propriété 1) ci-dessous). Compte tenu des propriétés formelles de la dérivation covariante exposées dans l'exercice 2) du paragraphe précédent, ceci rend naturelle la définition suivante, dans laquelle $\Gamma^\infty(TM)$ désigne l'espace des champs de vecteurs de classe C^∞ sur M (espace des sections de classe C^∞ du fibré tangent TM) :

Definition. Une connexion linéaire sur TM est une application $(X, Y) \mapsto \nabla_X Y$

$$\nabla : \Gamma^\infty(TM) \times \Gamma^\infty(TM) \rightarrow \Gamma^\infty(TM),$$

possédant les propriétés suivantes : quels que soient les fonctions f, g de classe C^∞ et les champs de vecteurs X, Y, Z sur M , on a

- 1) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z,$
- 2) $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z,$
- 3) $\nabla_X(fY) = f\nabla_X Y + (\partial_X f)Y.$

Choisissons des coordonnées locales $q_i, i = 1, \dots, n$, sur M , et notons $\frac{\partial}{\partial q_i}$ le champ de vecteurs constant sur \mathbb{R}^n obtenu en translatant le $i^{\text{ème}}$ vecteur $(0, \dots, 0, 1, 0, \dots, 0)$ de la base canonique. Etant donnés deux champs de vecteurs

$$X = (X_1, \dots, X_n) = \sum_{i=1}^n X_i \frac{\partial}{\partial q_i}, \quad Y = (Y_1, \dots, Y_n) = \sum_{i=1}^n Y_i \frac{\partial}{\partial q_i},$$

(les X_i et les Y_i sont des fonctions de q), on déduit des propriétés de ∇ que

$$\nabla_X Y = \sum_{j=1}^n X_j \nabla_{\frac{\partial}{\partial q_j}} Y = \sum_{j=1}^n X_j \left[\sum_{i=1}^n \frac{\partial Y_i}{\partial q_j} \frac{\partial}{\partial q_i} + \sum_{k=1}^n Y_k \nabla_{\frac{\partial}{\partial q_i}} \frac{\partial}{\partial q_k} \right],$$

donc

$$\nabla_X Y = \partial_X Y + \sum_{j,k=1}^n \nabla_{j k}^i X_j Y_k \frac{\partial}{\partial q_i},$$

où l'on a posé

$$\nabla_{\frac{\partial}{\partial q_j}} \frac{\partial}{\partial q_k} = \sum_{i=1}^n \nabla_{j k}^i \frac{\partial}{\partial q_i}.$$

Réciproquement, on vérifie que quelles que soient les fonctions $\nabla_{j k}^i$ de classe C^∞ à valeurs réelles définies sur le domaine de la carte, la formule ci-dessus définit une connexion ∇ sur ce domaine.

On en déduit que la dérivée covariante de la restriction à γ du champ Y , représentée dans la carte locale par $y(t) = Y(c(t))$, est

$$\frac{Dy}{dt}(t) = \left(\dots, \dot{y}_i(t) + \sum_{j,k=1}^n \nabla_{j k}^i(c(t)) \dot{c}_j(t) y_k(t), \dots \right).$$

La connexion de Levi-Civita. La connexion de Levi-Civita d'une variété riemannienne est définie dans chaque carte locale par les formules

$$\nabla_{j k}^i = \Gamma_{j k}^i.$$

On montre qu'elle est caractérisée par les propriétés suivantes :

1) la compatibilité avec la métrique, c'est-à-dire la conservation du produit scalaire par transport parallèle : rappelons que pour une surface de \mathbb{R}^3 le plan tangent est transporté par un déplacement rigide de \mathbb{R}^3 ;

2) la symétrie (ou nullité de la torsion), c'est-à-dire la validité pour toute application f de classe C^2 d'un ouvert de \mathbb{R}^2 dans M , de l'identité de Schwarz covariante

$$\forall (x_1, x_2) \in \Omega, \quad \frac{D}{\partial x_1} \frac{\partial f}{\partial x_2}(x_1, x_2) = \frac{D}{\partial x_2} \frac{\partial f}{\partial x_1}(x_1, x_2),$$

qui traduit l'absence de pivotement dans le roulement et qu'on en déduit du Lemme de Schwarz classique par projection sur TM . Notons que, compte tenu de la préservation de la métrique par la connexion, c'est cette condition qui

assure la validité du *Lemme de Gauss* (orthogonalité des rayons aux cercles riemanniens), c'est-à-dire la préservation par le flot géodésique de la forme de Liouville.

De nouveau l'équation des géodésiques. La compatibilité avec la métrique de la connexion de Levi-Civita se traduit par la propriété suivante de la dérivée covariante : pour tout couple de champs de vecteurs Y, Z le long d'une même courbe γ , on a

$$\frac{d}{dt} \langle Y(t), Z(t) \rangle_{g(\gamma(t))} = \left\langle \frac{DY}{dt}(t), Z(t) \right\rangle_{g(\gamma(t))} + \left\langle Y(t), \frac{DZ}{dt}(t) \right\rangle_{g(\gamma(t))}.$$

Obtenue par projection sur l'espace tangent dans le cas d'une surface de \mathbb{R}^3 , la formule se démontre dans le cas général à l'aide d'un champ parallèle de repères orthonormés le long de γ puisqu'alors la dérivation covariante se ramène à la dérivation ordinaire des coordonnées. On peut alors reprendre intrinsèquement le calcul de la variation première de l'intégrale d'énergie. Soit $u \mapsto \gamma_u$ une variation à extrémités fixées de $\gamma = \gamma_0$ à travers des courbes régulières et $X_u = \frac{\partial \gamma_u}{\partial u}$ la variation infinitésimale correspondante ; utilisant maintenant la symétrie de la connexion de Levi-Civita, on obtient

$$\begin{aligned} d\mathcal{E}_g(\gamma_u)X_u &= \frac{1}{2} \int_a^b \frac{d}{du} \|\dot{\gamma}_u\|_g^2 dt = \int_a^b \left\langle \frac{D\dot{\gamma}_u}{\partial u}, \dot{\gamma}_u \right\rangle_g dt = \int_a^b \left\langle \frac{D}{\partial u} \frac{\partial \gamma_u}{\partial t}, \dot{\gamma}_u \right\rangle_g dt \\ &= \int_a^b \left\langle \frac{D}{\partial t} \frac{\partial \gamma_u}{\partial u}, \dot{\gamma}_u \right\rangle_g dt = \int_a^b \left\langle \frac{DX_u}{\partial t}, \dot{\gamma}_u \right\rangle_g dt = [\langle X_u, \dot{\gamma}_u \rangle]_a^b - \int_a^b \left\langle X_u, \frac{D\dot{\gamma}_u}{\partial t} \right\rangle_g dt. \end{aligned}$$

On en déduit que γ est une extrémale de \mathcal{E}_g si et seulement si $\frac{D\dot{\gamma}}{dt} \equiv 0$.

4.2 Aspect géométrique (Ehresman)

Issue des travaux de E. Cartan, J.A. Schouten, et C. Ehresman, la définition géométrique des connexions allie une grande simplicité conceptuelle à une extrême généralité. Bien que notre propos soit local, commençons par une rapide description du monde des *espaces fibrés* car c'est là que la notion de connexion prend sa signification géométrique.

Fibrés vectoriels Le fibré tangent du cercle S^1 est l'exemple le plus simple. Afin de "voir" ensemble tous les vecteurs tangents en tous les points du cercle, on redresse chaque tangente orientée par une rotation de 90 degrés. On supprime ainsi les intersections deux à deux de ces tangentes et on fait apparaître la structure topologique naturelle de l'ensemble TS^1 des vecteurs tangents, qui est ici un cylindre, produit du cercle par la "fibre-type" que l'on peut identifier à \mathbb{R} une fois choisies une orientation du cercle et une mesure de longueur. C'est le triplet formé de S^1, TS^1 et de la projection $\pi : TS^1 \rightarrow S^1$ qui à un vecteur tangent fait correspondre son point de tangence, qu'on appelle le *fibré tangent du cercle*. On parle aussi du fibré π ou de la fibration π .

Un champ de vecteurs sur le cercle (i.e. un champ de vecteurs tangents au cercle) s'interprète maintenant comme une application $X : S^1 \rightarrow TS^1 \equiv S^1 \times \mathbb{R}$ qui fait correspondre à x un vecteur $X(x)$ tangent au point x , que l'on peut représenter par le couple du point x et d'une longueur algébrique portée sur la tangente orientée en x . On dit que X est une *section de π* . Ce qu'on a gagné est la possibilité de définir sans peine ce que signifie un champ de vecteurs dérivable, deux fois dérivable, etc...

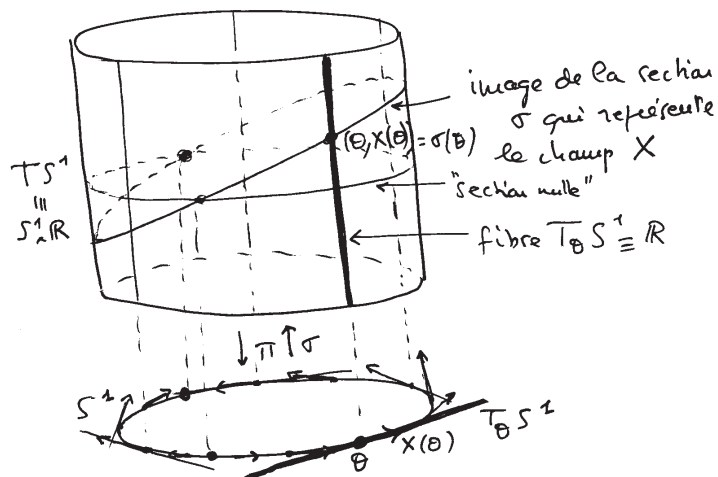


Figure 21 : le fibré tangent du cercle, un fibré trivial

De façon plus sophistiquée, on peut introduire la notion d'*espace fibré vectoriel* et définir deux variétés, le *fibré tangent* TM et le *fibré cotangent* T^*M en munissant l'ensemble des vecteurs tangents (resp. cotangents) d'une telle structure.

Un *fibré vectoriel* E sur M est la donnée d'une variété E (l'*espace total*) et d'une application différentiable $\pi : E \rightarrow M$ surjective (la *projection*) telle que la propriété suivante de *trivialité locale* soit satisfaite : les fibres $\pi^{-1}(x)$ sont munies d'une structure d'espace vectoriel et tout point de x possède un voisinage U tel que $\pi^{-1}(U)$ soit difféomorphe au produit $U \times F$ de U par un espace vectoriel F (la *fibre type*) par un difféomorphisme Φ dont la restriction Φ_y à la fibre $\pi^{-1}(y)$ est, pour tout $y \in U$, un isomorphisme linéaire de $\pi^{-1}(y)$ sur F . Le fibré est dit *trivial* si E est globalement difféomorphe au produit de M par F .

Une *section* du fibré est une application $\sigma : M \rightarrow E$ telle que $\pi \circ \sigma = \text{Id}_M$. On note $\Gamma^\infty(E)$ l'ensemble de sections de classe C^∞ de E . Cela a un sens dès que E et M sont des variétés de classe C^∞ .

Dans le cas du fibré tangent (resp. cotangent), la projection associée à un vecteur (resp. covecteur) tangent son point de tangence. Les fibres sont donc respectivement l'espace tangent $T_x M$ à M au point x et son dual $T_x^* M$ et les sections, dont on sait maintenant caractériser la régularité, sont respectivement les champs de vecteurs et les 1-formes différentielles. La trivialité locale vient de ce que si U est le domaine d'une carte locale de M , $\pi^{-1}(U)$ est isomorphe (i. e. difféomorphe

par un difféomorphisme linéaire sur les fibres) au fibré tangent (resp. cotangent) d'un ouvert Ω de \mathbb{R}^n . Mais ce dernier est canoniquement isomorphe au produit $\Omega \times \mathbb{R}^n$ (resp $\Omega \times (\mathbb{R}^n)^*$) car la translation de x identifie canoniquement $T_x\Omega$ à \mathbb{R}^n .

L'exemple caractéristique de fibré tangent non trivial est le fibré tangent de la sphère de dimension deux : la non-trivialité se traduit par l'impossibilité d'en trouver une section qui ne s'annule pas, c'est-à-dire un champ de vecteurs sur la sphère partout non nul : on ne peut pas peigner une sphère, il reste toujours un épi. La figure 22 donne un exemple plus simple et bien connu, le ruban de Mœbius ; comme dans le cas de TS^1 , les fibres sont de dimension 1

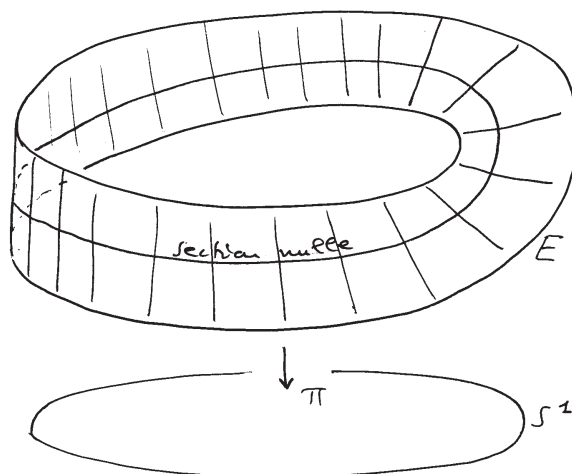


Figure 22 : la bande de Mœbius, un fibré vectoriel non trivial

Les connexions d'Ehresman. Soit M une variété riemannienne. Comme notre propos est local, nous pouvons aussi bien supposer que $M = \Omega$ est un ouvert de \mathbb{R}^n muni d'une connexion linéaire ∇ , c'est-à-dire de fonctions $\nabla_{jk}^i(q)$. Un point de TM s'écrit $(q, \dot{q}) \in T\Omega \equiv \Omega \times \mathbb{R}^n$. Introduisons le *tangent du tangent*, non pas pour faire plaisir à Edgar Morin et sa méthode des méthodes mais parce que cette inflation apparente va rendre géométriquement transparentes les notions de connexion, de dérivation covariante et de transport parallèle.

Puisque $T\Omega$ s'identifie à un ouvert de \mathbb{R}^{2n} , le fibré $TT\Omega$ est lui aussi trivial : $TT\Omega \equiv \Omega \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ et on y dispose de coordonnées naturelles (q, \dot{q}, Q, \dot{Q}) .

Appelons *vertical* le sous-espace vectoriel $V_{(q, \dot{q})}$ de $T_{(q, \dot{q})}T\Omega$ tangent à la fibre $T_q\Omega$ de la projection $\pi : T\Omega \rightarrow \Omega$. C'est le sous-espace vectoriel d'équation $Q = 0$.

Appelons *horizontal* le sous-espace vectoriel $H_{(q, \dot{q})}$ de $T_{(q, \dot{q})}T\Omega$, transverse au

sous-espace vertical, défini par les équations linéaires

$$\dot{Q}_i + \sum_{j,k=1}^n \nabla_{jk}^i(q) \dot{q}_k Q_j = 0, \quad i = 1, \dots, n.$$

Voici une bonne occasion d'utiliser le langage des formes différentielles : en effet, $H_{(q,\dot{q})} = \cap_{i=1}^n \text{Ker } \alpha_i(q, \dot{q})$, où les α_i sont les 1-formes différentielles définies localement sur TM par

$$\alpha_i \equiv d\dot{q}_i + \sum_{j,k=1}^n \nabla_{jk}^i(q) \dot{q}_j dq_k, \quad i = 1, \dots, n.$$

Pour parler en physicien, les équations $\alpha_i = 0$, $i = 1, \dots, n$ décrivent la "variation infinitésimale" ($d\dot{q}_1, \dots, d\dot{q}_n$) d'un vecteur tangent en q à Ω transporté parallèlement, en fonction de la "variation infinitésimale" (dq_1, \dots, dq_n) de son origine – après tout, on a simplement multiplié par dt les composantes de l'équation $\frac{DY}{dt}(t) = 0$, qui définit les champs de vecteurs parallèles.

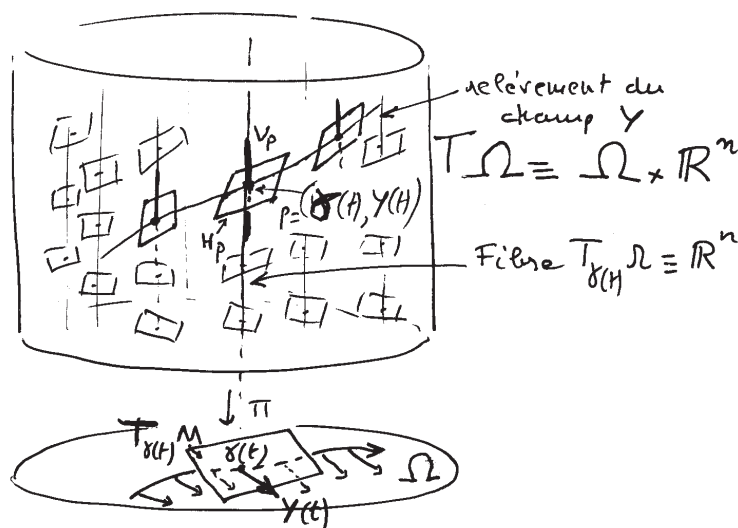


Figure 23 : sous-espaces verticaux et horizontaux, transport parallèle

En particulier, un champ parallèle le long d'une courbe n'est autre qu'une section horizontale de $T\Omega$ au-dessus de cette courbe et le transport parallèle d'un vecteur le long d'un chemin devient le relèvement horizontal du chemin à partir du point (q, \dot{q}) qui représente le vecteur.

Enfin, en termes des α_i , la $i^{\text{ème}}$ composante de la dérivée covariante d'un champ de vecteurs quelconque $Y(t) = (c_1(t), \dots, c_m(t), y_1(t), \dots, y_m(t))$ tangent à M le long d'une courbe $\gamma(t) = (c_1(t), \dots, c_m(t))$, s'écrit manifestement

$$\left(\frac{DY}{dt} \right)_i = (\alpha_i(Y)) \left(\frac{dY}{dt} \right),$$

où $\frac{dY}{dt}(t) = (\dot{c}_1(t), \dots, \dot{c}_m(t), \dot{y}_1(t), \dots, \dot{y}_m(t))$ est la “dérivée ordinaire” de l’application $Y : R \rightarrow E = TM$. La dérivée covariante $\frac{DY}{dt}(t)$ de Y apparaît ainsi comme la composition de la dérivée ordinaire $\frac{dY}{dt}(t)$ de l’application $Y : R \rightarrow E = TM$ avec le projecteur $\Phi : T_{Y(t)}E \rightarrow V_{Y(t)} \subset T_{Y(t)}E$ sur le sous-espace vertical $V_{Y(t)}$ parallèlement au sous-espace horizontal $H_{Y(t)}$, suivie de l’identification canonique par translation de $V_{Y(t)}$ à $E_{\gamma(t)} = T_{\gamma(t)}M$ (figure 22).

En résumé : le transport parallèle sur une variété riemannienne n’est que la manifestation intégrée d’une structure géométrique infinitésimale – un champ de sous-espaces horizontaux – sur le fibré tangent de la variété considérée, structure dont la donnée équivaut à celle de la connexion de Levi-Civita. La dérivation covariante qui lui est associée s’obtient en composant la dérivation ordinaire d’un champ de vecteurs, considéré comme une application de M dans TM , avec la projection sur les sous-espaces verticaux (i.e. sur les espaces tangents aux fibres de TM) parallèlement aux sous-espaces horizontaux (figure 24).

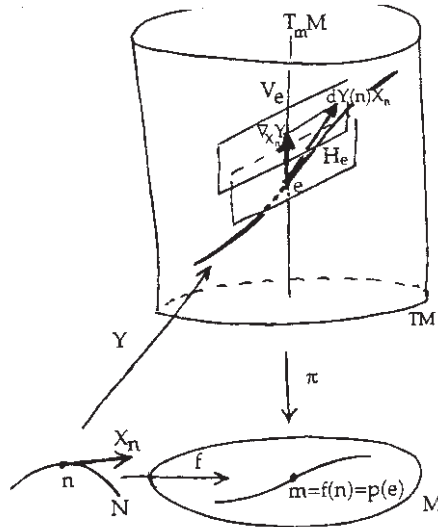


Figure 24 : la dérivé covariante comme projection (2)

4.3 Connexion dans un fibré

Comme il arrive souvent, c’est en généralisant une notion qu’on la rend plus immédiatement transparente. La notion de connexion conserve un sens si l’on remplace le fibré tangent TM par n’importe quel fibré vectoriel : mais définissons d’abord la notion générale d’espace fibré vectoriel.

Notant $\Gamma^\infty(E)$ l’espace des sections de classe C^∞ de $\pi : E \rightarrow M$, on définit une connexion linéaire sur le fibré vectoriel $E \rightarrow M$ comme une application

$$\nabla : \Gamma^\infty(TM) \times \Gamma^\infty(E) \rightarrow \Gamma^\infty(E),$$

notée $(X, \sigma) \mapsto (\nabla_X \sigma)$, $C^\infty(M, R)$ -linéaire en X et telle que $\nabla_X \sigma$ soit, pour chaque champ de vecteurs X , une dérivation de $\Gamma^\infty(E)$, ce qui est une façon sophistiquée de nommer les propriétés que nous avons décrites dans le cas où E est le fibré tangent.

Une définition géométrique équivalente est la donnée d'un *champ horizontal* $e \mapsto H_e$ de sous-espaces vectoriels tangents à E , qui équivaut comme on vient de le voir à celle d'une opération de *relèvement horizontal*.

En conclusion, si l'on prend E pour ce qu'il est, une famille d'espaces vectoriels E_m dépendant de façon C^∞ d'un paramètre $m \in M$, une *connexion apparaît comme une manière naturelle de généraliser la notion de dérivée aux applications à valeurs dans une famille d'espaces vectoriels, c'est-à-dire de définir dans cette situation une dérivée qui, comme la dérivée usuelle, soit à valeurs dans E* . Une bonne référence pour ce point de vue est [D]

Généralisation aux fibrés différentiables. La notion de connexion comme champ de sous-espaces tangents horizontaux garde un sens dans un fibré différentiable quelconque (les fibres sont maintenant des variétés et n'ont pas en général de structure vectorielle). La dérivée covariante est encore définie à valeurs dans $V_e = T_e E$ mais on ne dispose plus de l'identification de ce dernier espace à la fibre E_m de E . C'est la même situation que lorsqu'on définit la dérivée en un point n d'une application Y d'une variété N dans une autre E (cas où M est réduit à un point) comme étant à valeurs dans l'espace tangent $T_{Y(n)} E$, seule une structure vectorielle sur E permettant de définir une dérivée à valeurs dans E .

L'exemple le plus simple est celui de la restriction de la connexion de Levi-Civita au fibré unitaire tangent $T_1 M$ d'une variété riemannienne, c'est-à-dire à l'ensemble des vecteurs tangents à M de norme 1. Le fibré tangent unitaire de la sphère S^2 de dimension deux est difféomorphe au groupe $SO(3)$ des rotations, dont le revêtement à deux feuilletés s'identifie à la sphère S^3 (celle des quaternions de norme 1 par exemple). L'exemple qui suit est, sous un certain déguisement le même que celui qu'on obtiendrait par revêtement à deux feuilletés à partir de la connexion de Levi-Civita sur le fibré unitaire tangent de la sphère ronde.

4.4 Un exemple : la connexion naturelle de l'espace des triangles...

... ou comment une déformation à moment cinétique nul imite une rotation.

L'espace des *triangles orientés du plan*, c'est-à-dire l'espace des triplets ordonnés de points dans \mathbb{R}^2 s'identifie canoniquement au produit $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ de trois exemplaires de \mathbb{R}^2 c'est-à-dire à \mathbb{R}^6 .

L'espace des *triangles orientés du plan à translation près*, c'est-à-dire le quotient du précédent par l'action des translations, peut être représenté par les triangles dont le centre de gravité est fixé à l'origine lorsqu'on donne une masse unité à chacun des sommets. Il s'identifie naturellement à \mathbb{R}^4 par la construction

suivante, due à Jacobi : de la donnée du triangle orienté ABC , on ne retient que les deux vecteurs \overrightarrow{AB} et \overrightarrow{GC} , où G est le milieu du segment AB .

L'espace des *triangles orientés du plan à translation et rotation près* peut être identifié à \mathbb{R}^3 par l'*application de Hopf* de \mathbb{R}^4 sur \mathbb{R}^3 définie, après identification de \mathbb{R}^4 à \mathbb{C}^2 et de \mathbb{R}^3 à $\mathbb{R} \times \mathbb{C}$, par $H(z_1, z_2) = (|z_1|^2 - |z_2|^2, 2\bar{z}_1 z_2)$. On vérifie en effet que deux points (z_1, z_2) et (ζ_1, ζ_2) de \mathbb{C}^2 ont même image par H si et seulement s'il existe θ tel que $\zeta_1 = e^{i\theta} z_1$ et $\zeta_2 = e^{i\theta} z_2$.

Définissons la *taille* d'un triangle comme étant le tiers de la somme des carrés des longueurs de ses côtés (c'est-à-dire le *moment d'inertie* par rapport à son centre de gravité du solide formé par trois masses unité concentrées aux sommets). Les *triangles orientés du plan à translation et rotation près et de taille fixée* forment une sphère S^2 et la formule montre que cette sphère est l'image par H de la sphère S^3 de \mathbb{C}^2 d'équation $|z_1|^2 + |z_2|^2 = 1$, c'est-à-dire de l'espace des triangles orientés du plan à translation près et de taille fixée. La restriction de H à S^3 est la célèbre *fibration de Hopf* dont les fibres sont des cercles deux à deux enlacés dans S^3 : chaque cercle est l'ensemble des triangles qui se déduisent d'un triangle fixe par une rotation. On définit une connexion sur S^3 de la manière suivante : le relèvement horizontal d'un chemin de triangles est un mouvement $A(t), B(t), C(t)$ des trois masses unité A, B, C dans le plan dont le moment cinétique est identiquement nul : si O est le centre de gravité (supposé fixe),

$$\overrightarrow{OA} \wedge \dot{\vec{A}} + \overrightarrow{OB} \wedge \dot{\vec{B}} + \overrightarrow{OC} \wedge \dot{\vec{C}} = 0.$$

Le plan horizontal en chaque point n'est autre que le plan orthogonal à la fibre passant par ce point, pour le produit scalaire euclidien et le champ de plans ainsi obtenu est la *forme de contact standard* sur la sphère S^3 . La figure 25 illustre le revêtement horizontal d'un lacet de S^2 . Le chemin obtenu dans S^3 ne se ferme pas et le défaut de fermeture s'interprète comme une rotation du triangle de départ : une déformation a remplacé une rotation. L'angle de la rotation est proportionnel à l'aire entourée sur S^2 par le lacet.

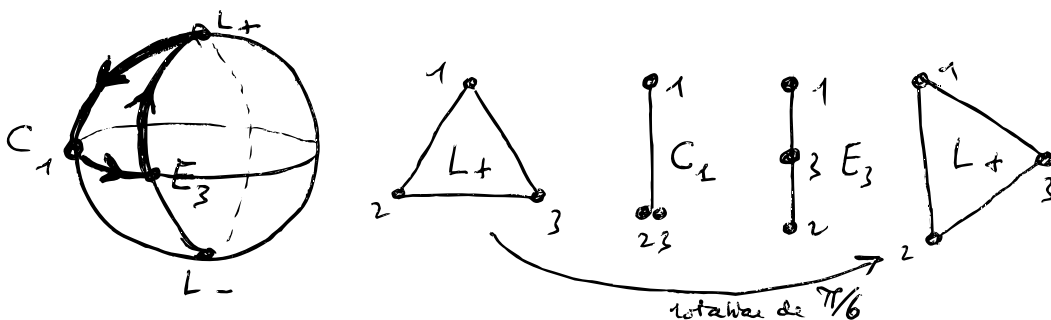


Figure 25 : bien connue des chats qui tombent d'un toit, la figure illustre l'holonomie de la connexion naturelle de l'espace des triangles orientés du plan.

5 Courbures : du théorème de Frobenius au tenseur de Riemann

The curvature tensor of a riemannian manifold is a little monster of (multilinear) algebra whose full geometric meaning remains obscure.

M. Gromov, *Sign and geometric meaning of curvature*, 1990.

Spacetime curvature manifests itself as gravitation, by means of the deviation of one geodesic from a nearby geodesic (relative acceleration of test particles).

C. Misner, K. Thorne, J.A. Wheeler, *Gravitation*, 1973.

La citation de Misha Gromov n'est là que pour effrayer le lecteur ; au niveau élémentaire auquel nous nous plaçons, le monstre s'apprivoise aisément. Par l'autre citation, il s'agit de rappeler l'importance fondamentale de la géométrie riemannienne dans les théories relativistes de la gravitation.

Soit $\pi : E \rightarrow M$ un fibré vectoriel sur une variété M de dimension n , par exemple le fibré tangent $E = TM$. Nous avons vu qu'une connexion ∇ sur E pouvait être définie par un champ de *sous-espaces horizontaux* $(H_\zeta)_{\zeta \in E}$ de dimension n tangents à E transverses aux fibres. Un tel champ n'est en général pas *intégrable* : il n'existe pas de sous-variétés de dimension n tangentes aux *sous-espaces horizontaux*. La traduction de cette non-intégrabilité est que le *relèvement horizontal* d'un lacet de M n'est en général pas un lacet. Pour la connexion de Levi-Civita, cela signifie que par transport parallèle d'un vecteur le long d'un lacet de M , un vecteur ne revient pas sur lui-même. On dit alors que la connexion ∇ a de la courbure. Techniquement, la courbure est une mesure de la non-intégrabilité, basée sur la notion de *crochet* de deux champs de vecteurs.

5.1 Un critère d'intégrabilité : le théorème de Frobenius

Le flot d'un champ de vecteurs. Commençons par quelques rappels sur les champs de vecteurs. Nous avons dit au début de ce cours qu'un vecteur tangent à une variété M est la version infinitésimale d'une courbe paramétrée. Intégrer un champ de vecteurs sur M , $x \rightarrow X(x) \in T_x M$, c'est déterminer une famille de courbes paramétrées $\varphi_t(x)$ dont la vitesse à chaque instant soit donnée par le vecteur du champ au point correspondant :

$$\frac{d\varphi_t}{dt}(x) = X(\varphi_t(x)).$$

Le théorème d'existence et d'unicité des solutions d'équations différentielles, base du déterminisme, affirme que l'application $(x, t) \mapsto \varphi_t(x)$ est bien définie et régulière. C'est le *flot* de X , dont on montre qu'il est défini pour tous temps lorsque M est compacte. La courbe $t \mapsto \varphi_t(x)$ est la solution (resp. la courbe

intégrale) de donnée initiale x . Une bonne référence pour la notion de flot est le livre [M].

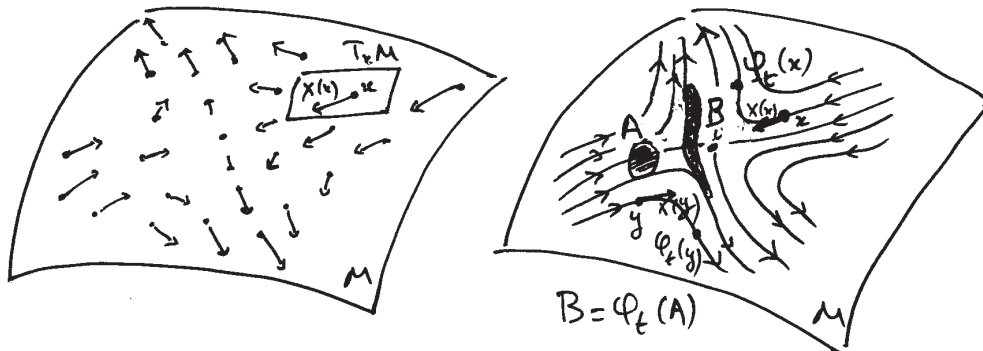


Figure 26 : un champ de vecteurs et son flot

Les champs comme dérivations; le crochet. Au champ de vecteurs X sur M est associé l'opérateur différentiel du premier ordre sur les fonctions $f \in C^\infty(M, \mathbb{R}) : f \mapsto \partial_X f$, où

$$\partial_X f(x) = \lim_{t \rightarrow 0} \frac{1}{t} (f(\varphi_t(x)) - f(x)).$$

Il suit du théorème de dérivation des applications composées que la fonction $\partial_X f$ n'est autre que la dérivée $df \cdot X$ de f suivant X ; elle est encore appelée *dérivée de Lie* de f suivant X et notée, $L_X f$, la multiplicité même des notations pour désigner cette notion est le garant de son importance. En coordonnées locales (q_1, \dots, q_n) , le champ X s'écrit $X = \sum_{i=1}^n X_i \frac{\partial}{\partial q_i}$ et

$$\partial_X f = df \cdot X = L_X f = \sum_{i=1}^n X_i \frac{\partial f}{\partial q_i}.$$

Un tel opérateur est une *dérivation* de l'algèbre $C^\infty(\Omega, \mathbb{R})$. Autrement dit, $\partial_X f$ est $C^\infty(\Omega, \mathbb{R})$ -linéaire en X , \mathbb{R} -linéaire en f , et vérifie

$$\partial_X (fg) = (\partial_X f)g + (\partial_X g)f.$$

De plus, il est *local* : la valeur de $\partial_X f$ en un point q ne dépend que des valeurs de f sur un voisinage arbitrairement petit de q . Réciproquement, on montre que toute dérivation locale de l'algèbre $C^\infty(\Omega, \mathbb{R})$ est définie par un champ de vecteurs. Ceci donne un sens à la

Définition. Le *crochet* de deux champs de vecteurs X, Y sur M est l'unique champ de vecteurs $[X, Y]$ sur M tel que, quelle que soit la fonction $f : M \rightarrow \mathbb{R}$ de classe C^∞ , on ait

$$\partial_{[X, Y]} f = \partial_X (\partial_Y f) - \partial_Y (\partial_X f).$$

En coordonnées locales, si $X = \sum_{i=1}^n X_i \frac{\partial}{\partial q_i}$ et $Y = \sum_{i=1}^n Y_i \frac{\partial}{\partial q_i}$, on trouve que

$$[X, Y] = \sum_{i=1}^n \left(\sum_{j=1}^n \left(X_j \frac{\partial Y_i}{\partial q_j} - Y_j \frac{\partial X_i}{\partial q_j} \right) \right) \frac{\partial}{\partial q_i}.$$

Le crochet est une obstruction à la commutation des flots φ_t et ψ_s engendrés par les deux champs : $[X, Y] = 0$ si et seulement si $\varphi_s \circ \psi_t = \psi_t \circ \varphi_s$. En effet, un calcul direct en coordonnées locales montre que, si s et t sont petits et si Φ_τ désigne le flot de $[X, Y]$,

$$\psi_{-s} \circ \varphi_{-t} \circ \psi_s \circ \varphi_t(x) = \Phi_{-ts}(x) + \text{termes d'ordre supérieurs.}$$

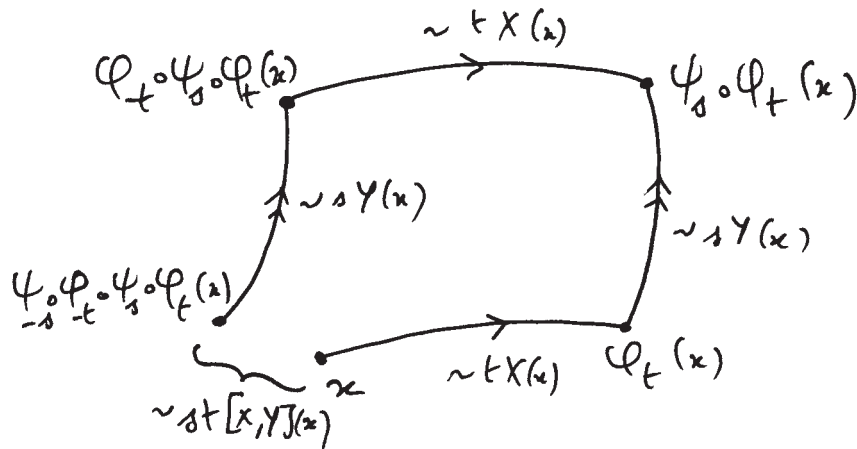


Figure 27 : crochet et non-commutation

L'exemple le plus simple d'un ensemble de champs de vecteurs qui commutent deux à deux est celui des $X_i = \frac{\partial}{\partial q_i}$, $i = 1, \dots, k$, dans \mathbb{R}^n (coordonnées q_1, \dots, q_n). Il est en fait paradigmatique : on montre qu'étant donné k champs de vecteurs X_1, \dots, X_k sur une variété M dont les crochets deux à deux sont tous nuls, il existe, au voisinage de tout point de M en lequel les X_i sont linéairement indépendants, des coordonnées locales q_1, \dots, q_n dans lesquelles les X_i deviennent les champs constants $\frac{\partial}{\partial q_i}$. On en déduit que, sous ces hypothèses, le champ des k -plans tangents à M engendrés en chaque point par les X_i est *intégrable* : il existe un *feuilletage* local de M par des sous-variétés de dimension k tangentes en chaque point au k -plan correspondant. Une conséquence de ceci est le

Théorème de Frobenius. Un champ de k -plans tangents à une variété M est intégrable si et seulement si le crochet de deux champs de vecteurs tangents à ce champ est également tangent à ce champ.

La démonstration est un simple choix de base : l'hypothèse permet en effet de choisir k champs de vecteurs qui engendrent en chaque point le k -plan donné et vérifient l'hypothèse du Théorème précédent.

5.2 La courbure de Riemann

Le théorème de Frobenius montre que la non-intégrabilité d'une connexion sur un fibré vectoriel est complètement caractérisée par la projection

$$[\tilde{X}, \tilde{Y}] - [\tilde{X}, Y]$$

sur l'espace vertical (i.e. l'espace tangent à la fibre, que l'on peut identifier à la fibre elle-même par translation) du crochet des relèvements horizontaux \tilde{X} et \tilde{Y} de deux champs de vecteurs X et Y dans la base M .

Dans la pratique, on peut choisir des champs X et Y qui commutent, ce qui implique que le crochet $[\tilde{X}, \tilde{Y}]$ est vertical et qu'aucune projection n'est plus nécessaire. De plus, on peut montrer que la valeur de cette projection en un point ζ de $E_m = \pi^{-1}(m)$ ne dépend de X et Y que par l'intermédiaire de leurs valeurs $\xi = X(m)$ et $\eta = Y(m)$ au point m , à l'exclusion de toute dérivée. Cela signifie que ce que l'on définit ci-dessous est un *tenseur* : après identification à E_m de l'espace tangent $T_\zeta E_m$ en l'un quelconque de ses points ζ , on associe à $\xi, \eta \in T_m M$ une application linéaire $R(\xi, \eta) : E_m \rightarrow E_m$:

$$\zeta \mapsto R(\xi, \eta)\zeta.$$

Tout ce ci vaut dans un fibré vectoriel sur M muni d'une connexion. Le tenseur (trois fois covariant et une fois contravariant) R qui à un triplet ξ, η, ζ de vecteurs tangents à M en un même point fait correspondre le vecteur $R(\xi, \eta)\zeta$ tangent en ce point.

Dans le cas riemannien où la connexion est celle de Levi-Civita, on définit souvent la *courbure de Riemann* comme le tenseur (quatre fois covariant) qui, à quatre vecteurs ξ, η, ζ, τ tangents à M au même point m associe

$$R(\xi, \eta, \zeta, \tau) = \langle R(\xi, \eta, \zeta), \tau \rangle_{g(m)} \in \mathbb{R}.$$

On trouve dans le paragraphe 9 de [Mi] une belle description des symétries de ce "little monster of (multilinear) algebra".

Le cas d'une surface riemannienne : variations d'angle et courbure de Gauss. La courbure $R(\xi, \eta, \zeta, \tau)$ acquiert une interprétation simple lorsque M est une surface riemannienne.

Puisque le transport parallèle préserve le produit scalaire, il suffit de considérer la restriction de la connexion au fibré tangent unitaire. Etant donné $m \in M$ et trois vecteurs tangents $\xi, \eta, \zeta \in T_m M$ qui vérifient

$$\|\xi\|_{g(m)} = 1, \|\eta\|_{g(m)} = 1, \|\zeta\|_{g(m)} = 1, \quad \langle \xi, \eta \rangle_{g(m)} = 0,$$

choisissons des coordonnées locales (q_1, q_2) centrées sur m , telles que les deux champs de vecteurs $X = \frac{\partial}{\partial q_1}$ et $Y = \frac{\partial}{\partial q_2}$ vérifient $X(m) = \xi$ et $Y(m) = \eta$. Puisque $[X, Y] = 0$, on a par définition $R(\xi, \eta)\zeta = [\tilde{X}, \tilde{Y}](\zeta)$, où \tilde{X} et \tilde{Y} sont les relèvements horizontaux de X et Y . Pour calculer ce crochet, relevons horizontalement le vecteur ζ le long du bord du "carré"

$$\mathcal{A}_\epsilon = \{(q_1, q_2), 0 \leq q_1 \leq \epsilon, 0 \leq q_2 \leq \epsilon\}.$$

Cela revient à transporter parallèlement le vecteur ζ le long du bord de \mathcal{A}_ϵ . Le résultat de ce transport parallèle est un vecteur $\hat{\zeta}$ de norme 1 qui se déduit de ζ par une rotation d'un angle (riemannien) $\Delta\theta_\epsilon$ dans le plan tangent $T_m M$. On a donc

$$\hat{\zeta} - \zeta = \epsilon^2[\tilde{X}, \tilde{Y}](\zeta) + O(\epsilon^3) = \Delta\theta_\epsilon + o(\Delta\theta_\epsilon).$$

La préservation de la métrique par transport parallèle a pour conséquence que l'angle $\Delta\theta_\epsilon$ ne dépend pas du choix de ζ . Choisissons $\zeta = \xi$ et divisons par l'aire (riemannienne) $\epsilon^2 + O(\epsilon^3)$ du carré, on obtient que

$$R(\xi, \eta, \xi, \eta) = \lim_{\epsilon \rightarrow 0} \frac{\Delta\theta_\epsilon}{\text{aire de } \mathcal{A}_\epsilon}.$$

Avec un peu d'algèbre linéaire on en déduit que, quels que soient les vecteurs linéairement indépendants $\xi, \eta \in T_m M$,

$$K(m) := \lim_{\epsilon \rightarrow 0} \frac{\Delta\theta_\epsilon}{\text{aire de } \mathcal{A}_\epsilon} = \frac{R(\xi, \eta, \xi, \eta)}{\|\xi \wedge \eta\|^2}$$

Cette quantité est appelée *courbure de Gauss* de la surface M .

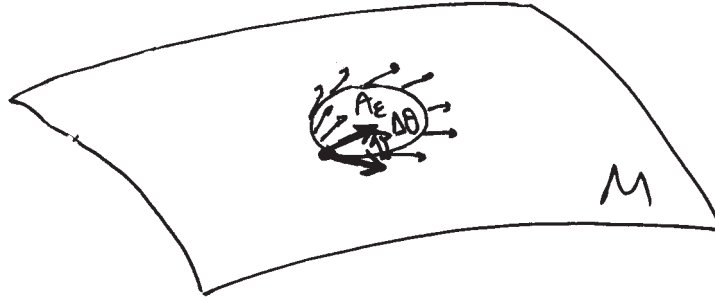


Figure 28 : courbure de Gauss (intrinsèque)

Dans une carte orthogonale $\varphi : \Omega \ni q = (q_1, q_2) \rightarrow \varphi(q) \in M$, la courbure $K = K \circ \varphi(q)$ est donnée par la formule

$$K = -\frac{1}{\sqrt{g_{11}g_{22}}} \left\{ \frac{\partial}{\partial q_1} \left(\frac{1}{2\sqrt{g_{11}g_{22}}} \frac{\partial g_{22}}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{1}{2\sqrt{g_{11}g_{22}}} \frac{\partial g_{11}}{\partial q_2} \right) \right\}.$$

Le signe de la courbure en un point d'une surface se lit sur la somme d'un petit triangle voisin de ce point : suivant que la courbure est positive ou négative, cette somme est inférieure ou supérieure à 2π : dans le premier cas, on déchire pour aplatisir, dans le deuxième on fait des plis.

5.3 Le theorema egregium de Gauss

Il est surprenant que la quantité ci-dessus, attachée à la seule métrique de la surface M coïncide, lorsque cette métrique est induite de la métrique euclidienne par un plongement de M dans \mathbb{R}^3 , avec le produit des *rayons de courbure principaux*, quantité éminemment extrinsèque déjà introduite par Euler mais dont il revient à Gauss d'avoir montré le caractère intrinsèque.

Définition extrinsèque de la courbure. La courbure de Gauss extrinsèque en un point m de M est définie par

$$K_e(m) = \lim_{\mathcal{A} \rightarrow m} \epsilon(m) \frac{\text{aire } \mathcal{N}(\mathcal{A})}{\text{aire } \mathcal{A}},$$

où $\mathcal{N} : M \rightarrow S^2$ est l'application de Gauss qui, à un point de M , fait correspondre la normale en ce point et l'aire sur S^2 est définie par la métrique induite par R^3 et $\epsilon(m)$ vaut $+1$ ou -1 suivant que \mathcal{N} préserve ou non l'orientation au voisinage de m . [si $K_e(m) = 0$ il faut prendre l'aire algébrique de $\mathcal{N}(\mathcal{A})$ qui, lorsque \mathcal{N} n'est pas trop dégénérée, s'interprète géométriquement comme la somme avec signes ϵ des aires des images par \mathcal{N} des parties de \mathcal{A} où le déterminant de $d\mathcal{N}$ ne s'annule pas].

Gauss appelle *courbure totale* de \mathcal{A} l'aire algébrique de $\mathcal{N}(\mathcal{A})$.

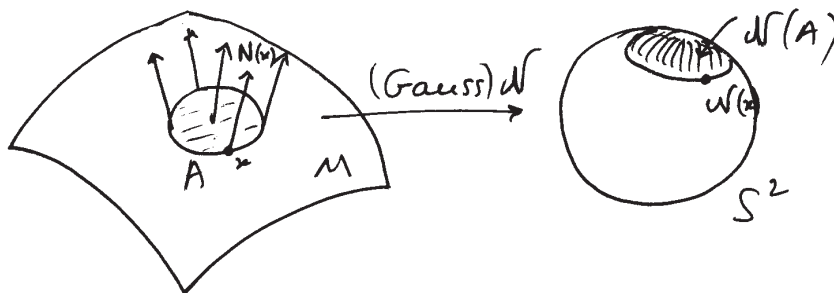


Figure 29 : courbure de Gauss (extrinsèque)

Le *theorema egregium*, textuellement le “théorème merveilleux”, affirme que les deux courbure K et K_e que nous venons de définir sont les mêmes pour une surface M de \mathbb{R}^3 . La démonstration analytique n'est pas difficile mais elle reste mystérieuse. Le théorème est profond, même dans le cas des polyèdres, où il relie les angles des faces et l'angle solide des normales extérieures (voir [HCV] et [Ba]). Je me contenterai d'en donner une illustration que je tiens de Mark Levi (voir [L]) : considérons le cas où M est un cône. Les courbures intrinsèque et extrinsèque sont des *masses de Dirac* concentrées au sommet. Ceci signifie simplement que quel que soit le domaine \mathcal{A} de M , ou bien \mathcal{A} ne contient pas le sommet et la courbure totale de \mathcal{A} calculée avec l'une ou l'autre des courbures vaut 0, ou bien \mathcal{A} contient le sommet et chacune des courbures totales est indépendante de \mathcal{A} . Coupant le cône par une sphère de rayon 1

centrée en son sommet, on obtient un cercle. Développant le cône sur le plan, on voit que la variation d'angle $\Delta\theta$ par déplacement parallèle le long de ce cercle vaut $\Delta\theta = 2\pi - L$. Soit A l'angle solide défini par le cône *dual* M^* , défini comme enveloppe des plans orthogonaux au cône M . A n'est autre que la surface découpée sur la sphère par l'un des demi-cônes définis par M^* (figure 27). Le théorème de Gauss affirme l'égalité

$$2\pi - L = A.$$

Remarque. Même dans le cas polyédral ce théorème reste non évident : un polyèdre peut avoir beaucoup de déformations isométriques provenant de pivotelements des faces autour des arêtes (penser à une pyramide de base carrée que l'on peut aplatir en passant par des pyramides à base un losange de plus en plus plat), ce qui montre que l'angle solide *intérieur*, dont la définition est claire dans le cas convexe, n'est pas intrinsèquement défini ; il peut varier et même s'annuler sans que les angles des faces varient !

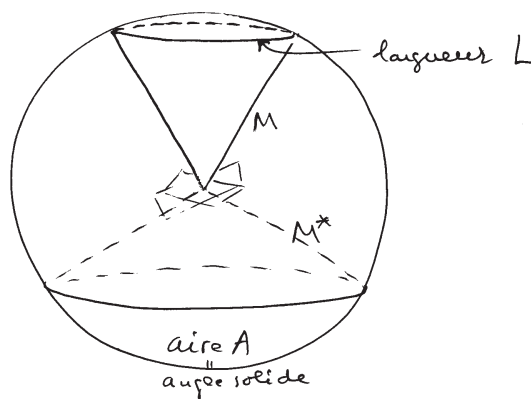


Figure 30 : cônes duaux

Quelques dates.

- 1827 Carl-Friedrich Gauss, Disquisitiones generales circa superficies curvas
- 1854 Bernhard Riemann, Über die Hypothesen, welche der Geometrie zu Grunde liegen
- 1869 Elwin Christoffel
- 1895 Gregorio Ricci Curbastro, Sulla teoria intrinseca delle superficie ed in specie di quelle di 2^o grado
- 1900 Gregorio Ricci Curbastro & Tullio Levi-Civita, Méthodes de Calcul différentiel absolu et leurs applications
- 1917 Tullio Levi-Civita, Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana, Rendiconti del Circolo Matematico di Palermo, vol. XLII, p. 173-215
- 1922 Elie Cartan, Sur une généralisation de la notion de courbure de Riemann et les espaces à torsion, C.R.Acad. Sc. t. 174, p. 734
- 1950 Charles Ehresman, Les connexions infinitésimales dans un espace fibré différentiable, Colloque de topologie (espaces fibrés), Bruxelles
- 1951 Jean-Louis Koszul

6 De la courbure à la forme : le théorème de Gauss-Bonnet

6.1 Triangles géodésiques et formule locale

Généralisant les formules classiques de trigonométrie sphérique ou hyperbolique, la formule de Gauss-Bonnet est le parangon des théorèmes faisant le lien entre propriétés globales et propriétés locales. Pour l'établir, décomposons une surface M en "triangles géodésiques" suffisamment petits pour être chacun contenu dans un domaine de carte; "géodésiques" veut dire que les côtés sont des segments de géodésiques. Remarquons que la courbure totale $\int_T K(m)d\sigma(m)$ d'un tel triangle T n'est autre que la variation d'angle $\Delta\theta_\epsilon$ définie par le transport parallèle le long de son bord orienté. Puisque ce bord est formé de segments de géodésiques, le transport parallèle d'un vecteur le long d'un côté conserve l'angle qu'il fait avec ce côté; on constate alors que $\Delta\theta_\epsilon + \sum_{i=1}^3 \beta_i = 2\pi$, où les angles extérieurs $\beta_i = \pi - \alpha_i$ du triangle sont définis sur la figure 31.

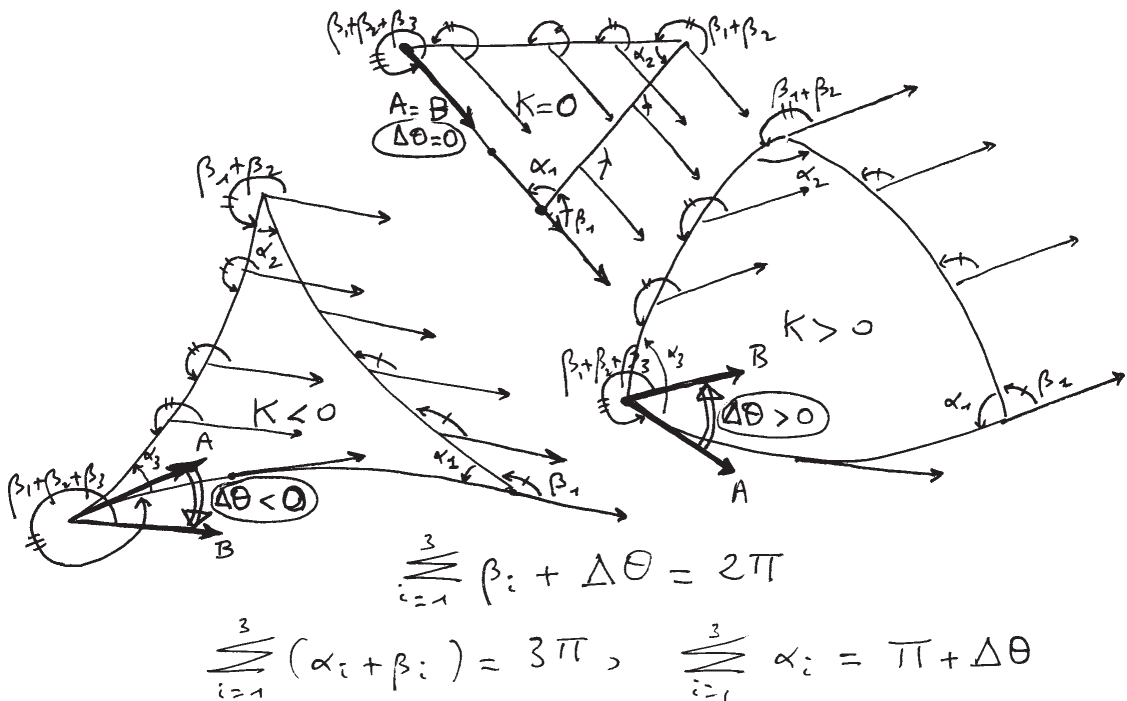


Figure 31 : triangles géodésiques

Autrement dit, pour tout petit triangle à bords géodésiques, on a la formule de Gauss

$$\sum_{i=1}^3 \alpha_i = \pi + \int_T K(m)d\sigma(m).$$

6.2 Caractéristique d'Euler-Poincaré et formule globale

Pour une subdivision de la surface compacte sans bord M en petits triangles géodésiques, notons S, A, F respectivement le nombre de sommets, le nombre de côtés (arêtes), le nombre de triangles (faces), et

$$\mathcal{X}(M) = S - A + F$$

la *caractéristique d'Euler* de M . Ajoutant les courbures totales de chaque triangle, on obtient pour la courbure totale de M la somme

$$(2\pi)F - \sum \beta_i = (2\pi)F - \sum (\pi - \alpha_i) = (2\pi)F - ((2A)\pi - S(2\pi)),$$

car chaque arête appartient à deux triangles et la somme des angles en un sommet vaut 2π . On vient de démontrer la *Formule de Gauss-Bonnet* :

$$\int_M K(m)d\sigma(m) = 2\pi\mathcal{X}(M).$$

Le lecteur devra se convaincre de l'invariance par subdivision de la *caractéristique d'Euler* (voir [H]). En fait, celle-ci ne dépend que de la topologie de M . La bonne démonstration passe par la théorie de l'*homologie*; on la trouvera dans tout livre élémentaire de *Topologie algébrique* ou dans le premier volume du traité de *Spivak*. Pour les surfaces *compactes orientables sans bord*, la caractéristique d'Euler est un invariant complet : elle suffit à les distinguer. Une telle surface M est une "sphère avec g anses" (g est le *genre* de M) et sa caractéristique vaut $\mathcal{X}(M) = 2 - 2g$. Si $g > 0$, on parle encore d'un "tore à g trous" ou de la "somme connexe" de g tores (figure 32). Quant aux surfaces compactes sans bord *non orientables*, qui ne peuvent se représenter dans \mathbb{R}^3 sans auto-intersections, elles s'obtiennent en attachant à la sphère q rubans de Moebius ($q = 1$ pour le plan projectif, $q = 2$ pour la bouteille de Klein, etc...) et sont également distinguées par leur caractéristique d'Euler-Poincaré $\mathcal{X}(M) = 2 - q$.

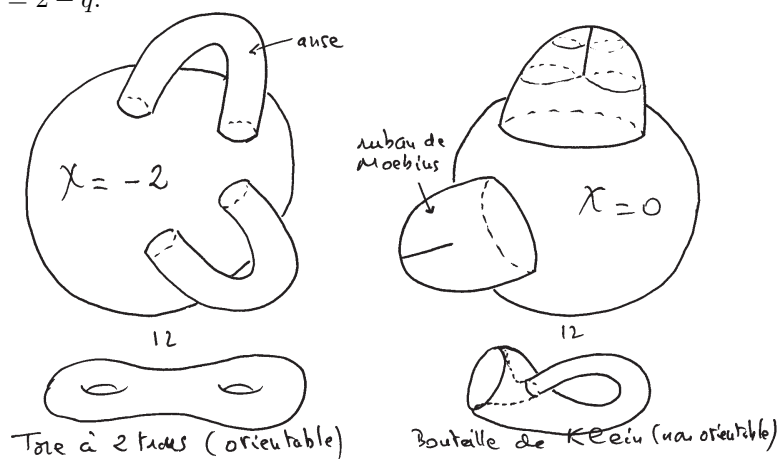


Figure 32 : surfaces

Considérons le cas où M est une surface compacte sans bord de \mathbb{R}^3 (donc orientable) munie de la métrique induite par la métrique euclidienne. Si on remplace la courbure totale par son interprétation extrinsèque, à savoir l'aire algébrique de l'image de M par l'application de Gauss, et si on se rappelle que l'aire de la sphère unité de \mathbb{R}^3 est 4π , on constate que la formule de Gauss-Bonnet équivaut à l'affirmation que l'image d'une surface compacte sans bord de \mathbb{R}^3 de genre g par son application de Gauss "enveloppe $(1-g)$ fois" la sphère S^2 .

Remarque. L'identité de Gauss-Bonnet interdit qu'on puisse "écraser" une famille de sphères riemanniennes S^2 sur un point tout en maintenant la courbure de Gauss bornée inférieurement. C'est au contraire très facile avec une famille de tores plats : il n'y a pas de contradiction puisque la caractéristique d'un tore est nulle.

Lecture. On trouvera dans [L] une autre interprétation de la formule $2\pi = L + A$ des cônes duaux dans laquelle L représente l'intégrale sur le cercle qui borde le cône de la courbure géodésique k_g de celui-ci (cette courbure s'annule identiquement si la courbe est une géodésique). La formule devient alors un cas particulier ($M = S^2$, $\mathcal{D} = A$) de la formule générale de Gauss-Bonnet pour un domaine à bord \mathcal{D} d'une surface riemannienne orientée M :

$$\int_{\mathcal{D}} K(m) d\sigma(m) + \int_{\partial\mathcal{D}} k_g(m) ds(m) + \sum_i \beta_i = 2\pi\chi(\mathcal{D}),$$

où $k_g(m)$ est la courbure géodésique du bord \mathcal{D} en l'un de ses points m , et ds l'élément de longueur riemannien sur ce bord. En particulier, la caractéristique d'Euler-Poincaré d'un anneau étant nulle, on en déduit que deux géodésiques périodiques plongées dans une sphère de courbure positive s'intersectent nécessairement.

Références

- [A] V.I. Arnold, Méthodes mathématiques pour la mécanique classique, Editions Mir 1976
- [Ba] Banchoff, Critical points and curvature for embedded polyedral surfaces, American Mathematical Monthly, may 1970
- [BG] M. Berger & B. Gostiaux, Géométrie différentielle : variétés, courbes et surfaces, P.U.F. 1987
- [C] A. Chenciner, Géométrie et mécanique, livre en préparation
- [D] J. Dieudonné, Eléments d'Analyse, chapitre XVII (tome troisième)
- [G] F. Gramain, Topologie des surfaces, P.U.F. 1971
- [H] H. Hopf, Differential geometry in the large, Lecture Notes in Mathematics 1000, Springer-Verlag 1983
- [HCV] Hilbert & Cohn-Vossen, Geometry and the imagination, chapitre IV, paragraphe 29, pages 194,195

- [L] M. Levi, Lectures on Geometrical Methods in Mechanics, in “Classical and Celestial Mechanics, the Recife lectures”, H. Cabral and F. Diacu ed., Princeton University Press 2002
- [LS] D. Lehmann & C. Sacré, Géométrie et topologie des surfaces, PUF 1982
- [M] J. Milnor, Morse theory, Princeton University press, 1963
- [MTW] C.W. Misner, K.S. Thorne & J.A. Wheeler, Gravitation, Freeman 1970
- [Mi] E. Miller, Metric combinatorics of convex polyhedra : cut loci and nonoverlapping unfoldings, preprint 2003
- [P] F. Pham, Géométrie et Calcul différentiel sur les variétés, InterEditions 1992
- [P1] Pham, Fonctions d’une ou deux variables, Dunod 2003
- [S] M. Spivak, A comprehensive introduction to differential geometry, volumes 1 and 2, Publish or Perish 1970
- [S1] M. Spivak, Calculus on manifolds, Benjamin 1965