Université de Strasbourg
École doctorale mathématiques, sciences
de l'information et de l'ingénieur - ED 269

Université Abou Bakr Belkaid - Tlemcen
Faculté de Sciences
Département de Mathématiques

Thèse de doctorat en cotutelle

Spécialité : Mathématiques

Option : Géométrie différentielle

Présentée par : Mohammed Gorine

Intitulée :

Influence de la courbure sur la taille du barycentre convexe dans les variétés différentiables

---

Soutenue le : 24 Janvier 2015

Jury :

A. LANSARI	Prof à l'université de Tlemcen, ALGERIE	Président
M. BENALILI	Prof à l'université de Tlemcen, ALGERIE	Rapporteur
M. ARNAUDON	Prof à l'université de Bordeaux, FRANCE	Rapporteur
J. FRANCHI	Prof à l'université de Strasbourg, FRANCE	Examinateur
M. BELKHELFA	Prof à l'université de Mascara, ALGERIE	Directeur
M. EMERY	Directeur de recherche IRMA Strasbourg, FRANCE	Directeur

Table des matières

Remerciements

Ce travail a été réalisé en cotutelle aux seins de l'institut de recherche mathématique avancée (IRMA) à Strasbourg et le département de mathématique à l'université de Tlemcen.

Je remercie trés cordialement les responsables de cette coopération universitaire Franco-Algerienne.

Je tiens à exprimer ma profonde gratitude et ma scincère reconnaissance a mes directeurs de thèse qui m'ont encadré durant ces années. Monsieur Mohammed Belkhelfa, directeur du laboratoire LPQ3M à l'université de Mascara, pour ses précieux conseils, son aide inestimable et son optimisme contagieux. Monsieur Michel Emery, directeur de recherche à l'IRMA à l'université de Strasbourg pour m'avoir fait partagé ses nombreuses connaissances et qui m'a souvent donné le courage d'avancer dans mes recherches, notamment en me motivant lorsque j'en éprouvais le besoin et sans qui cette thèse n'aurait jamais pu être menée a bien.

Merci aux professeurs, Monsieur Mohammed Benallili et Monsieur Marc Arnaudon qui ont accepté de rapporter cette thèse, je les remercie du temps qu'ils y ont consacré.

Je remercie également les professeurs, Monsieur A.Lansari et Monsieur Jacques Franchi pour avoir bien voulu faire partie du jury.

Je tiens a souligner que ces travaux n'auraient pu être réalisés sans l'aide financière du PHC Tassili 08 MDU737, qui nous a permis de fructueux contacts avec les probabilistes de Strasbourg et de Poitiers, a qui je suis reconnaissant.

Mes remerciements vont aux dirigeants de la faculté des sciences et de la technologie à Mascara et de la faculté des sciences de l'université de Tlemcen qui ont contribué à la réussite de cette collaboration interuniversitaire.

Enfin je ne saurais terminer cette liste sans adresser un remerciement particulier a ma petite famille qui m'a soutenu dans l'ombre sans qui ce travail n'aurait jamais vu le jour. Je leur dédie ce travail en témoignage de ma profonde affection pour toute la patience et les sacrifices qu'ils ont convertis pour moi et dont je ne serais à jamais redevable, et d'avoir porté ce travail à terme représente pour moi aujourd'hui la plus belle des récomponses.

Introduction

0.1 Eléments de synthèse sur les travaux antérieurs

La notion d'espérance pour une variable aléatoire à valeurs dans un espace métrique M trouve ses premières formulations chez Fréchet [26] et Doss [18]. Les propriétés de cette espérance sont développées par Benès [7], Doss [19] et Herer [38] et [41]. Ce dernier introduit dans [40] et [42] une autre définition et s'intéresse aux espaces à courbure négative. Plus récemment le cas M variété a été abordé par Arnaudon [3] et Émery et Mokobodzki [23]. Ces différentes approches ont en commun que l'espérance d'une varible aléatoire est un sous-ensemble fermé en général non réduit à un point.
Dans une variété, les martingales continues ont été définies par Duncan [20], Meyer [50] et Darling [16]. Meyer et Darling ont mis en évidence la structure géométrique qui permet cette définition : ce sont les connexions. Il est plus difficile de définir les martingales discontinues, ou même simplement les martingales à temps discret, parce qu'en l'absence de structure linéaire (ou affine) il y a au moins autant de définitions possibles que pour l'espérance. Si la variété est riemannienne, on peut chercher à construire le barycentre en minimisant la moyenne du carré de la distance ; ce procédé, déja employé par Cartan [11] pour trouver le point fixe d'un groupe d'isométries, a été efficacement utilisé par Kendall [45] et Picard [54] pour la construction approchée d'une martingale continue par discrétisation du temps.
Cette définition se généralise sans peine aux variétés munies d'une connexion (c'est le "barycentre exponentiel" que nous verrons plus loin), mais souffre d'un grave défaut : Les barycentres ainsi construits n'ont pas la propriété d'associativité.
Il est cependant un cas où la définition du barycentre d'une probabilité ne pose pas de problème. Dans une variété munie d'une connexion et telle que deux points quelconques sont joints par une géodésique et une seule, on peut définir de façon évidente le barycentre d'une mesure μ portée par au plus deux points : si $\mu = (1-t)\varepsilon_x + t\varepsilon_y$, son barycentre est bien sûr $\gamma(t)$ où γ est

la géodésique telle que $\gamma(0) = x$ et $\gamma(1) = y$.
Pour des mesures plus générales, nous suivrons la voie empruntée par Émery et Mokobodzki dans [23] : sacrifier l'unicité pour sauvegarder l'associativité, en définissant le barycentre comme étant non pas un point, mais tout un ensemble de points. Leur définition présente cependant un inconvénient : elle n'est pas locale, au sens où modifier la variété très loin du support de la mesure peut affecter le barycentre. La définition que nous utilisons (donnée au début du chapitre 3) est une variante de celle de [23], que nous empruntons à Arnaudon [4], et qui corrige ce défaut.
La non-unicité de l'espérance d'une variable aléatoire dans un espace métrique n'est pas une nouvelle idée ; voir Fréchet [25], Doss [18], Herer [38], [39] et [40].
La première définition formelle des moyennes pour des mesures de probabilité sur les variétés riemanniennes a été donnée par H. Karcher ; il a montré, à l'aide d'une estimation des champs de Jacobi, que l'energie locale fonctionnelle

$$F_\mu : \overline{B}(a, \rho) \to \mathbb{R}_+,\ x \mapsto \int_M d^2(x, p)\mu(dp)$$

est strictement convexe. Dans cette formule, d est la distance riemannienne sur la variété, $B(a, \rho) = \{x \in M : d(x, a) < \rho\}$ la boule géodésique de centre a et de rayon ρ vérifiant $\rho < min\{\frac{\pi}{4\sqrt{\Delta}}, \frac{inj}{2}\}$, où Δ est un majorant des courbures sectionnelles de $B(a, \rho)$ et inj son rayon d'injectivité. (Si $\Delta \leq 0$, alors $\frac{\pi}{4\sqrt{\Delta}}$ sera interprété comme $+\infty$.)
La convexité stricte de cette fonctionnelle F_μ implique l'existence d'un minimum unique $b(\mu)$ que Karcher a appellé centre riemannien de la mesure de probabilité μ. En outre $b(\mu)$ est également l'unique solution x de l'équation suivante :

$$\int_M exp_x^{-1} p\mu(dp) = 0_x\ , \qquad x \in \overline{B}(a, \rho).$$

Dès lors les moyennes locales des mesures de probabilité sur les variétés riemanniennes sont également appelées moyennes de Karcher, alors que les moyennes globales sont souvent appelées moyennes de Fréchet.
Ce dernier a défini dans son célèbre article [26] une p-moyenne d'une variable aléatoire X comme étant le point qui minimise l'éloignement de ce point à la puissance p de X ; ceci permet de définir différentes valeurs typiques parmi lesquelles on retrouve deux cas importants : $p = 1$ et $p = 2$, correspondant respectivement aux notions de médiane et moyenne.
W.S. Kendall a prouvé dans [44] l'unicité des moyennes locales avec la condition $\rho < min\{\frac{\pi}{2\sqrt{\Delta}}, inj(a)\}$, où $inj(a)$ est le rayon d'injectivité de a en prenant en considération le fait que dans $B(a, \rho)$, chaque paire de points peut

être jointe par une et une seule géodésique incluse dans $B(a,\rho)$ et que la distance $d(x,p)$ est la longueur géodésique de x à p (elle peut être plus grande que la distance induite par la métrique riemannienne de M).
En particulier, la condition $\rho < \frac{1}{2} min\{\frac{\pi}{\sqrt{\Delta}}, inj\}$, est suffisante pour garantir l'unicité de la moyenne de Karcher de μ.
La méthode de Kendall pour prouver l'unicité de la moyenne de Karcher est ingénieuse est intéressante ; son idée de base est une relation entre la moyenne de Karcher et les fonctions convexes sur $B(a,\rho)$: Pour chaque moyenne de Karcher x de μ, on a $\varphi(x) \leq \int_M \varphi(p)\mu(dp)$ pour chaque fonction φ convexe sur $B(a,\rho)$.
Compte tenu de ça et du fait que le produit de deux moyennes de Karcher de μ est une moyenne de Karcher du produit de mesure $\mu \otimes \mu$, la preuve de l'unicité est réduite à la construction d'une fonction convexe sur $B(a,\rho) \times B(a,\rho)$ qui s'annule exactement sur la diagonale.
Kendall a donné dans [44] un exemple de telles fonctions, des discussions détaillées sur la construction sont dans [45].
M. Émery et G. Mokobodzki ont défini dans [23] les barycentres exponentiels et les barycentres convexes pour les mesures sur les variétés affines comme généralisation de la moyenne de Karcher. Ils ont également montré qu'un point x est un barycentre convexe d'une probabilité μ si et seulement s'il existe une martingale continue à partir de x avec la loi terminale μ.
L'unicité des barycentres exponentiels a été généralisée par M. Arnaudon et X.M. Li dans [5] aux mesures de probabilité sur les variétés affines convexes à géométrie convexe semilocale, ils ont aussi étudié le comportement des barycentres exponentiels lorsque des mesures sont transportées par des flots stochastiques.
M. Arnaudon a donné dans [3] une majoration des barycentres convexes. J. Picard a aussi donné dans [53] une notion généralisée de barycentre pour étudier les applications harmoniques entre variétés riemanniennes par des méthodes probabilistes.
Si μ est une mesure discrète portée par un nombre fini de points dans l'hémisphère supérieur fermé, S.R Buss et J.P Fillmore ont montré dans [10] que si le support de μ n'est pas totalement contenu dans l'équateur alors la moyenne de Karcher de μ est unique, elle se trouve dans l'hémisphère ouvert et est égale à la moyenne de Fréchet. Inspiré par la méthode de Buss et Fillmore, B. Afsari a montré dans [1] que si le majorant des courbures sectionnelles et le rayon d'injectivité sont remplacés par ceux de la grande boule $B(a, 2\rho)$, alors toutes les p-moyennes de Fréchet de μ sont à l'intérieur de $B(a,\rho)$. La preuve de Afsari repose sur un argument de comparaison qui

prend en considération, non seulement la comparaison des triangles avec ceux des espaces modèles, mais toute la géométrie de la variété. L'idée de Afsari conduit à une nouvelle démonstration géométrique du résultat de Kendall sur l'unicité des moyennes de Karcher.

0.2 Objectif et plan de travail

On considère une variété différentiable M, sans bord, de classe C^∞, munie d'une connexion affine, et telle que deux points quelconques de M sont toujours reliés par une géodésique et une seule, qui en outre dépend de façon C^∞ des deux points.
Etant donné une mesure de probabilité μ sur M, à support compact suffisamment petit, notre but est de définir le barycentre $b(\mu)$ comme un ensemble de points, puis d'étudier l'influence de la structure de M (courbure ou tous autres éléments géométriquement significatifs) sur la taille de $b(\mu)$.
La définition de $b(\mu)$ est évidente, par interpolation géodésique, lorsque μ est portée par un ou deux points de M ; mais pour des mesures plus générales, nous travaillerons avec le barycentre convexe $b(\mu)$, qui n'est pas unique (c'est donc un ensemble de points de M), mais qui possède une propriété d'associativité ; notre définition sera une variante du barycentre convexe introduit dans [23].
Lorsque μ est à support fini, des points de $b(\mu)$ peuvent être obtenus en remplaçant une partie de la mesure portée par deux points par son barycentre géodésique, et en itérant ce procédé jusqu'à ce qu'il ne reste plus qu'un seul point ; chaque point de $b(\mu)$ ainsi obtenu est associé à la racine d'un arbre dyadique dont les sommets sont des points de M pondérés par des masses. Si μ est une probabilité plus générale, nous montrerons que tout point de $b(\mu)$ peut être obtenu comme une limite de tels points de $b(\mu_n)$ pour des mesures μ_n à supports finis qui convergent convenablement vers μ.
Inversement, des approximations de $b(\mu)$ par l'extérieur (nous les appellerons majorations) peuvent être obtenues en construisant des fonctions convexes ; ces approximations seront d'autant plus précises que les fonctions convexes utilisées sont plus près d'être affines.

Ce mémoire est divisé en six chapitres :

Le présent chapitre d'introduction présente des travaux antérieurs concernant ce sujet, et contient aussi un résumé des résultats que nous avons obtenus.

Le chapitre suivant (chapitre 1) est une étude détaillée des comportements des géodésiques, notamment dans les espaces homogènes de dimension 3 dont le groupe des isométries est de dimension 4. L'importance de cette étude réside dans le fait que ces courbes jouent un rôle central dans les notions de convexité d'ensembles ou de fonctions sur la variété.
Comme ce sujet est à l'interface entre la théorie de probabilité et la géométrie différentielle, des notions géométriques élémentaires sont rappelées dans cette partie à l'intention des lecteurs probabilistes.

Le chapitre 2 est consacrée à la convexité riemannienne. Elle est au cœur du sujet, puisque les barycentres convexes sont définis en utilisant les fonctions convexes sur les variétés d'une part, et que les encadrements optimaux des barycentres dépendent des comportements de ces fonctions d'autre part ; l'étude du défaut d'affinité de ces fonctions sera essentiel pour chercher des majorations aussi fines que possible du barycentre convexe.

Les propriétés du barycentre font l'objet du chapitre 3. La plupart d'entre elles ont été obtenues par M. Émery et G. Mokobodzki dans [23] et M. Arnaudon dans [3] ; nous les adaptons au contexte dans lequel nous nous plaçons ici, puisque le barycentre que nous utilisons n'est pas tout à fait le même que dans [23].
Nous nous intéressons aussi à l'enveloppe convexe inférieure d'une fonction, outil indispensable dans certaines démonstrations, et qui devrait également jouer un rôle dans la recherche de descriptions plus fines du barycentre.

Quant au chapitre 4, il présente les encadrements de barycentres que nous avons obtenus en utilisant l'associativité des barycentres (voir [23]) et des techniques de majoration du diamètre de $b(\mu)$ empruntées à [3] jointes à des estimations de hessiennes de fonctions liées à la distance riemannienne.

Enfin le chapitre 5, qui conclut ce travail, présente nos projets d'étude du bord du barycentre convexe en vue d'une description exacte, et non plus seulement approchée, de $b(\mu)$ dans certaines situations simples. Une famille de fonctions définies comme des enveloppes convexes inférieures y tient un rôle fondamental.

0.3 Résultats

Proposition 0.3.1. *Soient (M,g) une variété riemannienne de dimension 2, O un point de M, T_OM l'espace tangent en O, k la courbure de M en O. Pour toute constante réelle $C > \frac{2}{9}|k|$, et pour toute forme linéaire λ, de norme 1, définie sur T_OM, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ est convexe sur un voisinage de O, où $\overrightarrow{Ox} \in T_OM$ désigne la vitesse initiale de la géodésique allant de O à x, et $r = r(x) = \|\overrightarrow{Ox}\|$ la distance riemannienne de O à x.*

Corollaire 0.3.2. *Soient (M,g), O et k comme dans la proposition. Pour toute constante $C > \frac{2}{9}|k|$ il existe un voisinage V de O tel que pour toute forme linéaire λ sur T_0M de norme 1, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ soit convexe sur V.*

Théorème 0.3.3. *(M,g) est une variété riemanienne de dimension 2, O est un point de M et k désigne la courbure de M en O. Alors*

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{|k|}{18},$$

$E(D)$ désignant l'ensemble des mesures de probabilité portées par la boule géodésique $B(O, \frac{D}{2})$ et de barycentre exponentiel O, et $|b(\mu)|$ étant le diamètre du barycentre $b(\mu)$.

Ce théorème montre l'influence de la courbure de M sur la taille de $b(\mu)$ pour des mesures à support compact suffisamment petit.

Encadrement du diamètre du barycentre convexe de la loi uniforme portée par un petit cercle sur l'hémisphère

Dans le théorème 0.3.4 ci-dessous, la variété M est une boule géodésique ouverte de centre O et de rayon $R < \frac{\pi}{2}$ dans la sphère unité de dimension 2. Pour $r \in]0, R[$, on note μ_r la probabilité uniforme sur le cercle de centre O et de rayon r. L'ensemble $b(\mu_r)$ est une boule géodésique de centre O dont nous allons estimer le rayon.

Théorème 0.3.4. *Si c et C sont des réels positifs tels que $c < \frac{\sqrt{3}}{8\pi}$ et $C > \frac{2}{9}$, il existe $r_0 \in]0, R[$ tel que pour tout $r \in]0, r_0[$, on ait la double inclusion :*

$$B(O, cr^3) \subset b(\mu_r) \subset B(O, Cr^3),$$

où $B(O,a)$ est la boule géodésique de centre O et de rayon a.

Proposition 0.3.5. *Soient (M, g) une variété riemannienne de dimension 3, et O un point de M tel que toutes les courbures sectionnelles en O sont égales à une même valeur k. Pour tout $C > \frac{1+\sqrt{2}}{9}|k|$, il existe un voisinage de O sur lequel toutes les fonctions $f_{C,\lambda}$ sont convexes, où λ décrit la sphère unité de l'espace cotangent $T^*_O M$.*

Minoration du diamètre du barycentre convexe de la loi uniforme portée par une petite sphère S^{n-1} dans la n-sphère.

La variété M est maintenant de dimension n et à courbure constante 1 ; plus précisément, M est une boule géodésique régulière de rayon R et de centre O dans la n-sphère unité. Voici une minoration du rayon du barycentre convexe $b(\mu_r)$, où μ_r est la probabilité uniforme sur la sphère géodésique de petit rayon r et de centre O. Nous utiliserons la suite des nombres

$$I_k = \int_0^{\frac{\pi}{2}} (\cos\alpha)^k \, d\alpha = \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{k}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{k}{2} + 1\right)} .$$

Théorème 0.3.6. *Pour tout c tel que $0 < c < \frac{\sqrt{3}}{8\pi} I_{n-1}$ il existe $r_0 \in \,]0, R[$ tel que pour tout $r \in \,]0, r_0[$, on ait la minoration*

$$B(O, cr^3) \subset b(\mu_r).$$

Théorème 0.3.7. *Soient (M, g) une variété riemannienne de dimension 3, et O un point de M tel que toutes les courbures sectionnelles en O sont égales à une même valeur k. On a la majoration :*

$$\limsup_{D \to 0} \sup_{\mu \in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{1 + \sqrt{2}}{36} |k|,$$

où $E(D)$ désigne l'ensemble des probabilités portées par la boule géodésique $B(O, \frac{D}{2})$ et dont le barycentre exponentiel est O.

Encadrement du diamètre du barycentre convexe de la loi uniforme portée par une petite sphère S^2 dans la 3-sphère.

La variété M est une boule géodésique régulière de rayon R dans la 3-sphère unité ; soit O le centre de cette boule. Pour $r \in \,]0, R[$ on appelle μ_r la probabilité uniforme sur la sphère de centre O et de rayon r. L'ensemble $b(\mu_r)$ est une boule géodésique centrée en O ; nous allons estimer son rayon.

Théorème 0.3.8. *Si c et C sont des réels positifs tels que $c < \frac{\sqrt{3}}{32}$ et $C > \frac{1+\sqrt{2}}{9}$, alors il existe $r_0 \in \,]0, R[$ tel que pour tout $r \in \,]0, r_0[$, on ait la double inclusion :*

$$B(O, cr^3) \subset b(\mu_r) \subset B(O, Cr^3).$$

Minoration du barycentre convexe d'une loi uniforme portée par une petite sphère S^{n-1} dans l'espace hyperbolique H^n

Dans le cas de la dimension 2, μ est la probabilité uniforme sur un petit cercle C_r de centre o et de rayon r dans H^2. La minoration est obtenue en deux étapes : on construit d'abord un point de $b(\nu)$ aussi éloigné de o que possible, où ν est la probabilité uniforme sur les quatre sommets d'un carré inscrit dans C_r ; on considére ensuite μ comme un mélange de ces mesures ν, lorsque l'un des sommets du carré parcourt un quart du cercle. On trouve ainsi que $b(\mu)$ contient un disque centré en o et de rayon une certaine fonction $f(r)$, équivalente à $\frac{r^3}{8\sqrt{2}}$ lorsque $r \to 0$.

Pour les dimensions plus grandes, en procédant par récurrence sur n et en écrivant μ comme moyenne de mesures portés par des sphères de dimension plus petite, on montre que :

Proposition 0.3.9. *Le barycentre de la loi uniforme portée par une sphère de centre o et de rayon r dans l'espace hyperbolique H^n contient la boule géodésique de centre o et de rayon r_n, où r_n est défini par récurrence à partir de $r_1 = f(r)$ par*

$$r_{n+1} = \int_0^1 \operatorname{Arg\,th}\left(\sqrt{1-s^2}\,\operatorname{th} r_n\right) ds \; .$$

Lorsque $r \to 0$, on a $r_n \sim (\frac{\pi}{4})^{n-1} \frac{r^3}{8\sqrt{2}}$.

Majoration du barycentre convexe d'une mesure dans un espace homogène de dimension 3

Ici (M, g) est un espace homogène de dimension 3, O un point de M ; (E_1, E_2, E_3) une base orthonormée de $T_O M$.

Théorème 0.3.10. *On a les majorations asymptotiques suivantes :*
Si M est la sphère unité de courbure k, alors

$$\lim_{D\to 0} \sup_{\mu} \frac{|b(\mu)|}{D^3} \leq \frac{2+2\sqrt{2}}{9} k \; .$$

Si M est le produit $S^2 \times \mathbb{R}$ (respectivement : $H^2 \times \mathbb{R}$) dont la courbure sectionnelle du plan (E_1, E_2) est R (respt : $(-R)$), alors

$$\lim_{D\to 0} \sup_{\mu} \frac{|b(\mu)|}{D^3} \leq \frac{2+2\sqrt{2}}{9} |R| \; .$$

Si M est l'espace de Heisenberg, alors

$$\lim_{D\to 0}\sup_{\mu}\frac{|b(\mu)|}{D^3}\leq\frac{7+\sqrt{65}}{9}\tau^2\ ,$$

où τ est la courbure de fibration et où le sup *est pris sur les mesures μ portées par la boule géodésique $B(O,\frac{D}{2})$ et dont le barycentre exponentiel est O.*

Le dernier chapitre esquisse une étude des points du bord du barycentre d'une probabilité qui, au contraire des majorations et minorations mentionnées ci-dessus, ne soit plus approchée, mais exacte. Nous donnons une condition suffisante facile pour qu'un point de $b(\mu)$ soit sur le bord $\partial b(\mu)$:

Proposition 0.3.11. *Soit K un compact convexe d'une variété convexe M, μ une probabilité sur K, x un point de $b(\mu)$, et f une fonction convexe au voisinage de K, non constante au voisinage de x, et vérifiant $f(x)=\mu(f)$. Alors x est sur la frontière $\partial b(\mu)$.*

La réciproque de cette proposition n'est pas établie ; nous la proposons sous forme de conjecture :

Conjecture 0.3.12. *Soit M, K et μ comme dans la proposition 0.3.11. Pour tout point $x\in\partial b(\mu)$, il existe une fonction f, définie et convexe sur K, non constante au voisinage de x, et telle que $f(x)=\mu(f)$.*

Chapitre 1

Géodésiques sur les variétés riemaniennes

Dans ce chapitre il sera question des géodésiques sur les variétés riemaniennes. On débutera par des rappels sur les géodésiques : existence, unicité, minimisation locale des longueurs. A la fin du chapitre on décrira les géodésiques sur les espaces homogènes de dimension trois dont le groupe des isométries est de dimension quatre.
La courbure de la variété permet de contrôler les comportements des géodésiques, car si σ est un un 2-plan inclus dans T_pM, la courbure sectionnelle $K(p,\sigma)$ mesure la vitesse de déviation par rapport à σ d'une géodésique passant en $t = 0$ par p et tangente à σ. Le formalisme de cette vitesse de déviation nécessite l'introduction des champs de Jacobi ; la formule

$$|J(t)| = t - \frac{1}{6}K(p,\sigma)t^3 + \widetilde{R}(t), \qquad \lim_{t\to 0} \frac{\widetilde{R}(t)}{t^3} = 0$$

dont la démonstration sera donnée dans ce chapitre, montre que les géodésiques $t \to exp_p(tv(s))$ s'écartent de la géodésique $\gamma(t) = exp_p tv(0)$ avec une vitesse qui diffère de t par un terme d'ordre 3 en t, donné par $-\frac{1}{6}K(p,\sigma)t^3$. Cela veut dire que, localement les géodésiques s'éloignent moins vite les unes des autres que les rayons dans T_pM si $K_p(\sigma) > 0$, et s'éloignent plus vite que les rayons dans T_pM si $K_p(\sigma) < 0$. Pour t suffisamment petit, la valeur $K(p,\sigma)t^3$ fournit l'ampleur approximative de cet éloignement, avec une erreur d'ordre de t^3.

1.1 Préliminaires

Mentionnons que tout au cours du chapitre, lorsqu'on parlera d'une variété M, on la considèrera de dimension n, $n \geq 1$; on notera $\Gamma(M)$ l'ensemble des champs de vecteurs différentiables sur M.
L'analogue dans la variété des mouvements uniformes dans $\mathbb{R}^n$ sont les géodésiques, dont la définition fait intervenir une structure géométrique supplémentaire, la connexion.

Définition 1.1.1. *Une connexion sur une variété M de classe C^∞ est une application bilinéaire*

$$\nabla : \Gamma(M) \times \Gamma(M) \to \Gamma(M)$$

$$(X, Y) \mapsto \nabla_X Y$$

vérifiant les propriétés suivantes :
Soient X et $Y \in \Gamma(M)$ et $f \in C^\infty(M)$ alors,

$$\nabla_{fX} Y = f \nabla_X Y$$

$$\nabla_X (fY) = f \nabla_X Y + X(f) Y$$

Si on se donne un système de coordonnées locales $(x_1, ..., x_n)$ autour d'un point p, on peut écrire

$$X = \sum_i x_i \frac{\partial}{\partial x_i}, Y = \sum_i y_i \frac{\partial}{\partial x_i}$$

et on trouve

$$\nabla_X Y = \sum_k (\sum_{i,j} x_i x_j \Gamma_{ij}^k + X(y_k)) \frac{\partial}{\partial x_k}$$

où

$$\nabla_{\partial_{x_i}} \partial_{x_j} = \sum_k \Gamma_{ij}^k x_k$$

Proposition 1.1.1. *Soit M une variété lisse munie d'une connexion ∇. Alors il existe une unique correspondance qui associe à un champ de vecteurs V le long d'une courbe différentiable $c : I \subset \mathbb{R} \to M$ un autre champ de vecteurs $\frac{DV}{dt}$ le long de c, telle que, si V est induit par un champ de vecteurs $Y \in \Gamma(M)$, i.e, $V(t) = Y(c(t))$, alors*

$$\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} Y$$

On nomme cette opération dérivée covariante le long de c.

La dérivée covariante vérifie

$$\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$$

$$\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$$

où W est un champ de vecteurs le long de c et $f \in C^\infty(M)$. Vu en coordonnées locales,

$$\begin{aligned}\frac{DV}{dt} &= \sum_j \frac{dv_j}{dt}\frac{\partial}{\partial x_j} + \sum_j v_j \frac{D}{dt}(\frac{\partial}{\partial x_j}) \\ &= \sum_j \frac{dv_j}{dt}\frac{\partial}{\partial x_j} + \sum_{i,j} \frac{dx_i}{dt} v_j \nabla_{\partial x_i}\partial x_j \qquad (1.1.1) \\ &= \sum_k (\frac{dv_k}{dt} + \sum_{i,j}\frac{dx_i}{dt} v_j \Gamma^k_{ij})\frac{\partial}{\partial x_k}\end{aligned}$$

où

$$V = \sum_j v_j \frac{\partial}{\partial x_j}$$

Les Γ^k_{ji} qu'on nomme symboles de Christoffel peuvent être quelconques, mais si on travaille sur une variété riemannienne il y a un choix naturel, assuré par le théorème de Levi-Civita 1.1.3.

Définition 1.1.2. *Soit M une variété lisse munie d'une connexion ∇ et d'une métrique riemanienne $\langle\ ,\ \rangle$. La connexion est dite compatible avec la métrique si pour toute courbe différentiable $c : I \to M$ et pour toute paire de champs de vecteurs P et P' parallèles le long de c, on a : $\langle P, P'\rangle =$ constante.*

Proposition 1.1.2. *Une connexion sur la variété M est compatible avec la métrique $\langle\ ,\ \rangle$, si et seulement si pour n'importe quels champs de vecteurs V et W le long d'une courbe différentiable $c : I \to M$, on a :*

$$\frac{d}{dt}\langle V, W\rangle = \langle \frac{DV}{dt}, W\rangle + \langle V, \frac{DW}{dt}\rangle, t \in I \qquad (1.1.2)$$

Définition 1.1.3. *Une connection ∇ sur une variété lisse M est dite symétrique si pour tout $X, Y \in \Gamma(M)$ on a :*

$$\nabla_X Y - \nabla_Y X = [X, Y]$$

Ceci est équivalent à dire que la torsion $T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$ est nulle. En coordonnées locales $(x_1, ..., x_n)$, le fait que ∇ soit symétrique implique $\forall i, j = 1, .., n$

$$\nabla_{\partial_{x_i}} \partial_{x_i} - \nabla_{\partial_{x_j}} \partial_{x_i} = [\partial_{x_i}, \partial_{x_j}] = 0$$

$$\sum_k (\Gamma_{ij}^k - \Gamma_{ji}^k) \frac{\partial}{\partial x_k} = 0$$

$$\text{d'où} \qquad \Gamma_{ij}^k = \Gamma_{ji}^k \ .$$

Théorème 1.1.3. *(Levi-Civita) Étant donné une variété riemannienne M, il existe une unique connexion ∇ sur M satifaisant les conditions :*
1 – ∇ est symétrique,
2 – ∇ est compatible avec la métrique riemannienne.
Cette connexion est appelée connexion de Levi-Civita ou connexion riemannienne sur M.

Dans ce qui suit, on considère que M est une variété riemannienne C^∞ munie de sa connexion riemannienne ; à moins d'un avis contraire.

1.2 Géodésiques

Définition 1.2.1. *Une courbe paramétrée $\gamma : I \subset \mathbb{R} \to M$ est une géodésique en $t_0 \in I$ si*

$$\nabla_{\dot\gamma} \dot\gamma \mid_{t=t_0} = 0$$

Si γ est une géodésique en $t, \forall t \in I$, on dit que γ est une géodésique. Si $[a,b] \subset I$, la restriction de γ à $[a,b]$ est appelée le segment géodésique reliant $\gamma(a)$ à $\gamma(b)$. Géométriquement, une courbe sur M est une géodésique, si celle-ci transporte parallèlement son vecteur vitesse le long d'elle-même.

Un difféomorphisme entre deux variétés M et $(\widetilde{M}, \widetilde{\nabla})$ induit une connexion ∇ sur M.
Une application affine ϕ d'une variété affine (M, ∇) dans une variété affine $(\widetilde{M}, \widetilde{\nabla})$ est une application différentiable telle que pour toute fonction différentiable $f : \widetilde{M} \to \mathbb{R}$, $\nabla d(f \circ \phi) = \phi^*(\widetilde{\nabla} df)$. Une application affine sur M préserve les géodésiques.
Si $\gamma : I \to M$ est une géodésique, alors par (1.1.2) on a :

$$\frac{d}{dt}\langle \frac{d\gamma}{dt}, \frac{d\gamma}{dt} \rangle = 2\langle \frac{D}{dt}(\frac{d\gamma}{dt}), \frac{d\gamma}{dt} \rangle = 2\langle 0, \frac{d\gamma}{dt} \rangle = 0 \ ,$$

c'est-à-dire que la longueur du vecteur tangent $\dot{\gamma}$ est constante. Dans ce qui suit, on supposera que $|\dot{\gamma}| = c \neq 0$, excluant les géodésiques constantes qui se réduisent à un point. On dit qu'une géodésique est normalisée si elle est paramétrée par la longueur d'arc s calculée à partir d'un point origine $t = t_0$ par :

$$s(t) = \int_{t_0}^{t} |\frac{d\gamma}{dt}| dt = c(t - t_0) \ .$$

En utilisant (1.1.1), on trouve localement les équations d'une géodésique

$$0 = \frac{D}{dt}(\frac{d\gamma}{dt}) = \sum_k (\frac{d^2 x_k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{dx_i}{dt}\frac{dx_j}{dt})\frac{\partial}{\partial x_k}$$

Par conséquent, il faut résoudre le systéme d'équations différentielles donné par

$$\frac{d^2 x_k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{dx_i}{dt}\frac{dx_j}{dt} = 0, k = 1, .., n \tag{1.2.3}$$

Le système (1.2.3) nous permet de montrer l'exitence locale et l'unicité des courbes géodésiques, on peut se restreindre à un ouvert de carte au voisinage d'un point p.

Théorème 1.2.1. *(Existence et unicité d'une courbe géodésique.)*
Soit $p \in M$, il existe un voisinage U_p de p tel que pour tout $q \in U_p$ et $v_q \in T_qM$, il existe $\varepsilon > 0$ et une seule géodésique $\gamma :]-\varepsilon, \varepsilon[\subset \mathbb{R} \to M$, de classe C^∞, vérifiant $\gamma(0) = q$ et $\frac{d\gamma}{dt}|_{t=0} = v_q$.

Preuve : Utilisant un système de coordonnées locales au voisinage de p, on est ramené au système (1.2.3) et il suffit d'appliquer le théorème sur les équations différentielles ordinaires.

Corollaire 1.2.2. *Étant donné un point $p \in M$, il existe un ouvert $V \subset M$ contenant p, des nombres $\delta > 0$ et $\varepsilon_1 > 0$ et une application C^∞*

$$\gamma :]-\delta, \delta[\times U \to M, \qquad U = \{(q, v), q \in V, v \in T_qM, |v| < \varepsilon_1\}$$

tels que la courbe $t \mapsto \gamma(t, q, v)$ soit l'unique géodésique sur M vérifiant $\gamma(0, q, v) = q \in V$ et $\gamma'(0, q, v) = v \in T_qM$, pour tout v vérifiant $|v| < \varepsilon_1$; de plus l'application $t \mapsto \gamma(t, q, v)$ dépend de façon lisse des conditions initiales.

Preuve : analogue à la preuve de l'existence et l'unicité locale de la géodésique en considérant le point $(p, 0) \in TM$.

Proposition 1.2.3. *(Homogénéité des géodésiques.)*
Si une géodésique $\gamma(t,q,v)$ est définie sur un intervalle $]-\delta,\delta[$, alors pour $a>0$ la géodésique $\gamma(t,q,av)$, $a\in\mathbb{R}$, $a>0$, est définie sur l'intervalle $]-\frac{\delta}{a},\frac{\delta}{a}[$ et $\gamma(t,q,av)=\gamma(at,q,v)$.

Preuve :
La courbe $h:]-\delta,\delta[\to M$ donnée par $h(t)=\gamma(at,q,v)$, vérifie $h(0)=q$ et

$$\frac{dh(t)}{dt}=\frac{d}{d(at)}\gamma(at,q,v).\frac{d(at)}{dt}=a\gamma'(at,q,v)\ ,$$

donc $h'(0)=av$. De plus,

$$\begin{aligned}\frac{D}{dt}(\frac{dh(t)}{dt})&=\nabla_{h'(t)}h'(t)=\nabla_{a\gamma'(at,q,v)}a\gamma'(at,q,v)\\&=a^2\underbrace{(\nabla_{\gamma'(at,q,v)}\gamma'(at,q,v))}=0\ .\end{aligned}$$

Donc $h(t)$ est une géodésique qui en t vérifie $h(0)=q$ et $h'(0)=av$ et d'après l'unicité de la géodésique,

$$\gamma(at,q,v)=h(t)=\gamma(t,q,av)\ .$$

Le corollaire (1.2.2) et la proposition (1.2.3) permettent d'énoncer :

Proposition 1.2.4. *Soit $p\in M$. Il existe un voisinage V de p dans M, un nombre $\varepsilon>0$ et une application $\gamma:]-2,2[\times U\to M$ où $U=\{(q,w)\in TM, q\in V, w\in T_qM, |w|<\varepsilon\}$ tels que $t\mapsto\gamma(t,q,w), t\in]-2,2[$ est l'unique géodésique sur M qui passe par q en $t=0$ avec la vitesse w, pour tout $q\in V$ et $w\in T_qM$, $|w|<\varepsilon$.*

Preuve :
Soit la géodésique obtenue dans le corollaire (1.2.2)

$$\gamma(t,q,v):]-\delta,\delta[\to M,\qquad \gamma(0,q,v)=q,\qquad |\gamma'(0,q,v)|<\varepsilon_1$$

par l'homogénéité de la géodésique (proposition 1.2.3) on définit :

$$\tilde{\gamma}:]-2,2[\to M,\qquad \tilde{\gamma}=\gamma(\frac{\delta t}{2},q,v)=\gamma(t,q,\frac{\delta v}{2}).$$

On pose $w=\frac{\delta v}{2}$,en prenant $\varepsilon=\frac{\delta\varepsilon_1}{2}$ on a bien que $\tilde{\gamma}(t,q,w)$ est définie pour $|t|<2$ et $|w|<\varepsilon$.

Définition 1.2.2. *Soit (M,g) une variété riemannienne, N une sous variété de M.*
Pour tout $(p,v)\in TN$, on considère $\gamma_{(p,v)}:I\to M$ la géodésique dans M telle que $\gamma(0)=p$ et $\dot{\gamma}(0)=v$.
N est dite totalement géodésique si $\gamma_{(p,v)}(I)\subset N$ pour tout $(p,v)\in TN$.

1.2.1 L'application exponentielle

Définition 1.2.3. *Soit* $\mathfrak{D} = \{(q,v), q \in M, v \in T_qM\} \subset TM$ *tel que* $\gamma(t,q,v)$ *existe pour tout* $t \in [0,1]$*. On définit l'application exponentielle comme :*

$$exp : \mathfrak{D} \to M \qquad (q,v) \mapsto \gamma(1,q,v) = \gamma(|v|, q, \frac{v}{|v|}) \tag{1.2.4}$$

De plus, on note l'application exp restreinte à un sous ensemble ouvert du plan tangent T_qM centré en l'origine O, comme :

$$exp_q : B_\varepsilon(O) \subset T_qM \to M \tag{1.2.5}$$

où $exp_q(v) = exp(q,v)$ et $B_\varepsilon(O) = \{v \in T_qM, |v| < \varepsilon\}$, d'après la proposition (1.2.4), on peut toujours définir $B_\varepsilon(O)$ de tel sorte que $exp_q(v)$ existe pour $v \in B_\varepsilon(O)$. On peut aussi remarquer d'après le corollaire (1.2.2) que exp est différentiable.

Proposition 1.2.5. *Pour tout* $q \in M$*, il existe* $\varepsilon > 0$ *tel que* $exp_q : B_\varepsilon(O) \subset T_qM \to M$ *est un difféomorphisme de* $B_\varepsilon(O)$ *vers un voisinage* $V \subset M$ *contenant* q*.*

Preuve :
Calculons $d(exp_q)_O$:
Notons que $d(exp_q)_O : T_O(T_qM) \approx T_qM \to T_qM$ (l'identification est possible puisque les plans tangents sont isomorphes à $\mathbb{R}^n$). Soit $v \in T_O(T_qM)$, on peut considérer la courbe $\rho(t) = tv$ au voisinage de O. En $t = 0$, $\rho(0) = 0$ et $\rho'(0) = v \in T_O(T_qM)$. Ainsi,

$$d(exp_q)_O(v) = d(exp_q)_{\rho(0)}\rho'(0)$$

$$\frac{d}{dt}(exp_q(\rho(t)))|_{t=0} = \frac{d}{dt}(exp_q(tv))|_{t=0}$$

$$\frac{d}{dt}(\gamma(1,q,tv))|_{t=0} = \frac{d}{dt}(\gamma(t,q,v))|_{t=0} = v$$

Par conséquent, $d(exp_q)_O$ est l'identité sur T_qM, et d'après le théorème de la fonction inverse, il existe $\varepsilon > 0$ et un voisinage V de q tels que exp_q soit un difféomorphisme de $B_\varepsilon(O)$ vers V.

1.2.2 Propriété minimisante

Définition 1.2.4. *Une courbe différentiable par morceaux est une application continue* $c : [a, b] \subset \mathbb{R} \to M$ *satisfaisant la condition suivante : il existe une partition* $a < t_0 < t_1 < ... < t_{k-1} < t_k = b$ *de* $[a, b]$ *telle que les restrictions* $c|_{t_i, t_i+1}, i = 1, ..., k-1$ *sont différentiables. Les points* $c(t_i)$ *sont appelés les sommets de c.*

Définition 1.2.5. *Un segment de géodésique* $\gamma : [a, b] \to M$ *est dit minimisant si* $\mathfrak{L}(\gamma) \leq \mathfrak{L}(c)$*, où* $\mathfrak{L}(\ )$ *désigne la longueur d'une courbe et c décrit toutes les courbes différentiables par morceaux joignant les points* $\gamma(a)$ *et* $\gamma(b)$.

Définition 1.2.6. *Soit A un ensemble connexe de* $\mathbb{R}^2$*,* $U \subset A \subset \bar{U}$*, où U est un ouvert tel que* ∂A *soit une courbe différentiable par morceau dont l'angle formé par les sommets est différent de* π*. Une surface paramétrée dans M est une application différentiable* $s : A \subset \mathbb{R}^2 \to M$*. Un champ de vecteur V sur s est une application qui associe à chaque point* (u, v) *de A un vecteur* $V((u, v)) \in T_{s(u,v)}M$ *qui est différentiable dans le sens suivant : si f est une fonction différentiable sur M, alors l'application* $(u, v) \mapsto V(u, v)f$ *est différentiable.*
Pour v_0 *fixe, l'application* $A \cap \{v = v_0\} \to s(u, v_0)$ *est une courbe dans M. Sur celle-ci on définit le champ de vecteurs* $\frac{\partial s}{\partial u}$ *en posant* $\frac{\partial s}{\partial u} = ds(\frac{\partial}{\partial u})$*. En faisant varier v ceci définit* $\frac{\partial s}{\partial u}$ *pour tout* $(u, v) \in A$*. De façon analogue on définit* $\frac{\partial s}{\partial v}$.
Si V est un champ de vecteurs défini sur $s : A \to M$ *on définit les dérivées covariantes* $\frac{DV}{\partial u}$ *et* $\frac{DV}{\partial v}$ *de la façon suivante :* $\frac{DV}{\partial u}$ *est la dérivée covariante le long de la courbe* $u \mapsto s(u, v_0)$ *où on restreint V à celle-ci. En faisant varier v on définit ainsi* $\frac{DV}{\partial u}$ *sur tout A. De façon analogue on définit aussi* $\frac{DV}{\partial v}$.

Lemme 1.2.6. *(Symétrie)*
Soit M une variété différentiable munie d'une connexion symétrique et $s : (u, v) \subset A \to M$ *une surface paramétrée. Alors*

$$\frac{D}{\partial v}\frac{\partial s}{\partial u} = \frac{D}{\partial u}\frac{\partial s}{\partial v} .$$

Preuve :
Soient $(u_0, u_1) \in A$ et (U_α, ϕ_α) une carte telle que $s(u_0, v_0) \in U_\alpha$. On peut écrire : $\phi_\alpha \circ s(u, v) = (x_1(u, v), ..., x_n(u, v))$; alors en (u_0, v_0) on a

$$\frac{D}{\partial v}(\frac{\partial(\phi_\alpha \circ s)}{\partial u}) = \frac{D}{\partial v}(\sum_i \frac{\partial x_i}{\partial u}\frac{\partial}{\partial x_i})$$

$$= \sum_i \frac{\partial^2 x_i}{\partial u \partial v} \frac{\partial}{\partial x_i} + \sum_i \frac{\partial x_i}{\partial u} \nabla_{\sum_j (\frac{\partial x_j}{\partial v}) \frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_i}$$

$$= \sum_i \frac{\partial^2 x_i}{\partial u \partial v} \frac{\partial}{\partial x_i} + \sum_{i,j} \frac{\partial x_i}{\partial u} \frac{\partial x_j}{\partial v} \nabla_{\frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_i}$$

De même,

$$\frac{D}{\partial u}(\frac{\partial(\phi_\alpha \circ s)}{\partial v}) = \sum_j \frac{\partial^2 x_j}{\partial u \partial v} \frac{\partial}{\partial x_j} + \sum_{i,j} \frac{\partial x_j}{\partial u} \frac{\partial x_i}{\partial v} \nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j}$$

Et d'après la symétrie de la connexion,

$$\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = \nabla_{\frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_i}$$

Alors

$$\frac{D}{\partial v} \frac{\partial s}{\partial u} = \frac{D}{\partial u} \frac{\partial s}{\partial v} \; .$$

Lemme 1.2.7. *(Lemme de Gauss.)*
Soient $p \in M$ et $v \in T_pM$ tels que $exp_p v$ soit définie. Soit $w \in T_pM \approx T_v(T_pM)$. Alors

$$\langle (d\,exp_p)_v v, (d\,exp_p)_v w \rangle = \langle v, w \rangle \; .$$

Preuve :
Puisque

$$exp_p : T_pM \to M$$

alors

$$(d\,exp_p)_v : T_v(T_pM) \to T_{exp_p v} M$$

Soit $w = w_T + w_N$, où w_T est parallèle à v et w_N est normal à w. Comme $d\,exp_p$ est linéaire,

$$\langle (d\,exp_p)_v v, (d\,exp_p)_v (w_T + w_N) \rangle = \langle (d\,exp_p)_v v, (d\,exp_p)_v w_T \rangle + \langle (d\,exp_p)_v v, (d\,exp_p)_v w_N \rangle .$$

D'après la définition de $d\,exp_p$

$$\langle (dexp_p)_v v, (d\,exp_p)_v w_T \rangle = \langle v, w_T \rangle .$$

Considérons le cas $w = w_N$, comme $exp_p v$ est définie, il existe $\varepsilon > 0$ tel que $exp_p u$ est définie pour

$$u = tv(s), \qquad 0 \le t \le 1, \qquad -\varepsilon < s < \varepsilon$$

où $v(s)$ est une courbe dans T_pM telle que $v(0) = v$, $v'(0) = w_N$ et $|v(s)| =$ *constante*. On peut alors considérer la surface paramétrée

$$f : A \to M, \qquad A = \{(t,s), 0 \leq t \leq 1\}, -\varepsilon < s < \varepsilon$$

donnée par

$$f(t,s) = exp_p(tv(s)).$$

Les courbes $t \to f(t, s_0)$ sont des géodésiques. De plus, on a :

$$(d\,exp_p)_v(w_N) = (d\,exp_p)_v(v'(s))|_{s=0} = \frac{\partial}{\partial s}(exp_p(v(s)))|_{s=0} = \frac{\partial f}{\partial s}(1,s)|_{s=0} = \frac{\partial f}{\partial s}(1,0)$$

$$(d\,exp_p)(v) = (d\,exp_p)_v(\frac{\partial}{\partial t}(t(v(0))|_{t=1} = \frac{\partial}{\partial t}(exp_p(t(v(0))|_{t=1} = \frac{\partial}{\partial t}(t,0)|_{t=1} = \frac{\partial}{\partial t}(1,0)$$

Alors

$$\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle(1,0) = \langle (d\,exp_p)_v(w_N), (d\,exp_p)_v(v) \rangle \; ; \qquad (1.2.6)$$

de plus, quelque soit (t,s), on a

$$\frac{\partial}{\partial t}\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle = \langle \frac{D}{\partial t}\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle + \langle \frac{\partial f}{\partial s}, \frac{D}{\partial t}\frac{\partial f}{\partial t} \rangle.$$

Le deuxième terme est nul puisque $\frac{\partial f}{\partial t}$ est la vitesse d'une courbe géodésique, et comme la connexion est symétrique

$$\langle \frac{D}{\partial t}\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle = \langle \frac{D}{\partial s}\frac{\partial f}{\partial t}, \frac{\partial f}{\partial t} \rangle = \frac{1}{2}\langle \frac{\partial}{\partial s}\frac{\partial f}{\partial t}, \frac{\partial f}{\partial t} \rangle$$

$$= \frac{1}{2}\frac{\partial}{\partial s}|\frac{\partial f}{\partial t}|^2 = \frac{1}{2}\frac{\partial}{\partial s}|v(s)|^2 = 0$$

Par conséquent,

$$\frac{\partial}{\partial t}\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle = 0$$

C'est-à-dire

$$\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle$$

ne dépend pas de t. Comme $\lim_{t\to 0} \frac{\partial f}{\partial s}(t,0) = \lim_{t\to 0}(d\,exp_p)_{tv}(tw_N) = 0$, on en déduit par (1.2.6) que

$$\langle (d\,exp_p)_v(w_N), (d\,exp_p)_v(v) \rangle = \langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle(1,0) = \langle 0, \frac{\partial f}{\partial t}(0,0) \rangle = 0\ .$$

D'où

$$\langle v, w \rangle = \langle v, w_T \rangle = \langle (d\,exp_p)_v(v), (d\,exp_p)_v(w_T) \rangle$$

$$= \langle (d\,exp_p)_v(v), (d\,exp_p)_v(w)\rangle\ .$$

Une des conséquences directes du lemme de Gauss est que l'orthogonalité par rapport aux directions radiales est conservée d'un plan tangent à l'autre lorsqu'on suit une géodésique, et que les vecteurs radiaux sont de norme constante comme on s'y attendait.

Définition 1.2.7. *Si exp_p est un difféomorphisme dans un voisinage V de l'origine de T_pM, alors $exp_pV = U$ est appelé voisinage normal de p. Si $\overline{B_\varepsilon(O)} \subset V$, on appelle $exp_pB_\varepsilon(O) = B_\varepsilon(p)$ boule normale (ou boule géodésique) de centre p et de rayon ε.*

Par le lemme de Gauss, la frontière d'une boule normale est une hypersurface (sous-variété de codimension 1) dans M orthogonale à la géodésique issue de p notée $S_\varepsilon(p)$ et nommée sphère normale (ou sphère géodésique). Les géodésiques dans $B_\varepsilon(p)$ issues de p sont appelées géodésiques radiales.

Théorème 1.2.8. *Soit $p \in M$, U un voisinage normal de p, $B \subset U$ une boule normale de centre p et $\gamma : [0,1] \to B$ un segment de géodésique tel que $\gamma(0) = p$. Si $c : [0,1] \to M$ est une courbe différentiable par morceaux joignant les points $\gamma(0)$ et $\gamma(1)$ alors $\mathfrak{L}(\gamma) \leq \mathfrak{L}(c)$ et l'égalité a lieu seulement quand $\gamma([0,1]) = c([0,1])$.*

Supposons tout d'abord que $c([0,1]) \subset B$. Comme exp_p est un difféommorphisme sur U, la courbe $c(t)$, pour $t \neq 0$, peut être donnée de façon unique comme $exp_pr(t)v(t) = f(r(t),t)$ où $t \to v(t)$ est une courbe dans T_pM telle que $|v(t)| = 1$ et $r : [0,1] \to \mathbb{R}^+$ et une fonction positive et différentible par morceaux (coor. polaires). On peut supposer que si $t_1 \in]0,1]$ alors $c(t_1) \neq p$; autrement, on ignore l'intervalle $[0,t_1[$. Comme c est différentiable par morceaux, on a :

$$\frac{dc}{dt} = \frac{\partial f}{\partial r}r'(t) + \frac{\partial f}{\partial t}$$

sauf pour un nombre fini de points. Alors,

$$|\frac{dc}{dt}|^2 = \langle\frac{dc}{dt},\frac{dc}{dt}\rangle = \langle\frac{\partial f}{\partial r}r'(t) + \frac{\partial f}{\partial t}, \frac{\partial f}{\partial r}r'(t) + \frac{\partial f}{\partial t}\rangle$$

$$= \langle\frac{\partial f}{\partial r}r'(t), \frac{\partial f}{\partial r}r'(t)\rangle + 2\langle\frac{\partial f}{\partial r}r'(t), \frac{\partial f}{\partial r}\rangle + \langle\frac{\partial f}{\partial t}, \frac{\partial f}{\partial t}\rangle$$

$$= (r'(t))^2|\frac{\partial f}{\partial r}|^2 + 2r'(t)\langle\frac{\partial f}{\partial r}, \frac{\partial f}{\partial t}\rangle + |\frac{\partial f}{\partial t}|^2$$

Par le lemme de Gauss $\langle \frac{\partial f}{\partial r}, \frac{\partial f}{\partial t} \rangle = 0$. En effet :

$$\begin{aligned}(d\, exp_p)_{rv}(v(t)) &= (d\, exp_p)_{rv} \frac{\partial}{\partial r}((rv(t))) \\ &= \frac{\partial}{\partial r}(exp_p(rv(t))) = \frac{\partial}{\partial r} f(r,t)\ .\end{aligned}$$

D'autre part,

$$\begin{aligned}(d\, exp_p)_{rv}(rv'(t)) &= \frac{\partial}{\partial t}(exp_p(rv(t))) \\ &= \frac{\partial}{\partial t} f(r,t)\ .\end{aligned}$$

Alors,

$$\langle \frac{\partial f}{\partial r}, \frac{\partial f}{\partial t} \rangle = \langle (d\, exp_p)_{rv}(v(t)), (d\, exp_p)_{rv}(rv'(t)) \rangle = \langle v(t), rv'(t) \rangle = 0\ .$$

De plus comme $|\frac{\partial f}{\partial t}| = |v(t)| = 1$,

$$|\frac{dc}{dt}|^2 = (r'(t))^2 + |\frac{\partial f}{\partial t}|^2 \geq (r'(t))^2 \tag{1.2.7}$$

Donc

$$\int_\varepsilon^1 |\frac{dc}{dt}| dt \geq \int_\varepsilon^1 |r'(t)| dt \geq \int_\varepsilon^1 r'(t) dt = r(1) - r(\varepsilon). \tag{1.2.8}$$

En prenant la limite quand ε tend vers 0 on obtient $\mathfrak{L}(c) \geq \mathfrak{L}(\gamma)$, car :

$$\begin{aligned} & c(1) = exp(r(1)v(1)) = \gamma(1) \\ \Longrightarrow \quad & \mathfrak{L}(\gamma) = |r(1)v(1)| = r(1)|v(1)| = r(1) \\ & \Longrightarrow \quad r(1) = \mathfrak{L}(\gamma)\ .\end{aligned}$$

Si les inégalités (1.2.7) et (1.2.8) sont strictes alors $\mathfrak{L}(c) > \mathfrak{L}(\gamma)$. Si $\mathfrak{L}(c) = \mathfrak{L}(\gamma)$ alors $|\frac{\partial f}{\partial t}| = 0$ (1.2.7), i.e. $v(t) = cste$ et de plus $|r'(t)| = r'(t) > 0$ (1.2.8). Il s'ensuit que c est une reparamétrisation de γ, i.e. $c(t) = exp(r(t)v)$ et $\gamma([0,1]) = c([0,1])$.

Si $c([0,1])$ n'est pas contenue dans B, on considère le premier point $t_1 \in]0,1[$ tel que $c(t_1)$ appartienne à ∂B. En appelant ρ le rayon de la boule géodésique, on a :

$$\mathfrak{L}(c) \geq \mathfrak{L}_{[0,t_1]}(c) \geq \rho > \mathfrak{L}(\gamma).$$

Remarque 1.2.1. *D'après la proposition précédente et les propriétés minimisantes des géodésiques, il est clair que chaque deux points d'un voisinage normal $exp_p(B_\varepsilon(O))$ d'un point p, peuvent être joints par au moins une géodésique de longueur inférieure à 2ε.*

1.3 Géodésiques dans les espaces homogènes C-V

On considère une famille de variétés riemanniennes $(M, ds^2_{l,m})$, de dimension 3 et à deux paramètres, où les métriques sont exprimées par

$$ds^2_{l,m} = \frac{dx^2 + dy^2}{[1 + m(x^2 + y^2)]^2} + (dz + \frac{l}{2}\frac{ydx - xdy}{[1 + m(x^2 + y^2)]^2})^2 \tag{1.3.9}$$

avec $l, m \in \mathbb{R}$; la variété M est égale à $\mathbb{R}^3$ si $m \geq 0$ et à $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 < -\frac{1}{m}\}$ sinon.
On trouve ces métriques dans la classification des espaces homogènes de dimension 3, donnée par L. Bianchi en 1897. Leur ecriture sous la forme (1.3.9) est due à É. Cartan et à G. Vranceanu (voir [6]) ; pour cette raison on les appelle métriques de Cartan-Vranceanu (C-V metrics). L'intérêt géométrique de ces métriques réside dans le fait qu'elles décrivent tous les espaces homogènes de dimension 3 dont les groupes d'isométries sont de dimension 4 ou 6, à l'exception des espaces à courbure sectionnelle constante négative.

Si $l = 0$, alors M est le produit d'une surface S à courbure de Gauss constante $4m$, avec la droite réelle $\mathbb{R}$.

Si $4m - l^2 = 0$, alors M a une courbure sectionnelle constante positive.

Si $l \neq 0$ et $m > 0$, M est localement $SU(2)$.

De façon similaire, si $l \neq 0$ et $m < 0$, M est localement $\widetilde{SL}(2, \mathbb{R})$.

Le cas où $m = 0$ et $l \neq 0$, correspond à une métrique invariante à gauche sur le groupe de Lie de Heisenberg $\mathbb{H}_3$.

Théorème 1.3.1. *Si X est un champ de vecteurs de Killing sur la variété riemannienne (M, g) alors les géodésiques $\gamma(t)$ vérifient l'équation*

$$g(\dot{\gamma}, X) = c$$

où $\dot{\gamma}$ est le champ de vecteurs vitesse de $\gamma(t)$ et c est une constante.

Preuve :
La dérivée par rapport à t du produit scalaire $g(\dot{\gamma}, X)$ donne

$$\frac{d}{dt} g(\dot{\gamma}, X) = g(\nabla_{\dot{\gamma}}\dot{\gamma}, X) + g(\dot{\gamma}, \nabla_{\dot{\gamma}} X) = 0$$

car γ est une géodésique et X est un champ de Killing. Ainsi le produit scalaire $g(\dot{\gamma}, X)$ est constant.

Maintenant, nous allons chercher les géodésiques des espaces homogènes simplement connexes de dimension 3 dont les groupes des isométries sont de dimension 4.
La métrique C-V peut être mise sous la forme

$$ds^2_{l,m} = \sum_{i=1}^{3} w^i \otimes w^i$$

Où, en posant $D = 1 + m(x^2 + y^2)$,

$$w^1 = \frac{dx}{D} \qquad w^2 = \frac{dy}{D} \qquad w^3 = dz + \frac{l}{2}\frac{ydx - xdy}{D}$$

La base duale orthonormée des champs de vecteurs est

$$E_1 = D\frac{\partial}{\partial x} - \frac{l}{2}y\frac{\partial}{\partial z} \qquad E_2 = D\frac{\partial}{\partial y} + \frac{l}{2}x\frac{\partial}{\partial z} \qquad E_3 = \frac{\partial}{\partial z} \tag{1.3.10}$$

Le champ de Killing de la métrique (1.3.9) est le champ $X = \xi^i E_i$ tel que la dérivée de Lie de la métrique par rapport à X est nulle

$$L_X(ds^2_{l,m}) = 0 \ .$$

Une base de l'algère de Lie des champs de vecteurs de Killing est donnée par [6] :

$$X = \frac{2mxy}{D}E_1 + (1 - \frac{2mx^2}{D})E_2 - \frac{lx}{D}E_3$$
$$Y = (1 - \frac{2my^2}{D})E_1 + \frac{2mxy}{D}E_2 + \frac{ly}{D}E_3$$
$$Z = E_3$$
$$R = -\frac{y}{D}E_1 + \frac{x}{D}E_2 - \frac{l(x^2+y^2)}{2D}E_3 \ .$$

Soit $\gamma : I \to M$ une géodésique sur la variété $(M, ds^2_{l,m})$. Le champ de vecteurs tangents $\dot{\gamma}$ par rapport à la base (1.3.10) est donné par

$$\dot{\gamma} = \frac{\dot{x}}{D}E_1 + \frac{\dot{y}}{D}E_2 + [\dot{z} - \frac{l}{2}\frac{x\dot{y} - y\dot{x}}{D}]E_3$$

D'après le théorème (1.3.1), on a le système :

$$\begin{cases} \frac{2mxy\dot{x}}{D^2} + \frac{(1+m(y^2-x^2))\dot{y}}{D^2} - a_3\frac{lx}{D} = a_1 \\ \frac{(1+m(x^2-y^2))\dot{x}}{D^2} + \frac{2mxy\dot{y}}{D^2} + a_3\frac{ly}{D} = a_2 \\ \dot{z} - \frac{l}{2}\frac{x\dot{y}-y\dot{x}}{D} = a_3 \\ \frac{\dot{y}x}{D^2} - \frac{\dot{x}y}{D^2} - a_3\frac{l(x^2+y^2)}{2D} = a_4 \end{cases} \quad a_i \in \mathbb{R}$$

Supposons que la géodésique passe au temps $t = 0$ par une origine et telle que $\dot{\gamma}(0) = (u, v, w)$, les constantes a_i prennent les valeurs

$$a_1 = v \qquad a_2 = u \qquad a_3 = w \qquad a_4 = 0$$

et le système précédent devient :

$$\begin{cases} \frac{2mxy\dot{x}}{D^2} + \frac{(1+m(y^2-x^2))\dot{y}}{D^2} - \frac{lxw}{D} = v \\ \frac{(1+m(x^2-y^2))\dot{x}}{D^2} + \frac{2mxy\dot{y}}{D^2} + \frac{lyw}{D} = u \\ \dot{z} - \frac{l}{2}\frac{x\dot{y}-y\dot{x}}{D} = w \\ \frac{\dot{y}x}{D^2} - \frac{\dot{x}y}{D^2} - \frac{l(x^2+y^2)w}{2D} = 0 \ . \end{cases} \tag{1.3.11}$$

On remarque que le champ de vecteurs de Killing R qui engendre les rotations autour de l'axe des z peut être écrit en combinaison des autres champs de vecteurs de Killing :

$$R = \frac{-y}{m(x^2+y^2)-1}X + \frac{x}{m(x^2+y^2)-1}Y - \frac{l(x^2+y^2)}{2(m(x^2+y^2)-1)}Z$$

En appliquant le théorème (1.3.1) à R on trouve :

$$\frac{-y}{m(x^2+y^2)-1}a_1 + \frac{x}{m(x^2+y^2)-1}a_2 - \frac{l(x^2+y^2)}{2(m(x^2+y^2)-1)}a_3 = 0$$

Ce qui est équivalent à :

$$l(x^2+y^2)w - 2xu + 2yv = 0 \ . \tag{1.3.12}$$

On a

Proposition 1.3.2. *Pour les métriques C-V, les géodésiques partant d'une origine et telles que $\dot{\gamma}(0) = (u, v, w)$, sont définies comme intersection de deux surfaces :*
- Un cylindre circulaire dont l'axe de révolution est parallèle à l'axe des z, ou un plan parallèle à l'axe des z ;
- Une surface de révolution autour de l'axe des z.

Preuve :

D'après l'équation (1.3.12), si $l \neq 0$ et $w \neq 0$, on trouve que les géodésiques du groupe de Heisenberg $\mathbb{H}_3$, du groupe de Lie $SU(2)$ et du revêtement universel de $SL(2,\mathbb{R})$ sont contenues dans un cylindre dont l'axe est parallèle à l'axe des z.

Pour $l \neq 0$ et $w = 0$ les géodésiques sont contenues dans le plan $ux - vy = 0$ et pour $l = 0$ nous obtenons les géodésiques des produits $\mathbb{S}^2 \times \mathbb{R}$ et $\mathbb{H}^2 \times \mathbb{R}$ qui sont aussi dans le plan $ux - vy = 0$.

La surface de révolution est obtenue par rotation de la géodésique autour de l'axe des z.

On va décrire brièvement comment trouver les equations des géodésiques dans les cas de $SU(2)$ et $\widetilde{SL}(2,\mathbb{R})$. Écrivons le système (1.3.11) et l'égalité $|\dot{\gamma}| = 1$ en coordonnées cylindriques

$$\begin{cases} \frac{\dot{\rho}cos\theta}{1+m\rho^2} - \frac{\dot{\theta}sin\theta}{(1+m\rho^2)^2}(\rho - m\rho^3) + \frac{l\rho sin\theta}{1+m\rho^2}w = v \\ \frac{\dot{\rho}sin\theta}{1+m\rho^2} + \frac{\dot{\theta}cos\theta}{(1+m\rho^2)^2}(\rho - m\rho^3) - \frac{l\rho cos\theta}{1+m\rho^2}w = u \\ \dot{z} - \frac{l}{2}\frac{\rho^2\dot{\theta}}{1+m\rho^2} = w \\ \frac{\rho^2\dot{\theta}}{(1+m\rho^2)^2} - \frac{l}{2}\frac{\rho^2 w}{1+m\rho^2} = 0 \\ \frac{\dot{\rho}^2+\rho^2\dot{\theta}^2}{(1+m\rho^2)^2} = u^2 + v^2 \end{cases} \tag{1.3.13}$$

Pour $w = 0$ le système (1.3.13) devient :

$$\begin{cases} \dot{\rho}cos\theta = v(1+m\rho^2) \\ \dot{\rho}sin\theta = u(1+m\rho^2) \\ \dot{z} = 0 \\ \dot{\theta} = 0 \\ \frac{\dot{\rho}^2}{(1+m\rho^2)^2} = u^2 + v^2 \end{cases} \Rightarrow \begin{cases} \dot{\rho}cos\theta = v(1+m\rho^2) \\ \dot{\rho}sin\theta = u(1+m\rho^2) \\ z = 0 \\ \theta = a \\ \frac{\dot{\rho}}{1+m\rho^2} = \pm\sqrt{u^2+v^2} \end{cases}$$

Une intégration immédiate de la dernière équation donne les équations des géodésiques de $SU(2)$ et $\widetilde{SL}(2,\mathbb{R})$, respectivement :

$$\begin{cases} x = \frac{v}{\sqrt{u^2+v^2}}\frac{tan(\sqrt{m(u^2+v^2)}t)}{\sqrt{m(u^2+v^2)}} \\ y = \frac{u}{\sqrt{u^2+v^2}}\frac{tan(\sqrt{m(u^2+v^2)}t)}{\sqrt{m(u^2+v^2)}} \\ z = 0 \end{cases} \quad et \begin{cases} x = \frac{v}{\sqrt{u^2+v^2}}\frac{tanh(\sqrt{-m(u^2+v^2)}t)}{\sqrt{-m(u^2+v^2)}} \\ y = \frac{u}{\sqrt{u^2+v^2}}\frac{tanh(\sqrt{-m(u^2+v^2)}t)}{\sqrt{-m(u^2+v^2)}} \\ z = 0 \end{cases}$$

Si $w \neq 0$, de la dernière équation du système (1.3.13) on tire

$$d\theta = \frac{lw}{2}(1+m\rho^2)dt, \qquad \frac{d\rho}{(1+m\rho^2)\sqrt{(u^2+v^2)-\frac{l^2}{4}w^2\rho^2}} = \pm dt\ .$$

Posons

$$\frac{l^2}{4}w^2 = a^2, \qquad u^2+v^2 = b^2, \qquad a^2+b^2m \neq 0.$$

Par intégration on obtient

$$\rho^2 = \frac{b^2 tanAt}{A^2+a^2 tanAt} \qquad \theta = arctan\frac{lw\,tanAt}{A}$$

où $A = \sqrt{l^2w^2+4m(u^2+v^2)}$.

Pour $l \neq 0$ et $l^2w^2+4m(u^2+v^2) > 0$, on a les équations :

$$\begin{cases} x = \frac{2tan(\frac{At}{2})}{\sqrt{A^2+l^2w^2tan^2(\frac{At}{2})}}(vcosT - usinT) \\ y = \frac{2tan(\frac{At}{2})}{\sqrt{A^2+l^2w^2tan^2(\frac{At}{2})}}(ucosT + vsinT) \\ z = wt - \frac{l^2w}{4m}t - \frac{lw}{2m}T \end{cases}$$

Avec $T = arctan\frac{lw\,tan(\frac{At}{2})}{A}$

Si $m > 0$ et $4m \neq l^2$ les équations déterminent les géodésiques de $SU(2)$, quant au cas $4m = l^2$ on a les géodésiques de la sphère $\mathbb{S}^3$. Si $m < 0$ ces équations déterminent les géodésiques de $\widetilde{SL}(2,\mathbb{R})$.

Pour $l \neq 0$ et $l^2+w^2+4m(u^2+v^2) < 0$,$(m<0)$, les équations paramétriques des géodésiques de $\widetilde{SL}(2,\mathbb{R})$ dans le cas $w \neq 0$ sont :

$$\begin{cases} x = \frac{2tanh(\frac{Ct}{2})}{\sqrt{C^2+l^2w^2tanh^2(\frac{Ct}{2})}}(vcosT' - usinT') \\ y = \frac{2tanh(\frac{Ct}{2})}{\sqrt{A^2+l^2w^2tanh^2(\frac{Ct}{2})}}(ucosT' + vsinT') \\ z = wt - \frac{l^2w}{4m}t - \frac{lw}{2m}T' \end{cases}$$

où $C = \sqrt{-l^2w^2-4m(u^2+v^2)}$ et $T' = arctan\frac{lw\,tanh(\frac{Ct}{2})}{C}$.

Pour $l \neq 0$ et $l^2+w^2+4m(u^2+v^2) = 0$,$(m<0)$ les équations paramétriques des géodésiques de $\widetilde{SL}(2,\mathbb{R})$ dans le cas $w \neq 0$ sont :

$$\begin{cases} x = \frac{2t}{\sqrt{4+l^2w^2t^2}}(vcosT - usinT) \\ y = \frac{2t}{\sqrt{4+l^2w^2t^2}}(ucosT + vsinT) \\ z = wt - \frac{l^2w}{4m}t - \frac{l}{2m}T \end{cases} \qquad T = arctan\frac{lwt}{2}\ .$$

Si $m = 0$ et $l \neq 0$, les équations paramétriques des géodésiques du groupe de Heisenberg $\mathbb{H}_3$ pour $w \neq 0$ et $w = 0$ sont respectivement :

$$\left\{\begin{array}{l} x(t) = \frac{u}{lw}cos(lwt) + \frac{v}{lw}sin(lwt) - \frac{u}{lw} \\ y(t) = \frac{u}{lw}sin(lwt) - \frac{v}{lw}cos(lwt) + \frac{v}{lw} \\ z(t) = wt + \frac{u^2+v^2}{2w}t - \frac{u^2+v^2}{2w}sin(lwt) \end{array}\right. \quad et \quad \left\{\begin{array}{l} x = vt \\ y = ut \\ z = 0\ . \end{array}\right.$$

Si $m > 0$ et $l = 0$ on a les équations cartésiennes des géodésiques de $\mathbb{S}^2 \times \mathbb{R}$ dans les cas $w \neq 0$ et $w = 0$

$$\left\{\begin{array}{l} ux - vy = 0 \\ x^2 + y^2 = \frac{1}{m}tan^2(\frac{\sqrt{m(u^2+v^2)}}{w}z) \end{array}\right. \quad et \quad \left\{\begin{array}{l} ux - vy = 0 \\ z = 0\ . \end{array}\right.$$

Si $m < 0$ et $l = 0$ on a les équations cartésiennes des géodésiques de $\mathbb{H}^2 \times \mathbb{R}$ dans les cas $w \neq 0$ et $w = 0$

$$\left\{\begin{array}{l} ux - vy = 0 \\ x^2 + y^2 = -\frac{1}{m}tanh^2(\frac{\sqrt{-m(u^2+v^2)}}{w}z) \end{array}\right. \quad et \quad \left\{\begin{array}{l} ux - vy = 0 \\ z = 0\ . \end{array}\right.$$

1.4 Courbure et géodésiques

La notion de courbure a été intoduite par Riemann d'une manière assez géométrique, que nous allons maintenant décrire. Soit $p \in M$, $\sigma \subset T_pM$ un sous-espace de dimension 2 de l'espace tangent T_pM à la variété M en p. Considérons l'ensemble des géodésiques passant au temps $t = 0$ par p et tangentes à σ. Les segments de ces géodésiques dans un voisinage normal $U \subset M$ de p déterminent une sous-variété S de M de dimension 2, qui est l'image de exp_p restreinte à $\sigma \cap exp_p^{-1}(U)$, et qui hérite la métrique induite par l'inclusion. Puisque Gauss a prouvé que la courbure d'une surface peut être exprimée en termes de sa métrique, Riemann pouvait parler de la courbure de S en p ; il l'a notée $K(\sigma, p)$, et on l'appelle maintenant courbure sectionnelle de M en p par rapport à σ ; c'est la courbure considérée par Riemann.
Riemann n'a pas montré comment calculer cette quantité à partir de la métrique. Quelques années plus tard Christoffel a donné des formules permettant ce calcul et montrant que la courbure mesure la déviation de la variété M du cas euclidien ; cette notion de courbure généralise la notion de courbure de Gauss pour les surfaces et coïncide avec le concept de Riemann.

Définition 1.4.1. *La courbure R d'une variété riemannienne M est une correspondance qui associe à chaque paire de champs de vecteurs différentiables $X, Y \in \Gamma(M)$ une application $R(X,Y) : \Gamma(M) \to \Gamma(M)$ définie par :*

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, \qquad Z \in \Gamma(M)$$

où ∇ est la connexion riemannienne de M.

Observons que si $M = \mathbb{R}^n$, alors $R(X,Y)Z = 0$ pour tous $X, Y, Z \in \Gamma(\mathbb{R}^n)$.
En effet si le champ de vecteurs Z est donné par $Z = (z_1, z_2, ..., z_n)$ alors $\nabla_X Z = (Xz_1, Xz_2, ..., Xz_n)$ et $\nabla_Y \nabla_X Z = (YXz_1, YXz_2, ..., YXz_n)$ ce qui implique $R(X,Y)Z = 0$. On peut donc voir R comme un outil de mesure de la déviation de M du cas euclidien.
Une autre façon de voir la définition 1.4.1 est le point de vue local, en considérant un système de coordonnées $\{x_i\}$ autour de $p \in M$. Puisque $[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0$, on obtient :

$$R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})\frac{\partial}{\partial x_k} = (\nabla_{\frac{\partial}{\partial x_i}} \nabla_{\frac{\partial}{\partial x_j}} - \nabla_{\frac{\partial}{\partial x_j}} \nabla_{\frac{\partial}{\partial x_i}})\frac{\partial}{\partial x_k}$$

La courbure mesure donc la non-commutativité de la dérivée covariante.

Proposition 1.4.1. *la courbure R d'une variété riemannienne a les propriétés suivantes :*
i) R est bilinéaire dans $\Gamma(M) \times \Gamma(M)$

$$R(fX_1 + gX_2, Y_1) = fR(X_1, Y_1) + gR(X_2, Y_1),$$

$$R(X_1, fY_1 + gY_2) = fR(X_1, Y_1) + gR(X_1, Y_2),$$

$f, g \in D(M), X_1, X_2, Y_1, Y_2 \in \Gamma(M)$.
ii) Pour toute paire $X, Y \in \Gamma(M)$, l'opérateur $R(X,Y) : \Gamma(M) \to \Gamma(M)$ est linéaire

$$R(X,Y)(Z + W) = R(X,Y)Z + R(X,Y)W,$$

$$R(X,Y)fZ = fR(X,Y)Z,$$

$f \in D(M), Z, W \in \Gamma(M)$

Preuve de ii) :
La première égalité est évidente, pour la deuxième on a

$$\nabla_Y \nabla_X (fZ) = \nabla_Y (f \nabla_X Z + (Xf)Z)$$

$$= f\nabla_Y\nabla_X Z + (Yf)(\nabla_X Z) + (Xf)(\nabla_Y Z) + (Y(Xf))Z.$$

Et puisque

$$\nabla_Y\nabla_X(fZ) - \nabla_X\nabla_Y(fZ) = f(\nabla_Y\nabla_X - \nabla_X\nabla_Y)Z + ((YX - XY)f)Z,$$

alors

$$R(X,Y)fZ = f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_Y Z + ([X,Y]f)Z - f\nabla_{[X,Y]}Z + ([Y,X]f)Z$$

$$= fR(X,Y)Z \qquad \diamond$$

En regardant la preuve précédente, on remarque que la linéarité de $R(X,Y) : \Gamma(M) \to \Gamma(M)$ impose l'apparition du terme $\nabla_{[X,Y]}Z$ dans la définition de la courbure.

Proposition 1.4.2. *(Identité de Bianchi.)*

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$$

Preuve :
$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z + \nabla_Y\nabla_Z X - \nabla_Z\nabla_Y X - \nabla_{[Y,Z]}X + \nabla_Z\nabla_X Y - \nabla_X\nabla_Z Y - \nabla_{[Z,X]}Y$
$\nabla_X(\nabla_Y Z - \nabla_Z Y) + \nabla_Y(\nabla_Z X - \nabla_X Z) + \nabla_Z(\nabla_X Y - \nabla_Y X) - \nabla_{[Y,Z]}X - \nabla_{[X,Z]}Y - \nabla_{[X,Y]}Z$
$= \nabla_X[Y,Z] - \nabla_{[Y,Z]}X + \nabla_Y[Z,X] - \nabla_{[Z,X]}Y + \nabla_Z[X,Y] - \nabla_{[X,Y]}Z$
$= [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$
où on a utilisé deux fois la symétrie de la connexion pour arriver à l'identité de Jacobi.

Pour simplifier l'écriture nous allons poser $g(RX,Y)Z,T) = (X,Y,Z,T)$.

Proposition 1.4.3. *(a)* $(X,Y,Z,T) + (Y,Z,X,T) + (Z,X,Y,T) = 0$
(b) $(X,Y,Z,T) = -(Y,X,Z,T)$
(c) $(X,Y,Z,T) = -(X,Y,T,Z)$
(d) $(X,Y,Z,T) = (Z,T,X,Y)$.

Preuve :
(a) c'est juste l'identité de Bianchi ;
(b) découle directement de la dénition (1.4.1) ;
(c) est équivalente à $(X,Y,Z,Z) = 0$ dont la preuve est donnée par :

$$(X,Y,Z,Z) = g(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z, Z).$$

Mais
$$g(\nabla_X \nabla_Y Z, Z) = Xg(\nabla_Y Z, Z) - g(\nabla_Y Z, \nabla_X Z),$$
et
$$g(\nabla_{[X,Y]} Z, Z) = \frac{1}{2}[X,Y]g(Z,Z),$$
alors
$$(X,Y,Z,Z) = Xg(\nabla_Y Z, Z) - Yg(\nabla_X Z, Z) - \frac{1}{2}[X,Y]g(Z,Z)$$
$$= \frac{1}{2}X(Yg(Z,Z)) - \frac{1}{2}Y(Xg(Z,Z)) - \frac{1}{2}[X,Y]g(Z,Z)$$
$$= \frac{1}{2}[X,Y]g(Z,Z) - \frac{1}{2}[X,Y]g(Z,Z) = 0,$$
ce qui prouve (c).
Pour prouver (d), nous allons utiliser (a) en écrivant :
$$(X,Y,Z,T) + (Y,Z,X,T) + (Z,X,Y,T) = 0,$$
$$(Y,Z,T,X) + (Z,T,Y,X) + (T,Y,Z,X) = 0,$$
$$(Z,T,X,Y) + (T,X,Z,Y) + (X,Z,T,Y) = 0,$$
$$(T,X,Y,Z) + (X,Y,T,Z) + (Y,T,X,Z) = 0,$$
L'addition des quatre équations donne
$$2(Z,X,Y,T) + 2(T,Y,Z,X) = 0$$
Et par suite
$$(Z,X,Y,T) = (Y,T,Z,X) \qquad \diamond$$
Il est commode d'exprimer ce qu'on a vu ci-dessus dans un système de coordonnées (U,x) au voisinage d'un point p de la variété. On note $\frac{\partial}{\partial x_i}$ par X_i et on pose $R(X_i,X_j)X_k = \sum_l R^l_{ijk} X_l$, de sorte que R^l_{ijk} sont les composantes de R dans (U,x). Si $X = \Sigma_i u^i X_i$, $Y = \Sigma_j v^j X_j$, $Z = \Sigma_k w^k X_k$, on obtient, d'après la linéarité de R,
$$R(X,Y)Z = \Sigma_{i,j,k,l} R^l_{ijk} u^i v^j w^k X_l \ .$$
Pour exprimer R^l_{ijk} en fonction des coefficients Γ^k_{ij} de la connexion riemannienne, on écrit :
$$R(X_i,X_j)X_k = \nabla_{X_i}\nabla_{X_j} X_k - \nabla_{X_j}\nabla_{X_i} X_k$$

$$= \nabla_{X_i}(\sum_l \Gamma^l_{jk} X_l) - \nabla_{X_j}(\sum_l \Gamma^l_{ik} X_l)$$

Un calcul direct donne

$$R^s_{ijk} = \sum_l \Gamma^l_{jk}\Gamma^s_{il} - \sum_l \Gamma^l_{ik}\Gamma^s_{jl} + \frac{\partial}{\partial x_i}\Gamma^s_{jk} - \frac{\partial}{\partial x_j}\Gamma^s_{ik} \ .$$

Posons

$$g(R(X_i, X_j)X_k, X_s) = \sum_l R^l_{ijk} g_{ls} = R_{ijks} \ .$$

On peut donc écrire les identités de la proposition (1.4.3) sous la forme :

$$R_{ijks} + R_{jkis} + R_{kijs} = 0$$

$$R_{ijks} = -R_{jiks}$$

$$R_{ijks} = -R_{ijsk}$$

$$R_{ijks} = R_{ksij} \ .$$

Proposition 1.4.4. *Soit $\sigma \subset T_pM$ un sous-espace vectoriel de dimension 2 de T_pM et $x, y \in \sigma$ deux vecteurs linéairement indépendants, alors*

$$K(x, y) = \frac{(x, y, x, y)}{|x \wedge y|^2}$$

ne dépend pas du choix des vecteurs $x, y \in \sigma$.

Preuve :
On remarque qu'on peut passer d'une base $\{x, y\}$ de σ à une autre base $\{x', y'\}$ par itération des transformations élémentaires suivantes :
(a) $\{x, y\} \to \{y, x\}$,
(b) $\{x, y\} \to \{\lambda x, y\}$,
(c) $\{x, y\} \to \{x + \lambda y, y\}$,
et il est facile de remarquer que $K(x, y)$ est invariante par de telles transformations.

Définition 1.4.2. *Etant donné un point $p \in M$ et un sous-espace bidimensionnel $\sigma \subset T_pM$, le nombre réel $K(x, y) = K(\sigma)$, où $\{x, y\}$ est une base de σ, est appelé courbure sectionnelle de σ en p.*

L'importance géométrique de la courbure sectionnelle réside dans le fait que sa connaissance détermine complètement la courbure R.

Lemme 1.4.5. *Soit V un espace vectoriel de dimension $n \geq 2$, muni d'un produit scalaire $\langle \ , \ \rangle$. Soit $R : V \times V \times V \to V$ et $R' : V \times V \times V \to V$ deux applications tri-linéaires vérifiant les conditions (a),(b),(c) et (d) de la proposition (1.4.3), en posant :*

$$(x, y, z, t) = \langle R(x, y)z, t\rangle \qquad (x, y, z, t)' = \langle R'(x, y)z, t\rangle \ .$$

Si x, y sont deux vecteurs linéairement indépendants, on écrit,

$$K(\sigma) = \frac{(x, y, x, y)}{|x \wedge y|^2} \qquad K'(\sigma) = \frac{(x, y, x, y)'}{|x \wedge y|^2} \ ,$$

où σ est le sous-espace de dimension 2 engendré par x et y.
Si pour tout $\sigma \subset V$, $K(\sigma) = K'(\sigma)$ alors $R = R'$.

Preuve :
Pour tout $x, y \in V$, on a, par hypothèse $(x, y, x, y) = (x, y, x, y)'$, alors

$$(x + z, y, x + z, y) = (x + z, y, x + z, y)',$$

et par suite

$$(x, y, x, y) + 2(x, y, z, y) + (z, y, z, y) = (x, y, x, y)' + 2(x, y, z, y)' + (z, y, z, y)'$$

ce qui donne

$$(x, y, z, y) = (x, y, z, y)'.$$

Utilisons ce qu'on vient de prouver, nous obtenons

$$(x, y + t, z, y + t) = (x, y + t, z, y + t)',$$

ce qui entraine

$$(x, y, z, t) + (x, t, z, y) = (x, y, z, t)' + (x, t, z, y)',$$

qui peut être écrit :

$$(x, y, z, t) - (x, y, z, t)' = (y, z, x, t) - (y, z, x, t)'.$$

Il s'ensuit que l'expression

$$(x, y, z, t) - (x, y, z, t)'$$

est invariante par permutations circulaires des trois premiers éléments.
D'après (a) de la proposition (1.4.3), on a

$$3[(x, y, z, t) - (x, y, z, t)'] = 0$$

C'est-à-dire

$$(x, y, z, t) = (x, y, z, t)'$$

pour tout $x, y, z, t \in V$. $\diamond$

Lemme 1.4.6. *Soit (M, g) une variété riemannienne et p un point de M. $R' : T_pM \times T_pM \times T_pM \to T_pM$ est l'application tri-linéaire définie par*

$$g(R'(X, Y, W), Z) = g(X, W)g(Y, Z) - g(Y, W)g(X, Z)$$

pout tout $X, Y, W, Z \in T_pM$. Alors M a une courbure sectionnelle constante égale à K_0 si et seulement si $R = K_0R'$, où R est la courbure de M.

Preuve :
Supposons que $K(p, \sigma) = K_0$ pour tout $\sigma \subset T_pM$ et observons que R', tel que $g(R'(X, Y, W), Z) = (X, Y, W, Z)'$, satisfait les conditions (a), (b), (c) et (d) de la proposition (1.4.3). Puisque

$$(X, Y, X, Y)' = g(X, X)g(Y, Y) - (g(X, Y))^2,$$

on a, pour toute paire $X, Y \in T_pM$,

$$R(X, Y, X, Y) = K_0(g(X, X)g(Y, Y) - (g(X, Y))^2) = K_0R'(X, Y, X, Y).$$

Le lemme (1.4.5) implique que, pour tout X, Y, W, Z,

$$R(X, Y, W, Z) = K_0R'(X, Y, W, Z),$$

alors $R = K_0R'$. La réciproque est immédiate.

Corollaire 1.4.7. *Soit (M, g) une variété riemannienne, $n = dimM$, p un point de M et $\{e_1, e_2, ..., e_n\}$, une base orthonormée de T_pM. Posons $R_{ijkl} = g(R(e_i, e_j)e_k, e_l), i, j, k, l = 1, ..., n$. Alors $K(p, \sigma) = K_0$ pour tout $\sigma \subset T_pM$, si et seulement si*

$$R_{ijkl} = K_0(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}),$$

où

$$\delta_{ij} = \begin{cases} 1, & si \quad i = j \\ 0, & si \quad i \neq j. \end{cases}$$

En d'autres termes $K(p, \sigma) = K_0$ pour tout $\sigma \subset T_pM$, si et seulement si $R_{ijij} = -R_{ijji} = K_0$ pour tout $i \neq j$, et $R_{ijkl} = 0$ dans les autres cas.

Nous allons maintenant établir une relation trés utile dans la suite.
Soit $f : A \subset \mathbb{R}^2 \to M$ une surface paramétrée et (s, t) les coordonnées usuelles de $\mathbb{R}^2$. Soit $V = V(s, t)$ un champ de vecteurs le long de f. Pour tout (s, t) il est possible de définir $R(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t})V$ d'une manière évidente.

Lemme 1.4.8.

$$\frac{D}{\partial t}\frac{D}{\partial s}V - \frac{D}{\partial s}\frac{D}{\partial t}V = R(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t})V$$

Preuve :

Choisissons un système de coordonnées (U, x) autour d'un point $p \in M$. Soit $V = \sum_i v^i X_i$, où $v^i = v^i(s,t)$ et $X_i = \frac{\partial}{\partial x_i}$. Alors

$$\frac{D}{\partial s}V = \frac{D}{\partial s}V(\sum_i v^i X_i) = \sum_i v^i \frac{D}{\partial s}X_i + \sum_i \frac{\partial v^i}{\partial s}X_i,$$

et

$$\frac{D}{\partial t}(\frac{D}{\partial s}V) = \sum_i v^i \frac{D}{\partial t}\frac{D}{\partial s}X_i + \sum_i \frac{\partial v^i}{\partial t}\frac{D}{\partial s}X_i + \sum_i \frac{\partial v^i}{\partial s}\frac{D}{\partial t}X_i + \sum_i \frac{\partial^2 v^i}{\partial t \partial s}X_i.$$

Par permutation des rôles de s et t dans l'expression précédente puis par soustraction, on obtient

$$\frac{D}{\partial t}\frac{D}{\partial s}V - \frac{D}{\partial s}\frac{D}{\partial t}V = \sum_i v^i(\frac{D}{\partial t}\frac{D}{\partial s}X_i - \frac{D}{\partial s}\frac{D}{\partial t}X_i).$$

Calculons maintenant $\frac{D}{\partial t}\frac{D}{\partial s}X_i$. On pose

$$f(s,t) = (x_1(s,t), ..., x_n(s,t)).$$

Alors $\frac{\partial f}{\partial s} = \sum_j \frac{\partial x_j}{\partial s}X_j$ et $\frac{\partial f}{\partial t} = \sum_k \frac{\partial x_k}{\partial t}X_k$. On a

$$\frac{D}{\partial s}X_i = \nabla_{\sum_j \frac{\partial x_j}{\partial s}X_j}X_i = \sum_j \frac{\partial x_j}{\partial s}\nabla_{X_j}X_i$$

et

$$\frac{D}{\partial t}\frac{D}{\partial t}X_i = \frac{D}{\partial t}(\sum_j \frac{\partial x_j}{\partial s}\nabla_{X_j}X_i)$$

$$= \sum_j \frac{\partial^2 x_j}{\partial t \partial s}\nabla_{X_j}X_i + \sum_j \frac{\partial x_j}{\partial s}\nabla_{\sum_k \frac{\partial x_k}{\partial t}X_k}(\nabla_{X_j}X_i)$$

$$= \sum_j \frac{\partial^2 x_j}{\partial t \partial s}\nabla_{X_j}X_i + \sum_{jk} \frac{\partial x_j}{\partial s}\frac{\partial x_k}{\partial t}\nabla_{X_k}\nabla_{X_j}X_i,$$

$$(\frac{D}{\partial t}\frac{D}{\partial s} - \frac{D}{\partial s}\frac{D}{\partial t})X_i = \sum_{jk} \frac{\partial x_j}{\partial s}\frac{\partial x_k}{\partial t}(\nabla_{X_k}\nabla_{X_j}X_i - \nabla_{X_j}\nabla_{X_k}X_i).$$

Finalement

$$(\frac{D}{\partial t}\frac{D}{\partial s} - \frac{D}{\partial s}\frac{D}{\partial t})V = \sum_{ijk} v^i \frac{\partial x_j}{\partial s}\frac{\partial x_k}{\partial t}R(X_j, X_k)X_i = R(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t})V. \qquad \diamond$$

1.4.1 Champs de Jacobi

Dans cette partie on va tirer une relation entre les deux notions fondamentales introduites préalablement, à savoir les géodésiques et la courbure. Nous verrons que la courbure $K(p,\sigma)$, $\sigma \subset T_pM$, détermine à quelle vitesse les géodésiques passant en $t=0$ et tangentes à σ, s'écartent de σ. Afin de donner un formalisme à cette vitesse de déviation, il est nécessaire d'introduire les champs de vecteurs dits champs de Jacobi. Ces champs sont définis par une équation différentielle qui apparait lors de l'étude de la fonction exponentielle. Les champs de Jacobi permettent d'obtenir une caractérisation simple des singularités de l'exponentielle.
Soit (M,g) une variété riemannienne et $p \in M$. Dans la preuve du lemme de Gauss, nous avons vu que exp_p est définie en $v \in T_pM$, et si $w \in T_v(T_pM)$, alors

$$(d\,exp_p)_v w = \frac{\partial f}{\partial s}(1,0),$$

où f est une surface paramétrée donnée par

$$f(t,s) = exp_p tv(s), \qquad 0 \le t \le 1, \qquad -\varepsilon \le s \le \varepsilon,$$

et $v(s)$ est une courbe dans T_pM avec $v(o) = v$ et $v'(0) = w$.
Nous voulons acquérir de l'information sur $|(d\,exp_p)_v w|$, car cette quantité mesure intuitivement, la vitesse de déviation de la géodésique $t \mapsto exp_p tv(s)$ passant au temps $t=0$ par p. Nous verrons que cette vitesse dépend de la courbure sectionnelle en p par rapport au plan engendré par v et w. Un autre interêt de la connaissance de $|(d\,exp_p)_v w|$ est que son annulation pour $w \neq 0$ implique que v est un point critique de exp_p.
Etudions le champ

$$(d\,exp_p)_{tv}(tw) = \frac{\partial f}{\partial s}(t,0)$$

le long de la géodésique $\gamma(t) = exp_p(tv)$, $\qquad 0 \le t \le 1$.
Puisque γ est une géodésique, on a pour tout (t,s),$\frac{D}{\partial t}\frac{\partial f}{\partial t} = 0$.
D'après le lemme (1.4.8),

$$\begin{aligned} 0 = \frac{D}{\partial s}(\frac{D}{\partial t}\frac{\partial f}{\partial t}) &= \frac{D}{\partial t}(\frac{D}{\partial s}\frac{\partial f}{\partial t}) - R(\frac{\partial f}{\partial s},\frac{\partial f}{\partial t})\frac{\partial f}{\partial t} \\ &= \frac{D}{\partial t}(\frac{D}{\partial t}\frac{\partial f}{\partial s}) + R(\frac{\partial f}{\partial t},\frac{\partial f}{\partial s})\frac{\partial f}{\partial t}. \end{aligned}$$

Posons $\frac{\partial f}{\partial s}(t,0) = J(t)$, on trouve que J vérifie l'équation différentielle

$$\frac{D^2 J}{dt^2} + R(\gamma'(t), J(t))\gamma'(t) = 0 \ . \tag{1.4.14}$$

L'équation (1.4.14) est dite équation de Jacobi.

Définition 1.4.3. *Soit $\gamma : [0, a] \to M$ une géodésique dans M. Un champ de vecteurs le long de γ est appelé champ de Jacobi s'il vérifie l'équation (1.4.14), pour tout $t \in [0, a]$.*

Un champ de Jacobi est déterminé par ces conditions initiales $J(0)$ et $\frac{DJ}{dt}(0)$.
Soit $e_1(t), e_2(t), ..., e_n(t)$ un champ parallèle orthonormé le long de γ, avec $n = dimM$; on écrit :

$$J(t) = \sum_i f_i(t)e_i(t), \quad a_{ij} = g(R(\gamma'(t), e_i(t))\gamma'(t), e_j(t)), \quad i, j = 1, ..., n.$$

Alors

$$\frac{D^2 J}{dt^2} = \sum_i f_i''(t)e_i(t),$$

et

$$R(\gamma', J)\gamma' = \sum_j g(R(\gamma', J)\gamma', e_j)e_j = \sum_{ij} f_i g(R(\gamma', e_i)\gamma', e_j)e_j = \sum_{ij} f_i a_{ij} e_j.$$

L'équation (1.4.14) est equivalente au système

$$f_j''(t) + \sum_i a_{ij}(t)f_i(t) = 0, \quad j = 1, ..., n,$$

qui est un système linéaire de second ordre. Étant donné $J(0)$ et $\frac{DJ}{dt}(0)$, il existe une solution du système qui est de classe C^∞ et définie dans $[0, a]$; il existe donc $2n$ champs de Jacobi linéairement indépendants, le long de γ.

Remarque 1.4.1. *Il est important de remarquer que $\gamma'(t)$ et $t\gamma'(t)$ sont des champs de Jacobi le long de γ. Le premier champ a une dérivée nulle et il est de norme constante ; quant au second, il n'est nul qu'en $t = 0$. Suite à cette remarque on considérera les champs de Jacobi le long de γ qui sont normaux à γ'.*

Soit (M, g) une variété riemannienne à courbure sectionnnelle constante K, $\gamma : [0, l] \to M$ une géodésique sur M, et J un champ de Jacobi le long de γ normal à γ'. Puisque $|\gamma'(t)| = 1$ et d'après le lemme (1.4.6),

$$R(\gamma', J)\gamma' = KJ.$$

Pour tout champ de vecteurs T le long de γ, on a

$$g(R(\gamma', J)\gamma', T) = K\{g(\gamma', \gamma')g(J, T) - g(\gamma', T)g(J, \gamma')\} = Kg(J, T),$$

L'équation de Jacobi peut être écrite

$$\frac{D^2 J}{dt^2} + KJ = 0 \ . \tag{1.4.15}$$

Soit $w(t)$ un champ de vecteurs parallèle le long de γ avec $g(\gamma'(t), w(t)) = 0$ et $|w(t)| = 1$. Il est facile de vérifier que

$$J(t) = \begin{cases} \frac{\sin(t\sqrt{K})}{\sqrt{K}} w(t), & si \ \ K > 0, \\ tw(t), & si \ \ K = 0, \\ \frac{\sinh(t\sqrt{-K})}{\sqrt{-K}} w(t), & si \ \ K < 0, \end{cases}$$

est une solution de (1.4.15) avec les conditions initiales $J(0) = 0$, $J'(0) = w(0)$.

Etant donné un point $p \in M$, $v \in T_pM$ et $w \in T_v(T_pM)$, on peut construire un champ de Jacobi le long de la géodésique $\gamma : [0, 1] \to M$ donnée par $\gamma(t) = exp_p(tv)$. Pour cela on considère la surface paramétrée définie par $f(t, s) = exp_p tv(s)$, où $v(s)$ est une courbe dans T_pM avec $v(0) = v$, $v'(0) = w$, on prend $J(t) = \frac{\partial f}{\partial s}(t, 0)$. Remarquons que $J(0) = 0$.
Plus précisement, on a la proposition suivante

Proposition 1.4.9. *Soit $\gamma : [0, a] \to M$ une géodésique et J un champ de Jacobi le long de γ avec $J(0) = 0$ et $\frac{DJ}{dt}(0) = w$ et $\gamma'(0) = v$. En considérant w comme un élément de $T_{av}(T_{\gamma(0)}M)$, on construit une courbe $v(s)$ dans $T_{\gamma(0)}M$ avec $v(0) = av$, $v'(0) = w$, on pose $f(t, s) = exp_p(\frac{t}{a}v(s))$, $p = \gamma(0)$ et on définit $\overline{J}$ par $\overline{J}(t) = \frac{\partial f}{\partial s}(t, 0)$. Alors $\overline{J} = J$ sur $[0, a]$.*

Preuve :
Pour $s = 0$, on a

$$\begin{aligned} \frac{D}{dt}\frac{\partial f}{\partial s} &= \frac{D}{\partial t}((d\,exp_p)_{tv}(tw)) = \frac{D}{\partial t}(t(d\,exp_p)_{tv}(w)) \\ &= (d\,exp_p)_{tv}(w) + t\frac{D}{\partial t}((d\,exp_p)_{tv}(w)). \end{aligned}$$

Pour $t = 0$,

$$\frac{D\overline{J}}{dt}(0) = \frac{D}{\partial t}\frac{\partial f}{\partial s}(0, 0) = (d\,exp_p)_O(w) = w.$$

Puisque $J(0) = \overline{J}(0) = 0$ et $\frac{DJ}{dt}(0) = \frac{D\overline{J}}{dt}(0) = w$, on conclut, d'après le théorème d'unicité, que $J = \overline{J}$.

Corollaire 1.4.10. *Soit $\gamma : [0,a] \to M$ une géodésique. Un champ de Jacobi J le long de γ tel que $J(0) = 0$ est donné par*

$$J(t) = (d\,exp_p)_{t\gamma'(0)}(tJ'(0)), \qquad t \in [0,a].$$

Proposition 1.4.11. *Soit $p \in M$ et $\gamma : [0,a] \to M$ une géodésique avec $\gamma(0) = p$ et $\gamma'(0) = \gamma v$. Soit $w \in T_v(T_pM)$ tel que $|w| = 1$ et J un champ de Jacobi le long de γ donné par*

$$J(t) = (d\,exp_p)_{tv}(tw), \qquad t \in [0,a].$$

Le développement de Taylor de $|J(t)|^2$ au voisinage de $t = 0$ est donné par

$$|J(t)|^2 = t^2 - \frac{1}{3}g(R(v,w)v,w)t^4 + R(t) \tag{1.4.16}$$

où $\lim_{t\to 0} \frac{R(t)}{t^4} = 0$.

Preuve :
Puisque $J(0) = 0$ et $J'(0) = w$, on a

$$g(J,J) = 0,$$

$$(g(J,J))'(0) = 2g(J,J')(0) = 0,$$
$$(g(J,J))''(0) = 2g(J',J')(0) + 2g(J'',J)(0) = 2.$$

D'autre part, puisque $J''(0) = -R(\gamma',J)\gamma'(0) = 0$, on a

$$(g(J,J))''' = 6g(J',J'')(0) + 2g(J''',J)(0) = 0.$$

Nous allons démontrer puis utiliser la formule

$$\nabla_{\gamma'}(R(\gamma',J)\gamma')(0) = R(\gamma',J')\gamma'(0). \tag{1.4.17}$$

Pour prouver (1.4.17), notons que pour tout W, on a en $t = 0$,

$$g(\frac{D}{dt}(R(\gamma',J)\gamma'),W) = \frac{d}{dt}g(R(\gamma',W)\gamma',J) - g(R(\gamma',J)\gamma',W')$$

$$= g(\frac{D}{dt}(R(\gamma',W)\gamma'),J) + g(R(\gamma',W)\gamma',J')$$
$$= g(R(\gamma',J')\gamma',W).$$

De (1.4.17) et de l'équation de Jacobi, on déduit $J'''(0) = -R(\gamma', J')\gamma'(0)$ et par suite,

$$\begin{aligned}(g(J,J))'''(0) &= 8g(J', J''')(0) + 6g(J'', J'')(0) + 2g(J''', J)(0) \\ &= -8g(J', R\gamma', J')\gamma')(0) = -8g(R(v,w)v, w).\end{aligned}$$

En rassemblant le calcul ci-dessous, nous obtenons (1.4.16).

Corollaire 1.4.12. *Si $\gamma : [0,l] \to M$ est paramétrée par la longueur d'arc, et $g(w,v) = 0$, l'expression $g(R(v,w)v,w)$ est la courbure sectionnelle au point p par rapport au plan σ engendré par v et w. Dans ce cas*

$$|J(t)|^2 = t^2 - \frac{1}{3}K(p,\sigma)t^4 + R(t).$$

Corollaire 1.4.13. *Avec les mêmes conditions que dans le corollaire précédent, on a*

$$|J(t)| = t - \frac{1}{6}K(p,\sigma)t^3 + \widetilde{R}(t), \quad \lim_{t\to 0} \frac{\widetilde{R}(t)}{t^3} = 0. \tag{1.4.18}$$

Remarquons que l'expression (1.4.18) contient la relation entre les géodésiques et la courbure. En effet, si l'on considère la surface paramétrée

$$f(t,s) = exp_p tv(s), \quad t \in [0,\delta], \quad s \in \,]-\varepsilon, \varepsilon[,$$

où δ est choisi suffisamment petit pour que $exp_p tv(s)$ soit définie, et $v(s)$ est une courbe dans T_pM avec $|v(s)| = 1$, $v(0) = 0$, $v'(0) = w$, on voit que les rayons $t \to tv(s)$, $t \in [0,\delta]$, qui passent par l'origine O de T_pM en $t = 0$, s'écartent du rayon $t \to tv(0)$ avec la vitesse

$$|(\frac{\partial}{\partial s} tv(s))(0)| = |tw| = t.$$

D'autre part, l'expression (1.4.18) montre que les géodésiques $t \to exp_p(tv(s))$ s'écartent de la géodésique $\gamma(t) = exp_p tv(0)$ avec une vitesse qui diffère de t par un terme d'ordre 3 en t, donné par $-\frac{1}{6}K(p,\sigma)t^3$. Cela veut dire que, localement, les géodésiques s'écartent moins que les rayons dans T_pM si $K_p(\sigma) > 0$, et s'écartent plus que les rayons dans T_pM si $K_p(\sigma) < 0$. Pour t suffisamment petit, la valeur $K(p,\sigma)t^3$ fournit une approximation de l'ampleur de cet écartement avec une erreur d'ordre t^3.

1.4.2 Points conjugués

Dans cette partie on étudie la relation entre les singularités de l'application exponentielle et les champs de Jacobi.

Définition 1.4.4. *Soit $\gamma : [0, a] \to M$ une géodésique. Pour $t_0 \in]0, a]$, le point $\gamma(t_0)$ est dit conjugué à $\gamma(0)$ le long de γ, s'il existe un champ de Jacobi le long de γ, non identiquement nul, avec $J(0) = 0 = J(t_0)$. Le nombre maximal de ces champs linéairement indépendants est appelé la multiplicité du point conjugué $\gamma(t_0)$.*

Il est clair que si $\gamma(t_0)$ est conjugué à $\gamma(0)$ alors $\gamma(0)$ est conjugué à $\gamma(t_0)$.

Remarque 1.4.2. *Si la dimension de M est n, alors il existe exactement n champs de Jacobi linéairement indépendants le long de la géodésique $\gamma : [0, a] \to M$, qui sont nuls en $\gamma(0)$. On déduit que les champs de Jacobi $J_1, ..., J_k$ avec $J_i(0) = 0$ sont linéairement indépendants si et seulement si les vecteurs $J'_1(0), ..., J'_k(0)$ sont linéairement indépendants. De plus le champ de Jacobi $J(t) = t\gamma'(t)$ ne s'annule pour aucun $t \neq 0$, ce qui implique que la multiplicité d'un point conjugué ne dépasse pas $n - 1$.*

Définition 1.4.5. *L'ensemble des (premiers) points conjuguées du point $p \in M$, pour toutes les géodésiques passant en $t = 0$ par p, est appelé le conjugué locus de p.*

Définition 1.4.6. *Soit M une variété riemannienne complète, soit $p \in M$ et $\gamma : [0, \infty[\to M$ une géodésique telle que $\gamma(0) = p$. Si t est suffisamment petit, alors $d(\gamma(0), \gamma(t)) = t$, et $\gamma([0, t])$ est minimisante ; de plus, si $\gamma([0, t_1])$ n'est pas minimisante, il est de même pour tout $t > t_1$. Par continuité, l'ensemble des nombres $t > 0$ pour lesquels $d(\gamma(0), \gamma(t)) = t$ est de la forme $[0, t]$ ou $[0, \infty[$. Dans le premier cas, $\gamma(0)$ est dit le cut point de p le long de γ ; dans l'autre cas, on dit que ce cut point n'existe pas.*
On définit le cutlocus de p, noté $C_m(p)$, comme la réunion des cut points de p le long de toutes les géodésiques passant en $t = 0$ par p.

Définition 1.4.7. *Soit B une boule géodésique fermée de centre p et de rayon r dans une variété riemannienne. On dit que B est une boule géodésique régulière si $r\sqrt{k} < \frac{\pi}{2}$ et si B ne rencontre pas le cutlocus de p, k étant un majorant positif des courbures sectionnelles sur B.*

1.4.3 Courbure et métrique en coordonnées normales

Théorème 1.4.14. *Soit (M, g) une variété riemannienne, p un point de M. Notons S^{n-1} la sphère unité de T_pM.*

1. En coordonnées polaires $(r, v) \mapsto rv$, $\mathbb{R}_+ \times S^{n-1} \to T_pM$,

$$exp_p^*(g_M) = dr^2 + g_r$$

où, en un point $v \in S^{n-1}$ de la sphère unité, et pour un vecteur $w \in T_vS^{n-1}$ tangent à la sphère en v,

$$g_r(w) = |w|^2r^2 - \frac{1}{3}g(R(w,v)v,w)r^4 + o(r^4).$$

2. Choisissons des coordonnées cartésiennes $(x_1, ..., x_n)$ sur T_pM. Notons

$$R_{ijkl} = g(R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_l}).$$

Alors

$$exp_p^*(g_M) = \sum_{i,j}(\delta_{ij} - \frac{1}{3}\sum_{k,l} R_{iklj}x_kx_l + o(|x|^2)dx_idx_j).$$

Preuve :
Soit $s \mapsto v(s)$ une courbe tracée dans la sphère unité de T_pM, telle que $v(0) = v$ et $v'(0) = w$. Soit $\gamma_s(t) = exp_p(tv(s))$, soit W la variation de cette famille de géodésiques. C'est un champ de Jacobi de conditions initiales $W(0) = 0$ et

$$\nabla_T W(0) = \nabla_W T(0,0) = v'(0) = w$$

où T est le champ de vecteurs tangents à γ_0.
Comme

$$d_{rv}exp_p(w) = \frac{\partial}{\partial s}\gamma_s(r)|_{s=0} = W(r),$$

$$g_r(w) = exp_p^*(g_M)(v) = g_M(d_{rv}exp_p(w)) = |W(r)|^2.$$

Calculons les dérivées successives de la fonction $r \mapsto S(r) = |W(r)|^2$. On note $W' = \nabla_T W$, $W'' = \nabla_T\nabla_T W$, etc. les dérivées covariantes successives du champ W, de sorte que l'équation de Jacobi se lise $W'' = R(W,T)T$. Alors $S' = 2g(W', W)$, $S'' = 2gW'', W) + 2g(W', W')$, $S^{(3)} = 2g(W^{(3)}, W) + 6g(W'', W')$, $S^{(4)} = 2g(W^{(4)}, W)$+$8(W^{(3)}, W')$+6 g(W",W"). Il vient $S'(0) = 0$, $S''(0) = 2|w|^2$,

$$S^{(3)}(0) = 6g(W''(0), W'(0)) = -6g(R(W,T)T, w) = 0,$$

$$\begin{aligned} S^{(4)}(0) &= 8g(W^{(3)}(0), W'(0)) + 6g(W''(0), W''(0)) \\ &= 8g(W^{(3)}(0), W'(0)). \end{aligned}$$

car $W''(0) = (R(W,T)T)(0)$. Or

$$W^{(3)}(0) = -\nabla_T(R(W,T)T)(0)$$

$$= -((\nabla_T R)(W,T)T)(0) - (R(\nabla_T W,T)T)(0) - (R(W,\nabla_T T)T)(0) - (R(W,T)\nabla_T T)(0)$$

$$= R(W'(0),T(0))T(0) = -R(w,v)v,$$

d'où $S^{(4)}(0) = -8g(R(w,v)v,w)$. Avec la formule de Taylor, il vient

$$g_r(w) = S(r) = |w|^2 r^2 - \frac{1}{3} g(R(w,v)v,w) r^4 + o(r^4).$$

En coordonnées normales, la forme quadratique g_p à coefficients constants sur l'espace tangent s'écrit

$$g_p = dx_1^2 + ... + dx_n^2 = \sum_{i,j} \delta_{ij} dx_i dx_j.$$

Soit $x = \sum_k x_k \frac{\partial}{\partial x_k}$ un point de T_pM situé à distance r de l'origine, $x = rv$. Soit $y = \sum_i x_i \frac{\partial}{\partial x_i}$ un vecteur de T_pM, $y = rw$ où $w \in T_v S^{n-1}$. Alors

$$exp_p^* g_M(y) = |rw|^2 - \frac{1}{3} g(R(w,v)v,w) r^4 + o(r^4)$$

$$= g_p(y) - \frac{1}{3} \sum_{i,j,k,l} R_{iklj} y_i x_k x_l y_j + o(r^4),$$

ce qui s'écrit

$$exp_p^* g_M = \sum_{i,j} \delta_{ij} dx_i dx_j - \frac{1}{3} (\sum_{i,j,k,l} R_{iklj} x_k x_l + o(|x|^2)) dx_i dx_j.$$

Comme $g(R(y,x)x,y) = 0$ si y est colinéaire à x, et $exp_p^* g_M(y) = |y|^2$ dans ce cas, d'après le lemme de Gauss, la formule reste vraie même si y n'est pas orthogonal à x.
Si le terme $g(R(w,v)v,w)$ est positif, la métrique $exp_p^* g_M$ est plus petite que la métrique euclidienne, cela signifie que les géodésiques se rapprochent plus les unes des autres qu'en géométrie euclidienne.

1.5 Opérateurs différentiels

Soit $\mathcal{F}(M)$ l'algèbre de Lie des fonctions de classe C^∞ définies sur M et à valeurs réelles.
Le champ de vecteurs $grad\, f$ défini par :

$$g(X, grad\, f) = X(f) = df(X) \qquad \forall X \in \Gamma(M)$$

est appelé le gradient de f et noté ∇f.
Si on note les composantes locales de la différentielle df par $f_i = \frac{\partial f}{\partial x^i}$, alors les composantes locales de $grad\, f$ sont $f^i = g^{ij} f_j$.
La dérivée covariante seconde de f i.e $Hess\, f = \nabla(\nabla f) = \nabla df$ est appelée la hessienne de f. C'est le tenseur de type $(0, 2)$ défini par :

$$Hess\, f(X, Y) = g(\nabla_X grad\, f, Y) = \nabla_X(df)(Y)$$
$$= X(df(Y)) - df(\nabla_X Y) = X(Y(f)) - (\nabla_X Y)f \qquad \forall X, Y \in \Gamma(M).$$

Les composantes locales de $Hess\, f$ sont : $f_{jk} = \frac{\partial^2 f}{\partial x^j \partial x^k} - \Gamma^i_{jk} \frac{\partial f}{\partial x^i}$.
Nous rappelons que la hessienne d'une fonction vérifie les propriétés suivantes :
1- $Hess(f_1 f_2) = f_1 Hess\, f_2 + f_2 Hess\, f_1 + df_1 \otimes df_2$
2- $Hess\, \varphi \circ f = \varphi' \circ f Hess\, f + \varphi'' \circ f df \otimes df$,
$f_1, f_2 \in \mathcal{F}(M)$ et $\varphi : \mathbb{R} \to \mathbb{R}$.
L'opérateur $Hess : \mathcal{F}(M) \to \mathcal{F}^0_2(M)$, $f \mapsto Hess\, f$
s'appelle le hessien de f.
Si x_0 un point critique de f, alors $f_{jk}(x_0) = \frac{\partial^2 f}{\partial x^j \partial x^k}(x_0)$, et donc $Hess\, f$ ne dépend de la métrique qu'aux points qui ne sont pas critiques.

Chapitre 2

Convexité dans les variétés

Ce chapitre présente les notions de base et les théorèmes concernant la convexité d'une fonction réelle sur une variété munie d'une connexion.

Soit μ une mesure de probabilité à support compact dans $\mathbb{R}^n$. Si $x \in \mathbb{R}^n$ est le barycentre (au sens usuel) de μ, l'inégalité de Hölder affirme que $f(x) \leq \mu(f)$ pour toute fonction convexe f définie au moins sur l'enveloppe convexe K du support de μ et à valeurs réelles. (Ici et dans toute la suite, la notation $\mu(f)$ désigne l'intégrale $\int f \, \mathrm{d}\mu$, conformément au point de vue de l'analyse fonctionnelle selon lequel fonctions et mesures sont en dualité.)
Inversement, si un point x de K vérifie $f(x) \leq \mu(f)$ pour toute fonction convexe f sur K, cette inégalité devient une égalité lorsque f est affine (parce qu'alors f et $-f$ sont toutes deux convexes), et ceci signifie que x est le barycentre de μ.

Sur une variété riemannienne M (ou, plus généralement, une variété munie d'une connexion), il n'existe en général pas de fonctions affines non constantes, mais on peut définir des fonctions convexes, et, si la variété n'est pas trop grosse, les fonctions convexes sont suffisamment nombreuses pour permettre une définition raisonnable du barycentre de μ comme étant l'ensemble des points $x \in M$ tels que $f(x) \leq \mu(f)$ pour beaucoup de fonctions convexes. Ceci fera l'objet du chapitre suivant (chapitre 3) ; le présent chapitre a pour but de préparer tout cela par l'introduction des fonctions convexes sur M et l'étude de leurs propriétés.

2.1 Définitions et rappels

Soit M une variété réelle de dimension n, munie d'une connexion ∇ affine sans torsion. Nous supposons M et ∇ de classe C^∞. L'hypothèse d'absence

de torsion n'est pas une restriction, car nous n'utiliserons la connexion que par l'intermédiaire des géodésiques qu'elle permet de tracer sur M ; et, si l'on part d'une connexion générale, il suffit de la détordre, c'est-à-dire de lui retrancher la moitié de sa torsion, pour obtenir une connexion sans torsion ayant les mêmes géodésiques.
Nous n'imposons pas en général à ∇ d'être la connexion de Levi-Civita associée à une métrique riemanienne sur M. Toutefois, pour les calculs, il nous arrivera dans des démonstrations d'utiliser une métrique riemanienne auxiliaire sur M, la distance associée étant alors notée ρ.
Mais lorsqu'il sera spécifié que M est une variété riemanienne, ∇ sera toujours implicitement supposée être la connexion de Levi-Civita.

Définition 2.1.1. *Nous dirons qu'une partie A de M est ∇-convexe (ou convexe s'il n'y a pas d'ambiguïté sur la connexion) si, pour tous x et y dans A, il existe une géodésique $\gamma_{xy} : [0,1] \to M$ et une seule telle que $\gamma_{xy}(0) = x$, $\gamma_{xy}(1) = y$ et $\gamma([0,1]) \subset A$.*

Un exemple d'ensembles ∇-convexe dans une variété riemannienne est une boule géodésique ouverte

$$B(a,r) = \{x \in M : d(x,a) < r\}$$

telle que

$$r < min\{\frac{\pi}{4\sqrt{\Delta}}, \frac{inj\,(\overline{B}(a,r))}{2}\}$$

où Δ est la borne supérieure des courbures sectionnelles, $\overline{B}$ l'adhérence de B et $inj\,(\overline{B}(a,r)$ son rayon d'injectivité, et où $\frac{\pi}{4\sqrt{\Delta}}$ est pris égal à $+\infty$ quand $\Delta \leq 0$. Cheeger et Ebin montrent dans [14] que $B(a,\rho)$ est ∇-convexe.

Si une variété M de dimension n est elle-même convexe, elle est difféomorphe à un ouvert de $\mathbb{R}^n$ et il existe un voisinage U de la section nulle dans TM tel que :

$$(x,u) \in U \mapsto (x, exp_x u) \in M \times M$$

soit un difféomorphisme.
Toujours si M est convexe, pour x, y dans M, posons $\gamma(x,y) = \gamma_{xy}$, de sorte que $(\gamma(x,y)(s), 0 \leqslant s \leqslant 1)$ est la géodésique par rapport à ∇, dont les extrémités sont x et y. Soit

$$J(u,v)(s) = D(\gamma(.,.)(s))(u,v); \qquad u \in T_xM, \quad v \in T_yM$$

où $D(\gamma(s)(u,v))$ désigne la dérivée de $\gamma(s), s \in \mathbb{R}$ dans la direction de (u,v). L'application $(J(u,v)(s), 0 \leqslant s \leqslant 1)$ est le champ de Jacobi satisfaisant les conditions au bord $J(u,v)(0) = u$ et $J(u,v)(1) = v$. Notons

$$\dot{J}(u,v)(s) = (\nabla_{\frac{d}{ds}} J(u,v))(s)$$

pour tout $x, y \in M$. On peut définir une connexion ∇^c, dite relevée complète de ∇ (voir Milnor [51]) ; c'est une connexion sans torsion puisque ∇ l'est, elle est donc caractérisée par ses géodésiques, qui sont les champs de Jacobi de la connexion ∇.
La projection canonique

$$\pi : (TM, \nabla^c) \to (M, \nabla)$$

est affine ; on peut facilement voir que (TM, ∇^c) est convexe si (M, ∇) l'est et la seule géodésique joignant u et v dans TM est $(J(u,v)(s), 0 \leqslant s \leqslant 1)$.

2.2 Fonctions convexes sur les variétés

Définition 2.2.1. *Une fonction réelle f définie sur un ouvert O de M est dite ∇-convexe ou tout simplement convexe si, pour toute géodésique $\gamma : I \to O$ où I est un intervalle de $\mathbb{R}$, la fonction $f \circ \gamma$ est convexe sur I. La fonction f est dite strictement ∇-convexe (ou strictement convexe) si $f \circ \gamma$ est strictement convexe pour toute géodésique γ non constante dans O.*

On remarque que :
1 - Cette définition généralise aux variétés la notion de fonction convexe sur un ouvert d'un espace vectoriel ou affine.
2 - La ∇-convexité (des ensembles et des fonctions) est une notion intrinsèque, ne dépendant que de la connexion ∇ et non d'un système de coordonnées.
3 - Cette définition n'est intéressante que s'il existe suffisamment de géodésiques dans O, par exemple lorsque deux points sont toujours reliés par une géodésique. Dans le cas où O n'est pas connexe, elle n'a pas beaucoup de sens.
4 - Une fonction peut être convexe par rapport à une connexion mais pas par rapport à une autre.

Dans tout le chapitre, nous considérons une partie A de M, ∇-convexe et admettant un voisinage ouvert ∇-convexe.

Théorème 2.2.1. *Soit $f : A \to \mathbb{R}$ une fonction de classe C^2.
La fonction f est convexe si et seulement si, pour tous x et y dans A, la fonction $\varphi_{xy}(t) = f(\gamma_{xy}(t))$, $\quad t \in [0,1]$, vérifie*

$$\frac{d^2\varphi_{xy}}{dt^2} \geq 0 \qquad \forall t \in [0,1].$$

Si de plus

$$\frac{d^2\varphi_{xy}}{dt^2} > 0$$

pour tout t et tout couple (x,y) de points de A tels que $x \neq y$, alors f est strictement convexe.

Preuve
On a $d^2\varphi_{xy}/dt^2 \geq 0$ pour tout t si et seulement si la fonction d'une variable φ_{xy} est convexe. Par ailleurs, γ_{xy} décrit toutes les géodésiques de A lorsque x et y varient dans A. La positivité ci-dessus équivaut donc à la convexité de $f \circ \gamma$ pour toute géodésique γ, c'est-à-dire à la convexité de f.
Le cas de la convexité stricte est analogue.

Théorème 2.2.2. *Soit $f : A \to \mathbb{R}$ une fonction de classe C^2. La fonction f est convexe si et seulement si la forme quadratique $Hess\, f$ est partout positive.*

Si de plus $Hess\, f$ est partout définie positive, f est strictement convexe.

Preuve
Conservons la notation $\varphi_{xy}(t) = f(\gamma_{xy}(t))$ du théorème précédent. Puisque $\nabla_{\dot\gamma_{xy}}\dot\gamma_{xy} = 0$, on trouve

$$\frac{d^2\varphi_{xy}}{dt^2} = \dot\gamma_{xy}(\dot\gamma_{xy}(f))(\gamma_{xy}(t)) = \nabla_{\dot\gamma_{xy}}(df)(\dot\gamma_{xy}) = Hess\, f(\dot\gamma_{xy}, \dot\gamma_{xy})$$

Le théorème (2.2.1) montre alors que f est convexe si et seulement si $Hess\, f \geq 0$ et strictement convexe si $Hess\, f > 0$. $\diamond$

Les fonctions réelles f de classe C^2 sur M telles que $Hess\, f = 0$ sont dites affines ; les deux fonctions f et $-f$ sont alors convexes, et $f \circ \gamma$ est affine (au sens ordinaire) pour toute géodésique γ. Nous verrons avec le théorème 2.3.2 que peu de variétés admettent des fonctions affines non constantes.

Nous allons donner quelques propriétés des fonctions convexes ; mais auparavant, voici un exemple.

Exemple dans le plan de Poincaré

L'ensemble $H = \{(x,y) \in \mathbb{R}^2 | y > 0\}$ muni de la métrique riemannienne $g_{ij}(x,y) = \frac{1}{y^2}\delta_{ij}, i,j = 1,2$, s'appelle le plan de Poincaré (ou plan hyperbolique).

Les composantes de la connexion riemannienne sont

$$\Gamma^1_{11} = \Gamma^1_{22} = \Gamma^2_{12} = \Gamma^2_{21} = 0, \quad \Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -\frac{1}{y}, \quad \Gamma^2_{11} = \frac{1}{y}$$

Si $f : H \to \mathbb{R}$ est de classe C^2, la hessienne hyperbolique de f a les composantes

$$f_{11} = \frac{\partial^2 f}{\partial x^2} - \frac{1}{y}\frac{\partial f}{\partial y}, \qquad f_{12} = \frac{\partial^2 f}{\partial x \partial y} + \frac{1}{y}\frac{\partial f}{\partial x}, \qquad f_{22} = \frac{\partial^2 f}{\partial y^2} + \frac{1}{y}\frac{\partial f}{\partial y}.$$

Les géodésiques du plan de Poincaré sont les demi-droites $C_a : x = a, y > 0$ et les demi-cercles $C_{b,r} : (x-b)^2 + y^2 = r^2$, elle admettent les paramétrisations suivantes

$$C_{x_0} : x = x_0,\ y = y_0 e^t, \quad t \in]-\infty, +\infty[$$

$$C_{b,r} : x = b - r\tanh t,\ y = \frac{r}{\cosh t}, \quad t \in]-\infty, +\infty[.$$

La distance hyperbolique entre deux points $P_1(x_1, y_1)$ et $P_2(x_2, y_2)$ est

$$d_H(P_1, P_2) = \begin{cases} |\ln \frac{y_2}{y_1}| \text{ pour } x_1 = x_2, \\ |\ln \frac{x_1-b+r}{x_2-b+r}\frac{y_2}{y_1}| \text{ pour } P_1, P_2 \in C_{b,r}. \end{cases}$$

Le plan de Poincaré H est convexe et complet (pour tous points P_1 et P_2 la géodésique $\gamma_{P_1P_2}$ existe et est unique et peut être définie sur $\mathbb{R}$ tout entier).

La courbure du plan de Poincaré est constante, partout égale à -1. A isométrie près, H est la seule variété riemannienne de dimension 2, homéomorphe à $\mathbb{R}^2$, complète et à courbure constante -1. Les isométries opèrent transitivement, et le groupe des isométries fixant un point est tout le groupe orthogonal $SO(2)$. En tout point O, la fonction exponentielle exp_O est un difféomorphisme C^∞ entre T_OH et H.

Un autre système de coordonnées dans H est les coordonnées normales polaires. On fixe un point $O \in H$, pris comme origine, et on appelle $r(P)$ la distance $d(O,P)$ et $\theta(P)$ l'angle orienté fait dans T_OH par la géodésique γ_{OP} avec une direction fixe. Le point O est une singularité de ces coordonnées. Les coordonnées normales polaires d'un point $exp_O(v)$ sont simplement les coordonnées polaires usuelles de v dans l'espace euclidien $T_O H$. Puisque dans H la longueur des cercles de rayon r est $2\pi \operatorname{sh} r$, la métrique est donnée par

$$g_{rr} = 1, \qquad g_{r\theta} = g_{\theta r} = 0, \qquad g_{\theta\theta} = \operatorname{sh}^2 r$$

(les indices sont notés r et θ au lieu de 1 et 2 pour éviter la confusion avec les coordonnées x et y introduites plus haut). Les symboles de Christoffel se calculent très facilement :

$$\Gamma^r_{\theta\theta} = -\operatorname{sh} r \operatorname{ch} r, \qquad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{\operatorname{ch} r}{\operatorname{sh} r}, \qquad \text{les autres } \Gamma^i_{jk} \text{ sont nuls.}$$

Proposition 2.2.3. *Soit γ une droite (une géodésique complète) dans H ; elle partage H en deux demi-espaces h et k.*

La fonction $f : H \to \mathbb{R}$ nulle sur k et égale sur h à la distance à la droite γ, est convexe sur H. En outre, sa restriction à l'intérieur de h est C^∞, et affine sur chaque droite coupant γ à angle droit ; mais $Hessf$ est strictement positive dans la direction orthogonale à une telle droite.

La fonction f^2 (convexe puisque f est convexe et positive) est strictement convexe dans l'intérieur de h.

La fonction égale sur H à la distance à la droite γ, est convexe sur H ; son carré est strictement convexe sur les intérieurs de h et de k.

Preuve :
On choisit un système de coordonnées normales polaires (r, θ) tel que γ ait pour équation $\theta = 0 \bmod \pi$ et que l'intérieur de h soit obtenu pour $0 < \theta < \pi$. Un argument de trigonométrie hyperbolique montre que, dans h, la distance $f(r, \theta)$ à γ est donnée par

$$\operatorname{sh} f(r, \theta) = \operatorname{sh} r \sin\theta \ .$$

Cette formule montre que f est C^∞ dans l'intérieur de h.

En dérivant deux fois les deux membres par rapport à r et θ, on trouve

$$\frac{\partial f}{\partial r} = \frac{\operatorname{ch} r \sin\theta}{\operatorname{ch} f}, \qquad \frac{\partial f}{\partial \theta} = \frac{\operatorname{sh} r \cos\theta}{\operatorname{ch} f},$$

$$\frac{\partial^2 f}{\partial r^2} = \frac{\operatorname{sh} f \, \cos^2\theta}{\operatorname{ch}^3 f}, \qquad \frac{\partial^2 f}{\partial r \, \partial\theta} = \frac{\operatorname{ch} r \cos\theta}{\operatorname{ch}^3 f}, \qquad \frac{\partial^2 f}{\partial \theta^2} = -\frac{\operatorname{sh} f \, \operatorname{ch}^2 r}{\operatorname{ch}^3 f}.$$

Compte tenu de l'expression des symboles de Christoffel, on a

$$Hessf = \frac{\operatorname{sh} f}{\operatorname{ch}^3 f} \begin{pmatrix} \cos^2\theta & -\operatorname{ch} r \, \operatorname{sh} f \, \cos\theta \\ -\operatorname{ch} r \, \operatorname{sh} f \, \cos\theta & \operatorname{ch}^2 r \, \operatorname{sh}^2 f \end{pmatrix} .$$

C'est une matrice positive, mais elle n'est pas définie positive puisque l'on a $Hessf(\nabla f, \nabla f) = 0$, le gradient ∇f étant le vecteur unitaire de composantes

$$\nabla_r f = \frac{\operatorname{ch} r \sin\theta}{\operatorname{ch} f}, \qquad \nabla_\theta f = \frac{\cos\theta}{\operatorname{sh} r \operatorname{ch} f}.$$

La fonction $f : H \to \mathbb{R}$ est donc positive, continue, et convexe sur l'ouvert où elle est strictement positive. Ceci entraîne qu'elle est convexe (car les fonctions d'une variable ayant cette propriété sont convexes).

En échangeant les rôles de h et k, la fonction $\hat{f}$ nulle sur h et égale dans k à la distance à γ est elle aussi convexe ; la fonction distance à γ est convexe car c'est le maximum des deux fonctions convexes f et $\hat{f}$ (cet argument sera formalisé dans le corollaire 2.2.10).

Enfin, la fonction f^2 est strictement convexe dans l'intérieur de h car $Hess(f^2) = 2f\, Hessf + df \otimes df$ y est définie positive, puisque $df \otimes df > 0$ dans la direction de ∇f, qui est la seule direction où $Hessf$ s'annule. $\square$

Voici l'analogue en courbure constante positive de la proposition 2.2.3. Prenons comme variété M la sphère de rayon 1 ; la courbure est constante et vaut $+1$. La variété entière n'est pas convexe (deux points diamétralement opposés sont joints par une infinité de géodésiques) mais les hémisphères ouverts sont convexes. Sur un hémisphère ouvert h, centré en un pôle Q et limité par un grand cercle γ, la fonction $f =$ (distance à Q) est convexe ; de façon équivalente, $f - \frac{\pi}{2} = -$(distance à γ) est convexe. Ces fonctions sont encore convexes sur le compact (non convexe) $\bar{h}$, mais, contrairement au cas hyperbolique, elles ne peuvent être prolongées en fonctions convexes sur un voisinage de $\bar{h}$.

Nous revenons maintenant aux propriétés des fonctions convexes.

Théorème 2.2.4. *Soit $f : A \to \mathbb{R}$ une fonction convexe, I est un ensemble convexe dans $\mathbb{R}$ contenant $f(A)$ et $\varphi : I \to \mathbb{R}$ est une fonction convexe croissante, alors $\varphi \circ f$ est convexe sur A.*

Preuve

Cela résulte immédiatement de la même propriété pour les fonctions convexes d'une variable définies sur un intervalle réel.

Le théorème suivant montre que l'ensemble des fonctions convexes est fermé par rapport aux combinaisons linéaires positives et aux limites simples.

Théorème 2.2.5. *Si f_i, $i = 1, 2, ..., m$ sont des fonctions convexes sur $A \subset M$ et $c_i \geq 0$, alors $\Sigma_i c_i f_i$ est convexe sur A.*

Si une suite de fonctions convexes sur A converge point par point vers une fonction limite, celle-ci est aussi convexe.

Preuve
Vraies pour les fonctions convexes d'une variable, ces propriétés se transfèrent aussitôt aux fonctions convexes sur A.

Théorème 2.2.6. *Si f est une fonction convexe sur A et c est un nombre réel, alors l'ensemble $A^c = \{z \backslash z \in A, f(z) \leq c\}$ est un sous-ensemble ∇-convexe de A.*

Preuve
Soient x et y deux points de A^c ; on a $f(x) \leq c$ et $f(y) \leq c$. Soit $z = \gamma_{xy}(t)$ un point arbitraire sur la géodésique joignant x à y. Comme f est convexe on trouve

$$f(z) = f(\gamma_{xy}(t)) \leq (1-t)f(x) + tf(y) \leq (1-t)c + tc = c,$$

et z est aussi dans A^c.

Corollaire 2.2.7. *Soient $f_i, i = 1, 2, ..., m$ des fonctions convexes sur A et a_i des nombres réels. Le sous-ensemble*

$$\{z \backslash z \in A, f_i(z) \leq a_i, i = 1, 2, ..., m\}$$

est ∇-convexe dans A.

Preuve
Cela résulte du théorème précédent et du fait qu'une intersection d'ensembles ∇-convexe est ∇-convexe.

Proposition 2.2.8. *Si la variété M est convexe et ∇-complète (c'est-à-dire que toute géodésique $\gamma : I \to M$ peut être prolongée en une géodésique définie sur $\mathbb{R}$ tout entier), toute fonction convexe sur M et majorée est constante.*

Preuve
Soit f une telle fonction et x et y deux points de M. Il existe une géodésique $\gamma : \mathbb{R} \to M$ telle que $\gamma(0) = x$ et $\gamma(1) = y$. La fonction $f \circ \gamma$ est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ convexe et majorée, donc constante ; ceci entraîne que $f(x) = f(y)$.

Si deux variétés M' et M'' sont respectivement munies de connexions ∇' et ∇'', la variété produit $M = M' \times M''$ est munie de la connexion produit ∇, telle qu'une courbe $\gamma : I \to M$ est une ∇-géodésique si et seulement si les deux projections γ' et γ'' de γ sont des géodésiques respectivement pour ∇' et ∇''. Si de plus ∇' et ∇'' sont les connexions de Levi-Civita pour des métriques riemanniennes g' et g'' sur M' et M'', alors ∇ est la connexion de Levi-Civita associée à la métrique riemannienne produit $g' \oplus g''$ sur M.

En particulier, si M est une variété munie d'une connexion, on peut munir la variété produit $M \times \mathbb{R}$ de la connexion produit de la connexion de M par la connexion canonique de $\mathbb{R}$. Et une partie A de M est convexe si et seulement si le produit $A \times \mathbb{R}$ est convexe dans $M \times \mathbb{R}$.

Dans une carte, si $(x_1, ..., x_n)$ sont des coordonnées locales sur M, alors $(x_1, ..., x_n, t)$ sont des coordonnées locales sur $M \times \mathbb{R}$; et dans cette carte, les symboles de Christoffel $\overline{\Gamma}^{\alpha}_{\beta\gamma}$, $\alpha, \beta, \gamma = 1, 2, \cdots, n, n+1$ de la connexion produit sont donnés par $\overline{\Gamma}^{\alpha}_{\beta\gamma} = 0$ lorsque α, β ou γ vaut $n+1$, et $\overline{\Gamma}^{i}_{jk} = \Gamma^{i}_{jk}$ pour $i, j, k = 1, \cdots, n$, où Γ^{i}_{jk} sont les symboles de Christoffel de la connexion de M dans la carte $(x_1, ..., x_n)$.

Les fonctions convexes sont liées aux ensembles convexes :

Théorème 2.2.9. *Une fonction $f : A \to \mathbb{R}$ est convexe si et seulement si son épigraphe*

$$E(f) = \{(x, u)/f(x) \leqslant u\} \subset A \times \mathbb{R}$$

est convexe dans le produit $M \times \mathbb{R}$..

Preuve :

Supposons que f est convexe. Soit $(x, u) \in E(f)$, $(y, v) \in E(f)$. On a $u \geqslant f(x)$ et $v \geqslant f(y)$ et par suite

$$(1-t)u + tv \geqslant (1-t)f(x) + tf(y) \geqslant f(\gamma_{xy}(t)),$$

d'où $(\gamma_{xy}(t), (1-t)u + tv) \in E(f)$. Comme $t \mapsto (\gamma_{xy}(t), (1-t)u + tv)$ est la géodésique joignant (x, u) à (y, v) dans $M \times \mathbb{R}$, ceci montre que $E(f)$ est ∇-convexe.

Supposons maintenant que $E(f)$ est ∇-convexe. Pour x et y dans A, on a $(x, f(x)) \in E(f)$ et $(y, f(y)) \in E(f)$, d'où

$$(\gamma_{xy}(t), (1-t)f(x) + tf(y)) \in E(f),$$

i.e. $f(\gamma_{xy}(t)) \leqslant (1-t)f(x) + tf(y)$ et par suite f est convexe.

Corollaire 2.2.10. *Soit A_i des sous-ensembles convexes de M tels que*

$$A = \bigcap_{i \in I} A_i \neq \emptyset$$

et soit $f_i : A_i \to \mathbb{R}$ des fonctions convexes. On définit $f : A \to \mathbb{R} \cup \{+\infty\}$ par $f(x) = sup_{i \in I} f_i(x)$ et $S = \{x \in A/f(x) < \infty\}$. Alors S est ∇-convexe et f est convexe sur S.

Théorème 2.2.11. *Si F est un sous-ensemble convexe de $M \times \mathbb{R}$, ses deux projections $\pi_1 F$ et $\pi_2 F$ sur M et $\mathbb{R}$ sont convexes. Si de plus F est non vide et $\pi_2 F$ est bornée inférieurement, alors la fonction $f(x) = inf\{u/(x,u) \in F\}$ est convexe sur $\pi_1 F$.*

Preuve :
La convexité des projections découle directement du fait que les projections d'un segment géodésique sont des segments géodésiques.

Soit $x \in \pi_1 F$; par définition de f, il existe une suite (u_i) qui tend vers $f(x)$ et telle que chaque (x, u_i) est dans F. De même, pour $y \in \pi_1 F$, il existe des v_i tendant vers $f(y)$ et tels que $(y, v_i) \in F$. Pour $0 \leq t \leq 1$, la convexité de F entraîne que $(\gamma_{xy}(t), (1-t)u_i + tv_i)$ est dans F ; par définition de f, on en déduit $f(\gamma_{xy}(t)) \leq (1-t)u_i + tv_i$. Faisant tendre i vers l'infini, on obtient $f(\gamma_{xy}(t)) \leq (1-t)f(x) + tf(y)$, et f est convexe.

Nous plaçant maintenant dans le cadre d'une variété riemannienne, nous allons nous intéresser à la convexité de petites boules.

Lemme 2.2.12. *Soit x_0 un point d'une variété riemannienne M. Il existe $b > 0$ tel que, pour tout $r \in]0, b[$, pour tout x vérifiant $d(x_0, x) = r$ et pour toute géodésique γ incluse dans $B(x_0, b)$ et tangente à la sphère $S(x_0, r)$ en x, l'ensemble $\gamma \setminus \{x\}$ est extérieur à $\overline{B(x_0, r)}$.*

Preuve :
Si $(\xi^1, \cdots, \xi^n)$ est un système de coordonnées normales autour du point x_0 alors $\xi(x_0) = 0$ et

$$B(x_0, r) = \{x \in M / \sum_{i=1}^{n} (\xi^i(x))^2 < r^2\}.$$

Soit $\xi^i = \xi^i(t), i = 1, \cdots, n$, $t \in I$ les équations paramétriques d'une géodésique tangente à $S(x_0, r)$ au point $x = (\xi^1(0), \cdots, \xi^n(0))$ avec $0 \in I$.
On considère la fonction :

$$F(t) = \sum_{i=1}^{n} (\xi^i(t))^2, \qquad t \in I.$$

Par les conditions de tangence, on a :

$$F(0) = r^2, \qquad\qquad 0 = \sum_{i=1}^{n} \xi^i(0) \frac{d\xi^i}{dt} = \frac{1}{2} \frac{dF}{dt}(0);$$

ainsi $t = 0$ est un point critique de F.
En tenant compte des équations différentielles des géodésiques, nous trouvons :

$$\frac{1}{2}\frac{d^2F}{dt^2} = \sum_{j,k=1}^{n}(\delta_{jk} + \sum_{i=1}^{n}\Gamma^i_{jk}(\xi(t))\frac{d\xi^j}{dt}\frac{d\xi^k}{dt}.$$

Puisque $\Gamma^i_{jk}(x_0) = 0$, il existe $b > 0$ tel que la matrice

$$\left[\delta_{jk} + \sum_{i=1}^{n}\Gamma^i_{jk}\right]$$

est définie positive sur $B(x_0, b)$. Si $0 < r < b$, il existe un voisinage $J_0 \subset I$ tel que

$$\frac{d^2F}{dt^2} > 0, \qquad \forall t \in J_0$$

Et par suite $F(t) > F(0) = r^2, \forall t \in J_0 - \{0\}$, et donc il existe un arc de la géodésique qui se trouve à l'extérieur de $B(x_0, r)$.

Théorème 2.2.13. *Chaque point $x_0 \in M$ a un voisinage normal convexe sphérique $B(x_0, r)$.*

Preuve :
D'après la remarque (1.2.1), chaque point x_0 a un voisinage normal convexe W_{x_0}.
Il existe $B(x_0, r) \subset W_{x_0}$ et chaque deux points $x, y \in B(x_0, r)$ peuvent être joints par une géodésique unique $\gamma_{xy}(t), t \in [0, 1]$. Montrons que γ_{xy} est incluse dans $B(x_0, r)$. Pour cela on suppose que γ_{xy} est donnée par les équations paramétriques

$$\xi^i = \xi^i(t), i = 1, \cdots, n$$

et on considère la fonction :

$$F(t) = \sum_{i=1}^{n}(\xi^i(t))^2, t \in [0, 1]$$

$F(0) < r^2, F(1) < r^2$. Si γ_{xy} a un point à l'extérieur de $B(x_0, r)$, alors le maximum s^2 de F sur $[0, 1]$, atteint pour un $t_0 \in]0, 1[$, vérifiera $s^2 \geq r^2$. Puisque F a un maximum local en t_0,

$$0 = \frac{dF}{dt}(t_0) = 2\sum_{i=1}^{n}\xi^i(t_0)\frac{d\xi^i}{dt}(t_0),$$

et par suite γ_{xy} est tangente à la sphère $S(x_0, s)$ en $z = (\xi^1(t_0), \cdots, \xi^n(t_0))$. Mais le lemme précédent implique $F(t) > s^2$ pour $t \neq t_0$, t au voisinage de t_0, ce qui contredit la maximalité de F en t_0.

2.3 Changement de métrique et préservation de la convexité

Dans cette section, on s'intéresse à des changements de la métrique riemannienne g qui préserve la convexité d'une fonction réelle définie sur M.
Soit $f : M \to \mathbb{R}$ une fonction de classe C^∞.
Considérons la métrique

$$\widetilde{g} = g + df \otimes df$$

Théorème 2.3.1. *f est convexe par rapport à $\widetilde{g}$ si et seulement elle l'est par rapport à g.*

Preuve
Soit g_{ij}, $i, j = 1, ..., n$ les composantes locales de g. Un calcul direct donne

$$\widetilde{g}^{ij} = g^{ij} - (1 + \|\nabla f\|^2)^{-1} f^i f^j,$$

$$\widetilde{\Gamma}^i_{jk} = \Gamma^i_{jk} + (1 + \|\nabla f\|^2)^{-1} f^i f_{jk}, \qquad i, j, k = 1, ..., n,$$

où

$$f_r = \frac{\partial f}{\partial x_r}, \quad f^i = g^{ir} f_r, \quad \|\nabla f\|^2 = g^{ij} f_i f_j, \qquad f_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_k} - \Gamma^i_{jk} f_i.$$

On obtient

$$\widetilde{f}_{jk} = \frac{1}{1 + \|\nabla f\|^2} f_{jk}.$$

On remarque que si on prend une autre fonction de classe C^∞, $\varphi : M \to \mathbb{R}$, alors

$$Hess_{\widetilde{g}} \varphi = Hess_g \varphi - \frac{g(\nabla f, \nabla \varphi)}{1 + \|\nabla f\|^2} Hess_g f.$$

Remarque 2.3.1. *Soit $f : M \to \mathbb{R}$ une fonction convexe de classe C^2. Le changement conforme*

$$g \to \widetilde{g} = e^{2f} g$$

ne préserve pas toujours la convexité de f

En effet, la relation

$$\widetilde{\Gamma}^i_{jk} = \Gamma^i_{jk} + f_k\delta^i_j + f_j\delta^i_k - f^i g_{jk}$$

implique

$$\widetilde{f}_{jk} = f_{jk} - 2f_j f_k + g_{jk} f^h f_h. \qquad \square$$

Dans le cas où A est une sous-variété de M et $Hess_g f$ est définie positive, alors $(A, Hess_g f)$ est une nouvelle variété riemanienne.
Rappelons que les fonctions de hessienne nulle sont les fonctions affines ; et on a le résultat suivant :

Théorème 2.3.2. *Une variété riemanienne (M, g) de dimension n est produit d'une variété riemanienne de dimension $(n - p + 1)$ et de l'espace euclidien $\mathbb{R}^{p-1}$ (au moins localement) si et seulement si l'espace vectoriel de toutes les fonctions affines sur M est de dimension au moins p.*

Preuve
Soit (U, x^i) un voisinage de coordonnées et Γ^h_{ij}, $i, j, h = 1, ..., n$, les composantes de la connexion riemannienne déterminées par les composantes g_{ij} de la métrique g.
Supposons que

$$f_0 = const \neq 0, f_1, ..., f_{p-1}$$

sont des fonctions affine sur M linéairement indépendantes, on a

$$\frac{\partial^2 f_\alpha}{\partial x^i \partial x^j} - \Gamma^h_{ij} \frac{\partial f_\alpha}{\partial x^h} = 0 \qquad \text{pour} \quad \alpha = 0, 1, ..., p - 1.$$

Cette hypothèse est equivalente au fait que

$$grad\, f_1, ..., grad\, f_{p-1}$$

sont des champs de vecteurs parallèles non nuls. Par un changement de coordonnées [21] il est prouvé que la métrique $g = g_{ij} dx^i dx^j$ peut être écrite

$$g = g_{\alpha'\beta'} dx^{\alpha'} dx^{\beta'} + dx^{(n-p+2)'} \otimes dx^{(n-p+2)'} + ... + dx^{n'} \otimes dx^{n'},$$

$$\alpha', \beta' = 1, ..., n - p + 1.$$

(M, g) est donc, le produit riemannien entre

$$(M_1, g_1 = g_{\alpha'\beta'} dx^{\alpha'} dx^{\beta'})$$

et

$$(\mathbb{R}^{p-1}, g_2 = dx^{(n-p+2)'} \otimes dx^{(n-p+2)'} + ... + dx^{n'} \otimes dx^{n'}).$$

Réciproquement, supposons que M est le produit de telles variétés.
Les composantes $\Gamma^{h'}_{i'j'}$, $i', j', h' = 1, ..., n$ de la connexion riemannienne déterminée par $g = g_1 + g_2$ sont nuls sauf peut-être $\Gamma^{\alpha'}_{\beta'\gamma'} = \overline{\Gamma}^{\alpha'}_{\beta'\gamma'}, \alpha'.\beta'.\gamma' = 1, ..., n-p+1$, où $\overline{\Gamma}^{\alpha'}_{\beta'\gamma'}$ sont les symboles de Christoffel de g_1.
Dans ce cas les fonctions de coordonnées

$$x^{(n-p+2)'}, ..., x^{n'}$$

sont des fonctions affines, et en prenant en compte qu'une fonction constante non nulle est affine, il s'ensuit que l'espace des fonctions affines sur M est de dimension au moins p. $\square$

Revenons au cas d'une variété M sans structure riemannienne, mais pourvue d'une connexion ∇. Deux points assez voisins dans M sont reliés par une géodésique, qui dépend de façon C^∞ des deux points : nous appellerons γ la fonction, définie et C^∞ dans un ouvert de $M \times M \times \mathbb{R}$ contenant $M \times M \times [0,1]$ et à valeurs dans M, telle que $\gamma(x, y, t)$ soit égal à $\gamma_{x,y}(t)$, la géodésique joignant x à y.

Lemme 2.3.3. *[24]*
Soient a un point de M et V un voisinage de a ∇-convexe et relativement compact dans le domaine d'une carte locale et tel que $V \times V \times [0,1]$ soit relativement compact dans le domaine de γ. Après identification par la carte de V à une partie de $\mathbb{R}^n$, il existe c (dépendant de la carte) tel que, pour $(x, y, t) \in V \times V \times [0,1]$, on ait

$$\|\gamma_{xy}(t) - (1-t)x - ty\| \leq 2ct(1-t)\|x-y\|^2.$$

Dans cette formule, $\gamma_{xy}(t)$ fait intervenir la connexion et non la carte, $(1-t)x + ty$ la carte et non la connexion, et $\| \, \|$ est la norme euclidienne calculée dans $\mathbb{R}^n$ au moyen de la carte.

Preuve :
Posons $z = y - x$ et $\overline{\gamma}_{xz}(t) = \gamma_{xy}(t) - x - tz$.
Il faut montrer que $\|\overline{\gamma}_{xz}(t)\| \leq 2ct(1-t)\|z\|^2$.
Montrons cette inégalité pour chaque composante $\overline{\gamma}^i_{xz}$.
Soit g la géodésique qui, en $s = 0$ passe par x et avec le vecteur tangent $\frac{\partial}{\partial x^j}$, ($j$-ième vecteur de base de T_xM), on a donc $\gamma_{xg(s)}(t) = g(ts)$.
En dérivant par rapport à s, on trouve

$$\frac{\partial \gamma^i}{\partial y^k}(x, g(s), t)\dot{g}^k(t) = t\dot{g}^i(ts).$$

En $t = 0$, on a $\frac{\partial \gamma^i}{\partial y^j}(x, x, t) = t\delta_j^i$ car $\dot{g}^k(0) = \delta_j^k$, ce qui est équivalent à

$$\frac{\partial \overline{\gamma}^i}{\partial z^j}(x, 0, t) = 0 \; ; \tag{2.3.1}$$

une dérivation par rapport à t, donne :

$$\frac{\partial^2 \overline{\gamma}^i}{\partial t \partial z^j}(x, 0, t) = 0. \tag{2.3.2}$$

Or $\overline{\gamma}^i(x, 0, t) = 0$, donc $\frac{\partial \overline{\gamma}^i}{\partial t}(x, 0, t) = 0$.
Les relations (2.3.1) et (2.3.2) montrent que la fonction $z \mapsto \frac{\partial \overline{\gamma}^i}{\partial t}(x, z, t)$ et ses dérivées d'ordre un sont nulles en $z = 0$; et puisque le triplet (x, z, t) varie dans un compact par hypothèse, on déduit $|\frac{\partial \overline{\gamma}^i}{\partial t}(x, z, t)| \leq c^i \|z\|^2$.
En remarquant que $\overline{\gamma}(x, z, t) = 0$ pour $t \in \{0, 1\}$, on tire

$$|\overline{\gamma}^i(x, z, t)| \leq c^i \inf(t, 1-t)\|z\|^2 \leq c^i 2t(1-t)\|z\|^2. \qquad \square$$

Pour le lemme suivant, on définit des fonctions w_1, w_2, etc. par

$$w_1(x_0, x_1, t) = \gamma(x_0, x_1, t);$$

$$w_k(x_0, x_1, \ldots, x_k; t_1, \ldots, t_k) = \gamma(w_{k-1}(x_0, x_1, \ldots, x_{k-1}; t_1, \ldots, t_{k-1}), x_k, t_k).$$

Comme conséquence du lemme 2.3.3, on a le lemme suivant :

Lemme 2.3.4. *[24]*
Soient a et W comme dans le lemme (2.3.3). Il existe $n+1$ points $x_0, ..., x_n$ de W tels que le "simplexe"

$$\{w_n(x_0, ..., x_n; t_1, ..., t_n) : t_1, ..., t_n \in [0, 1]\}$$

obtenu par interpolation géodésique entre ces points, soit un voisinage de a.

Proposition 2.3.5. *[24] Soit f une fonction ∇-convexe sur M.*
a) Toute représentation locale de f est continue et même lipschitzienne.
b) Pour $a \in M$ et $u \in T_aM$, soit $\delta f_a(u)$ la dérivée à droite en 0 de la fonction convexe d'une variable $t \mapsto f(exp_a(tu))$. La fonction positivement homogène δf_a est convexe sur l'espace vectoriel T_aM.

La fonction δf définie sur la variété tangente TM n'est autre que la différentielle de f quand f est de classe C^1, dans ce cas δf_a est linéaire.
Preuve de (a)
Soit a un point de M et V un voisinage de a vérifiant les hypothèses du

lemme (2.3.3) et vu à travers une carte normale en a comme une boule de $\mathbb{R}^n$. (Dans une carte normale, les boules de rayon assez petit sont ∇-convexe [37] p.35 ; et c'est encore vrai pour une carte quelconque d'après le lemme 2.3.3). Nous identifions V à la boule ; le point a devient maintenant O.
D'après le lemme (2.3.4), il existe un voisinage de O, qu'on peut choisir symétrique, inclus dans un "simplexe". Pour tout point x du voisinage, on a $\gamma(x, -x, \frac{1}{2}) = O$.
La fonction f étant convexe, on a $f(\gamma_{xy}(t)) \leq sup(f(x), f(y))$ sur tout segment géodésique reliant deux sommets x et y ; ceci entraîne la majoration $f \leq sup\{f(x); x \text{ sommet du simplexe}\}$ sur le simplexe, et a fortiori au voisinage de O.
D'aure part, la convexité de f donne aussi $f(\gamma(x, -x, \frac{1}{2})) \leq \frac{1}{2}(f(x) + f(-x))$ c'est à dire $f(x) \geq 2f(O) - f(-x) \geq 2f(O) - sup(f)$ et f est aussi minorée au voisinage de O ; f est donc bornée sur un voisinage de O.
Ce qui précède montre l'existence de $\varepsilon > 0$ tel que la boule $\overline{B}(\varepsilon)$ de $\mathbb{R}^n$ soit ∇-convexe, incluse dans V et que f soit bornée sur cette boule.
On considère maintenant deux suites (x_n) et (y_n) dans $B(\varepsilon)$ tendant toutes deux vers O et telles que $x_n \neq y_n$.
Pour montrer que f est lipschitzienne au voisinage de O, il suffit de prouver que la suite $\frac{(f(y_n) - f(x_n))}{\|y_n - x_n\|}$ contient une sous-suite bornée.
On remarque que lorsque ε tend vers zéro, l'inf des rayons de courbure des géodésiques contenues dans $B(\varepsilon)$ tend vers l'infini, on peut donc prolonger la géodésique reliant x_n et y_n au delà de y_n jusqu'a rencontrer la frontière $\partial B(\varepsilon)$ en un point z_n.
Il existe donc une suite de réels $t_n \in]0, 1[$ tels que $y_n = \gamma_{x_n z_n}(t_n)$. Quitte à extraire une sous suite, on peut supposer que la suite de points z_n tend vers une limite $z \in \partial B(\varepsilon)$ et la suite des réels t_n vers un réel $t \in [0, 1]$.
Par passage à la limite ; $\gamma_{Oz}(t) = O$, ce qui implique $t = 0$.
Puisque f est convexe, on a :

$$\frac{f(z_n) - f(y_n)}{1 - t_n} \geq \frac{f(y_n) - f(x_n)}{t_n} = \frac{f(y_n) - f(x_n)}{\|y_n - x_n\|} \frac{\|y_n - x_n\|}{t_n}$$

Le premier membre est borné, car f est bornée sur $\overline{B}(\varepsilon)$ et t_n tend vers zéro.
D'autre part,

$$\frac{y_n - x_n}{t_n} = \frac{1}{t_n}(\gamma_{x_n z_n}(t_n) - \gamma_{x_n z_n}(0)) = \frac{1}{t_n} \int_0^{t_n} \frac{\partial \gamma_{x_n z_n}(t)}{\partial t} dt$$

tend vers $\frac{\partial \gamma_{Oz}(0)}{\partial t} \neq 0$, donc $\lim_{n \to +\infty} \frac{\|y_n - x_n\|}{t_n} > 0$. Ceci prouve que $\frac{f(y_n) - f(x_n)}{\|y_n - x_n\|}$ reste bornée.

Preuve de (b)
On se place dans une carte normale en a. Soient u et v dans T_aM et $\lambda \in [0,1]$; posons $r = (1-\lambda)u + \lambda v$. Il faut montrer que $(1-\lambda)\lambda\delta f(u) + \lambda\delta f(v) - \delta f(r) \geq 0$. Le premier membre est la limite, quand t tend vers zéro par valeurs positives, de

$$\frac{1}{t}\left[(1-\lambda)f(tu) + \lambda f(tv) - f(tr)\right]$$
$$= \frac{1}{t}[(1-\lambda)f(tu) + \lambda f(tv) - f(\gamma(tu,tv,\lambda))] + \frac{1}{t}[f(\gamma(tu,tv,\lambda)) - f(tr)]$$

Le premier terme est non négatif par ∇-convexité de f. Vérifions que le second tend vers zéro. Pour t assez petit, en utilisant la condition de Lipschitz et le lemme 1, on obtient

$$\frac{1}{t}|f(\gamma(tu,tv,\lambda)) - f(tr)| \leq \frac{k}{t}\|\gamma(tu,tv,\lambda) - tr\| \leq \frac{k}{t}c\|tu - tv\|^2 \to 0.$$

Corollaire 2.3.6. *Soit f une fonction convexe définie sur une partie convexe A de M d'intérieur A° convexe. Alors f est continue dans A°.*

Preuve :
Il suffit d'appliquer la proposition 2.3.5 à la variété A°.

Meyer a fait remarquer les auteurs que le (b) de la proposition 2.3.5 admet une réciproque :

Corollaire 2.3.7. *[24] Soit f une fonction sur M. Pour que f soit ∇-convexe, il faut et il suffit que, pour chaque $a \in M$, il existe une forme linéaire h sur T_aM et un voisinage V_a de a tels que pour tout $x \in V_a$, on ait $f(x) - f(a) \geq h(\overrightarrow{ax})$.*

Preuve :
Pour la nécessité, il suffit, en appliquant la proposition (2.3.5), de choisir h telle que $h \leq \delta f$ et de remarquer que, par convexité le long de chaque géodésique issue de a, on a $f - f(a) \geq \delta f \circ exp_a^{-1}$.
Réciproquement, si $\gamma : I \to M$ est une géodésique,$f \circ \gamma$ est une fonction sur I telle que pour tout $t \in I$, γ a un voisinage U sur lequel est définie une fonction affine h avec $h(t) = f \circ \gamma(t)$, $h \leq f \circ \gamma$ sur U. Ceci entraine $f \circ \gamma$ est convexe, donc f est ∇-convexe. □

Théorème 2.3.8. *Soit $f : A \to \mathbb{R}$ une fonction de classe C^1.*
La fonction f est convexe si et seulement si

$$f(x) + \dot{\gamma}_{xy}(f)(x) \leq f(y), \qquad \forall x, y \in A.$$

La forme linéaire du corollaire 2.3.7 est, dans ce théorème, la dérivée de f en x dans la direction $\dot{\gamma}_{xy}$.

Preuve
C'est un corollaire immédiat de la propriété suivante : pour qu'une fonction φ définie et C^1 sur un intervalle I de $\mathbb{R}$ soit convexe, il faut et il suffit que $\varphi(t) - \varphi(s) \geq (t-s)\varphi'(s)$ pour tous s et t de I.

On remarque que :
1) La fonction $y \mapsto f(x) + \dot{\gamma}_{xy}(f)(x) = f(x) + df(\dot{\gamma}_{xy}(0))$, x est un point fixe, est une approximation radiale linéaire de f car elle est de degré au plus 1 par rapport au paramètre t sur chaque géodésique passant par x en $t = 0$, et elle est affine au point x.
2) Si f est une submersion, alors le graphe

$$G(f) = \{(x, f(x)) | x \in A\} \subset A \times \mathbb{R}$$

est une hypersurface de $A \times \mathbb{R}$.
Soit x un point de A. Le graphe G_* de la fonction

$$y \mapsto f(x) + \dot{\gamma}_{xy}(f)(x) = f(x) + df(\dot{\gamma}_{xy}(0))$$

est une hypersurface totalement géodésique en x tangente à $G(f)$ en $(x, f(x))$. Le théorème précédent montre que f est convexe si et seulement l'hypersurface totalement géodésique G_* tangente à $G(f)$ se situe au dessous de $G(f)$. Cette remarque produit une nouvelle définition géométrique de la convexité d'une fonction de classe C^1.

Théorème 2.3.9. *Soit $f : A \to \mathbb{R}$ une fonction de classe C^1. La fonction f est convexe si et seulement si*

$$df(\dot{\gamma}_{xy})(x) - df(\dot{\gamma}_{xy})(y) \leq 0, \qquad \forall x, y \in A, \quad \forall \gamma_{xy} \in \Gamma$$

c'est-à-dire la différentielle est monotone.

Preuve
C'est une conséquence directe du fait qu'une fonction C^1 sur un intervalle est convexe si et seulement si sa dérivée est croissante.

2.4 Extrémums d'une fonction convexe

Remarquons d'abord que si une fonction f sur M a un minimum local en un point x_0, ce point est a fortiori un minimum local pour la restriction de f à tout sous-ensemble de M contenant x_0, et en particulier à toute géodésique passant par x_0. Mais la réciproque de cette propriété n'est pas vraie ; par exemple, la fonction

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) = (x - y^2)(2x - y^2)$$

a un minimum local en $(0, 0)$ sur toutes les droites passant par ce point, mais ce n'est pas un minimum local de f.

Théorème 2.4.1. *Soit $f : A \to \mathbb{R}$ une fonction convexe de classe C^1. Si x est un point critique de f, alors x est un minimum dans A (il est suffisant que f soit différentiable seulement au point x).*

Preuve

La convexité implique $f(x) + df(\dot{\gamma}_{xy}(x)) \leq f(y)$. Puisque x est un point critique de f, i.e., $df(x) = 0$, on trouve

$$f(x) \leq f(y), \qquad \forall y \in A.$$

La réciproque de ce théorème n'est pas vraie.

Théorème 2.4.2. *Soit A d'intérieur non vide. Si une fonction convexe $f : A \to \mathbb{R}$ admet un maximum global dans l'intérieur de A, alors f est constante.*

Preuve

Supposons que x' est un maximum global de f sur A, i.e., $f(y) \leq f(x')$, $\forall y \in A$, et que x' est un point intérieur de A.
Chosissons $y \in A \setminus \{x'\}$. Il existe $x \in A$ tel que x' soit un point intérieur de la géodésique γ_{xy} joignant x à y ; on a donc un $s \in]0, 1[$ tel que $\gamma_{xy}(s) = x'$, ce qui donne

$$f(x') \leq (1 - s)f(x) + sf(y).$$

Puisque x' est un maximum global, on a $f(x) \leq f(x')$. Si $f(y) < f(x')$, alors

$$f(x') < (1 - s)f(x') + sf(x') = f(x'),$$

ce qui est absurde. Donc $f(y) = f(x')$, $\forall y \in A$. $\square$

Ce théorème dit que si une fonction convexe f est non contante, alors tous les points maximums globaux appartiennent au bord de l'ensemble convexe A.

2.5 Construction de l'enveloppe convexe inférieure d'une fonction.

Étant donné un ensemble convexe A et une fonction $f : A \to \mathbb{R}$ minorée, considérons l'ensemble $\mathfrak{C}$ formé de toutes les fonctions convexes $\varphi : A \to \mathbb{R}$ majorées par f. Cet ensemble n'est pas vide, car il contient la fonction constamment égale à c, où c est un réel minorant f. Pour $x \in A$, on a $\varphi(x) \leq f(x)$ pour toute $\varphi \in \mathfrak{C}$, ce qui permet de définir

$$g(x) = \sup_{\varphi \in \mathfrak{C}} \varphi(x) \ .$$

La fonction $g : A \to \mathbb{R}$ ainsi définie est convexe sur A comme supremum d'une famille de fonctions convexes (corollaire 2.2.10) ; g est donc le plus grand élément de $\mathfrak{C}$, c'est-à-dire la plus grande fonction convexe minorant f. Elle est appelée enveloppe convexe inférieure de f sur A, ou plus grande minorante convexe de f sur A.

On peut construire cette enveloppe comme limite d'une suite décroissante de fonctions, de la façon suivante. Toujours pour $f : A \to \mathbb{R}$ minorée par une constante c, posons pour $x \in A$

$$Gf(x) = \inf_{\substack{(y,z,t) \,\in\, A \times A \times [0,1] \\ \gamma_{yz}(t) \,=\, x}} \left[(1-t)f(y) + tf(z)\right] \ ,$$

où l'inf porte sur tous les triplets $(y, z, t) \in A \times A \times [0, 1]$ tels que $\gamma_{yz}(t) = x$. Puisque f est minorée par c, on a $(1-t)f(y)+tf(z) \geq c$ pour $t \in [0, 1]$; ainsi l'inf existe dans $\mathbb{R}$ et est lui aussi minoré par c. On a donc un opérateur (non linéaire) G agissant dans l'ensemble des fonctions sur A minorées par c ; on peut l'itérer et définir G^k pour tout entier $k \geq 1$.

Une propriété de G est que $Gf \leq f$ (prendre $(y, z, t) = (x, x, \frac{1}{2})$ dans l'inf). Pour f donnée, la suite des fonctions $G^k f$ est donc décroissante, et comme elle est minorée, elle converge vers une limite $G^\infty f = \lim_k G^k f$.

Proposition 2.5.1. *On conserve les notations ci-dessus.*

i) *f est convexe si et seulement si $Gf = f$.*

ii) *Si $f \leq g$, alors $Gf \leq Gg$.*

iii) *Si $g : A \to \mathbb{R}$ est minorée et est la limite d'une suite décroissante de fonctions f_n sur A, alors $Gg = \lim_n \downarrow Gf_n$.*

iv) *$G^\infty f$ est l'enveloppe convexe inférieure de f.*

Preuve :
i) $Gf = f$ signifie que pour tout x de A, on a :

$$f(x) = \inf_{x \,=\, \gamma_{yz}(t)} [(1-t)f(y) + tf(z)]$$

autrement dit

$$f(x) \leq (1-t)f(\gamma(0)) + tf(\gamma(1))$$

pour toute géodésique γ vérifiant $\gamma(t) = x$ pour un $t \in [0,1]$; mais ceci exprime la convexité de f.
ii) Cette propriété est une conséquence immédiate de la définition de G.
iii) Si g est la limite d'une suite décroissante de fonctions f_n, alors la suite des Gf_n est décroissante d'après ii), et

$$\begin{aligned} Gg(x) &= \inf_{x=\gamma_{yz}(t)} [(1-t)g(\gamma_{yz}(0)) + tg(\gamma_{yz}(1))] \\ &= \inf_{x=\gamma_{yz}(t)} \inf_n [(1-t)f_n(\gamma_{yz}(0)) + tf_n(\gamma_{yz}(1))] \\ &= \inf_n \inf_{x=\gamma_{yz}(t)} [(1-t)f_n(\gamma_{yz}(0)) + tf_n(\gamma_{yz}(1))] = \lim_n \downarrow Gf_n(x). \end{aligned}$$

iv) La suite des $G^n f$ décroît et a pour limite $G^\infty f$; donc, en utilisant le point iii),

$$G(G^\infty f) = \lim_n \downarrow G(G^n f) = \lim_n \downarrow G^{n+1} f = G^\infty f,$$

et d'après i) $G^\infty f$ est convexe. Si h est n'importe quelle fonction convexe majorée par f, la fonction $h' = \sup(h, c)$ est convexe, majorée par f et minorée, et on peut écrire $h' \leq G^n h' \leq G^n f$ pour tout n. A la limite on obtient $h' \leq G^\infty f$, donc aussi $h \leq G^\infty f$; ceci montre que $G^\infty f$ est la plus grande minorante convexe de f.

(Les propriétés précédentes sont prouvées et utilisées dans [23] pour y démontrer le théorème 1.)
Ces minorantes convexes permettent de démontrer une propriété d'approximation de fonctions convexes sur des voisinages d'un compact de la variété ; il est montré dans [4] :

Proposition 2.5.2. *Soit K un compact convexe de la variété M. Il existe une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions convexes définies sur des voisinages de K, telles que pour tout $\varepsilon > 0$ et toute fonction f convexe sur un voisinage de K, il existe $n \in \mathbb{N}$ tel que* $\sup_K |f - f_n| < \varepsilon$.

Preuve :
On considère une suite décroissante W_p de voisinages ouverts de K, formant une base de voisinages de K, et telle que $\overline{W_{p+1}} \subset W_p$. Tout voisinage de K contient un $\overline{W_p}$. On choisit dans chaque compact $\overline{W_p}$ une suite dense pour la norme uniforme $(g_{p,n})_{n\in\mathbb{N}}$ de fonctions continues, et on considère la suite

$(h_{p,n})_{n\in\mathbb{N}}$ des plus grandes minorantes convexes de leurs restrictions à W_p ; en les renumérotant, on obtient une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions convexes.
Soit f une fonction convexe définie sur un voisinage W' de K, et $\varepsilon > 0$. Il existe un p tel que $\overline{W_p}$ soit inclus dans W', et n' tel que $\sup_{\overline{W_p}} |f - g_{p,n'}| < \varepsilon$.
Cela implique que $\sup_{W_p} |f - h_{p,n'}| < \varepsilon$, puisque $g_{p,n'}$ est supérieure à $f - \varepsilon$ qui est elle même convexe.
La fonction $f_n = h_{p,n'}$ répond à la question.

2.6 Géométrie convexe

Définition 2.6.1. *[3] On dit qu'une variété affine (M, ∇) est à géométrie convexe s'il existe une fonction convexe $\psi : M \times M \to \mathbb{R}$ de classe C^3, positive, qui s'annule exactement sur la diagonale et telle que, pour tout $a \in M$, l'application $\psi_a = \psi(a, .)$ ait une hessienne $\nabla d\psi_a$ strictement positive sur $M \backslash \{a\}$.*
Un compact K convexe d'une variété affine (M, ∇) est à géométrie convexe s'il possède un voisinage ouvert qui est à géométrie convexe.

Dans la définition précédente, la fonction ψ s'appelle fonction de séparation ou séparante.
D'après [22] lemme (4.59), de telles fonctions existent localement ; tout point de K a un voisinage à géométrie convexe.
Si K est suffisamment petit, il est à géométrie convexe.
D'autre part, Kendall a démontré dans [44] que les boules géodésiques régulières, dans les variétés riemannienne sont à géométrie convexe. Il a explicité dans le théorème suivant une famille de fonctions de séparation.

Théorème 2.6.1. *[44] Soit $\mathfrak{B}$ une boule géodésique régulière de centre p et de rayon r telle que $\sqrt{k}r < \frac{1}{2}\pi$, où k est un majorant positif des courbures sectionnelles de $\mathfrak{B}$.*
$\mathfrak{B}$ est à géométrie convexe, une fonction de séparation étant donnée par : pour tout $(x, y) \in \mathfrak{B} \times \mathfrak{B}$

$$\phi^{(k)}_{\nu,h}(x, y) = \left(\frac{1 - \cos\sqrt{k}\,\delta(x, y)}{\cos\sqrt{k}\,\delta(p, x)\cos\sqrt{k}\,\delta(p, y) - \widetilde{h}^2}\right)^{\nu+1}$$

où δ est la distance induite par la métrique riemannienne, $h = \cos(\sqrt{k}r)$, $\nu \geq 1$ et $2\nu\widetilde{h}^2(h^2 - \widetilde{h}^2) \geq 1$ et $\widetilde{h} \in [0, h]$.
De plus, si $2\nu\widetilde{h}^2(h^2 - \widetilde{h}^2) > 1$ alors

$$\frac{d^2}{dt^2}\phi^{(k)}_{\nu,h}(\gamma(t)) \geq C\left[\frac{d}{dt}\phi^{(k)}_{\nu,h}(\gamma(t))\right]^2$$

pour toute géodésique γ dans $\mathfrak{B} \times \mathfrak{B}$, pour une constante positive $C = C(\nu, h, \widetilde{h})$.

Arnaudon [4] a utilisé l'une de ces fonctions pour construire une famille de fonctions convexes sur $\mathfrak{B}''$, où $\mathfrak{B}''$ est une boule géodésique régulière de rayon ρ dont l'intérieur $\mathfrak{B}'$ contient $\mathfrak{B}$.
Rappelons la construction de ces fonctions [4] qui vont permettre de donner des majorations de la taille du barycentre convexe dans les chapitres suivants. Soit $h = \cos(\sqrt{k}\rho)$ (on suppose que k est aussi un majorant des courbures sectionnelles de $\mathfrak{B}''$) et ν un réel strictement supérieur à $\frac{4}{h^4}$.
Pour $a, x \in \mathfrak{B}''$, on pose

$$\psi_a(x) = \left(\frac{1 - \cos(\sqrt{k}\delta(a,x))}{\cos(\sqrt{k}\delta(p,a))\cos(\sqrt{k}\delta(p,x)) - \frac{h^2}{2}}\right)^{\nu+1}$$

On démontre dans [4] les résultats suivants :

Lemme 2.6.2. *Soit $\varepsilon > 0$, alors pour tous $a, x \in \mathfrak{B}''$ vérifiant $\delta(a,x) \geq \varepsilon$ on a $Hess\,\psi_a(x) \geq (1 - \cos\sqrt{k}\varepsilon)^\nu k\, g(x)$.*

Lemme 2.6.3. *Il existe un réel positif A tel que les fonctions $h_a(x) = (1 - \cos\sqrt{k}\delta(a,x))^{\frac{3}{2}} + A\psi_a(x)$ soient convexes sur $\mathfrak{B}''$, et tel qu'il existe des constantes $c > 0$ et $C > 0$ vérifiant pour tous $a, x \in \mathfrak{B}''$,*

$$c\delta(a,x)g(x) \leq Hess\, h_a(x) \leq C\delta(a,x)g(x)$$

Proposition 2.6.4. *Soit K un compact de la variété M. Il existe une constante $B > 0$, telle que pour tout point $a \in K$ et toute forme linéaire $\lambda_a \in T_a^*K$ de norme 1, les fonctions $x \mapsto \lambda_a(\overrightarrow{ax}) + B\, h_a(x)$ soient convexes sur $\mathfrak{B}''$.*

Nous allons reprendre avec un peu plus de détails quelques passages dans les preuves du lemme 2.6.3 et la proposition 2.6.4 pour constater comment les courbures sectionnelles interviennent dans le choix des constantes A et B dans le lemme 2.6.3 et la proposition 2.6.4 respectivement.
Preuve du lemme 2.6.3
Nous nous intéressons uniquement à la minoration de $Hess\, h_a(x)$ car le nombre c intervient dans le choix de la constante B de la proposition (2.6.4). Pour minorer $Hess\, h_a(x)$, on envisagera deux cas : $0 < \delta(a,x) \leq \frac{\pi}{3\sqrt{k}}$ et $\delta(a,x) \geq \frac{\pi}{3\sqrt{k}}$.
Posons $f_a(x) = \cos\sqrt{k}\delta(a,x)$.

$$Hess(1 - f_a(x))^{\frac{3}{2}} = \frac{3}{2}\sqrt{1 - f_a(x)}Hess(1 - f_a)(x)$$

$$+\frac{3}{4\sqrt{1-f_a(x)}}d(1-f_a)(x)\otimes d(1-f_a)(x)$$

• Si $0<\delta(a,x)\leq\frac{\pi}{3\sqrt{k}}$ on pose $X=\sqrt{k}\delta(a,x)$ et on considère la fonction $g(X)=\frac{\sqrt{1-\cos X}}{X}$. g est croissante sur $]0,\frac{\pi}{3}]$ et $\lim_{X\to 0^+}g(X)=\frac{\sqrt{2}}{2}$; on en déduit que $\sqrt{1-f_a(x)}\geq\frac{\sqrt{2}}{2}\sqrt{k}\delta(a,x)$ pour tous a,x vérifiant $0<\delta(a,x)\leq\frac{\pi}{3\sqrt{k}}$.
Or les fonctions f_a vérifient $Hess\, f_a+kf_a\, g\leq 0$ pour $x\neq a$ (voir [54], lemme 1.2.1), donc $Hess(1-f_a)(x)\geq\frac{k}{2}\, g(x)$.
Finalement :
On a $Hess\, h_a(x)>Hess(1-\cos\sqrt{k}\delta(a,x))^{\frac{3}{2}}>\frac{3}{2}\sqrt{1-f_a(x)}Hess(1-f_a)(x)$ car ψ_a est convexe, on déduit $Hess\, h_a(x)\geq c\delta(a,x)g(x)$ avec $c=\frac{3}{8}\sqrt{2k}k$.
• Si $\delta(a,x)\geq\frac{\pi}{3\sqrt{k}}$, $\sqrt{1-f_a}(x)=\sqrt{2}|\sin(\frac{\sqrt{k}}{2}\delta(a,x))|\geq\sqrt{2}$
et $Hess(1-f_a(x))\geq -kg(x)$ (Voir [54]), alors il existe un réel positif M vérifiant $Hess(1-f_a(x))^{\frac{3}{2}}\geq -Mg(x)$, $M=k\sqrt{2}$ et d'après lemme (2.6.2), on a $Hess\,\psi_a(x)\geq\frac{k}{2^\nu}g(x)$, $(\varepsilon=\frac{\pi}{3\sqrt{k}})$.
Alors $Hess\, h_a(x)\geq(A\frac{k}{2^\nu}-M)g(x)$, on choisit $A>2^{\nu+\frac{1}{2}}$
et $0<c<(A\frac{k}{2^\nu}-M)\frac{\sqrt{k}}{\pi}$. □

Preuve de la proposition 2.6.4

Il suffit de remarquer qu'il existe une contante B' positive, telle que pour tous λ_a et x, on ait $Hess\,\lambda_a(\overrightarrow{ax})\geq -B'\delta(a,x)g(x)$.
On pose $B=\frac{B'}{c}$ où c est la constante du lemme 2.6.3. □

On remarque que :
1 - Les fonctions $f_a(x)=\lambda_a(\overrightarrow{ax})+B\, h_a(x)$ vérifient $df_a(a)=\lambda_a$ et $Hess\, f_a(a)=0$.
2 - au voisinage de la diagonale Δ, les fonctions $h_a(x)=(1-\cos\sqrt{k}\delta(a,x))^{\frac{3}{2}}+A\psi_a(x)$ sont équivalentes au cube de la distance $\delta(a,x)^3$ et les fonctions $f_a(x)=\lambda_a(\overrightarrow{ax})+B\, h_a(x)$ sont équivalentes à $x\mapsto\lambda_a(\overrightarrow{ax})+D\delta(a,x)^3$, où D est une constante positive qui dépend de k, majorant des courbures sectionnelles.

Chapitre 3

Barycentre convexe d'une probabilité définie sur une variété

Nous travaillerons dans une variété M réelle, sans bord, C^∞, de dimension finie, munie d'une connexion C^∞ sans torsion. Nous supposerons la variété M convexe, au sens où deux points quelconques sont joints par une géodésique et une seule, qui dépend de façon C^∞ des deux points.

On considère toujours un compact convexe K de M admettant une base de voisinages convexes, et on note $C(K)$ l'ensemble des fonctions convexes définies dans un voisinage de K ouvert et convexe. $\overrightarrow{xy}$ désigne le vecteur vitesse en x de la géodésique passant en x au temps 0 et en y au temps 1.

3.1 Définitions et propriétés

Les définitions suivantes sont empruntées à Émery et Mokobodzki [23] et Arnaudon [4].

Définition 3.1.1. *Si μ est une probabilité sur K, on dira que $x \in K$ est un point du barycentre convexe de μ si pour toute fonction f appartenant à $C(K)$, on a $f(x) \leq \mu(f)$. Le barycentre convexe $b(\mu)$ est l'ensemble de ces points.*

On appellera barycentre exponentiel de μ tout point e de K qui vérifie $\int_K \overrightarrow{ey}\, \mu(dy) = 0$.

A priori, l'ensemble $b(\mu)$ ne dépend pas seulement de μ, mais aussi du choix du compact K. Il est cependant toujours possible en pratique de choisir le plus petit K possible, qui est l'enveloppe convexe du support de μ (c'est-à-dire l'intersection des compacts convexes qui portent μ).

Proposition 3.1.1. *[23] Tout barycentre exponentiel d'une probabilité μ est dans $b(\mu)$.*

Démonstration :
La preuve est empruntée à Arnaudon [3], elle est analogue à celle de [23] proposition 2.
Soit a un barycentre exponentiel de μ. Pour $f \in C(K)$, on sait par le corollaire 2.3.7 qu'il existe une forme linéaire $h \in T_a^*K$ telle que $f(x) \geq f(a) + h(\overrightarrow{ax})$ pour tout x dans le domaine de f. Une intégration par rapport à μ donne :

$$\mu(f) \geq f(a) + \int_K h(\overrightarrow{ax})\,\mu(dy)$$

et d'après la linéarité de h, on a :

$$\int_K h(\overrightarrow{ax})\,\mu(dy) = h\int_K (\overrightarrow{ax})\,\mu(dy).$$

Comme a est un barycentre exponentiel de μ, on a $\int_K (\overrightarrow{ax})\,\mu(dy) = 0$, ce qui donne $\mu(f) \geq f(a)$. $\square$

Proposition 3.1.2. *[23] S'il existe une fonction de séparation sur un voisinage de K ouvert convexe, alors toute probabilité sur K a au plus un barycentre exponentiel.*

Remarque 3.1.1. *La diagonale Δ de $M \times M$ est une sous-variété totalement géodésique pour la connexion produit caractérisée par ses géodésiques $t \mapsto (\gamma_1(t), \gamma_2(t))$ où $t \mapsto \gamma_i(t)$ $i = 1, 2$ est une géodésique de M.*

Lemme 3.1.3. *Si μ est une probabilité sur $K \times K$ ayant pour marges ν et π, et si x (respectivement y) est un barycentre exponentiel de ν (respectivement π) dans K, alors (x, y) est un barycentre exponentiel de μ dans $K \times K$.*

Démonstration : Soient α et β dans K ; posons $X = \overrightarrow{x\alpha}$ et $Y = \overrightarrow{y\beta}$. La courbe $t \mapsto (exp_x tX, exp_y tY)$ est une géodésique vérifiant $\gamma(1) = (\alpha, \beta)$ et $\dot\gamma(0) = (X, Y)$, donc $\overrightarrow{(x, y)(\alpha, \beta)} = (X, Y)$.
Si h est une forme linéaire sur $T_{(x,y)}(M \times M)$, alors il existe deux formes linéaires $f \in T_x^*M$ et $g \in T_y^*M$ telles que $h(u, v) = f(u) + g(v)$. On a :

$$h(\overrightarrow{(x, y)(\alpha, \beta)}) = h(X, Y) = f(X) + g(Y) = f(\overrightarrow{x\alpha}) + g(\overrightarrow{y\beta}).$$

En intégrant par rapport à μ dont les marges sont ν et π, on trouve

$$\begin{aligned} h\Big(\int_{K\times K} \overrightarrow{(x,y)(\alpha,\beta)}\, \mu(d(\alpha,\beta)\Big) &= \int_{K\times K} h(\overrightarrow{(x,y)(\alpha,\beta)})\, \mu(d(\alpha,\beta)) \\ &= \int_K f(\overrightarrow{x\alpha})\, \nu(d\alpha) + \int_K g(\overrightarrow{y\beta})\, \pi(d\beta) \\ &= f\Big(\int_K \overrightarrow{x\alpha}\, \nu(d\alpha)\Big) + g\Big(\int_K \overrightarrow{y\beta}\, \pi(d\beta)\Big) = 0 \end{aligned}$$

car x et y sont des barycentres exponentiels de ν et π respectivement.
Or h est arbitraire, donc $\int_{K\times K}(\overrightarrow{(x,y)(\alpha,\beta)})\mu(d(\alpha,\beta)) = 0$. □

Démonstration de la proposition 3.1.2 :
Soient x et y deux barycentres exponentiels de la même probabilité μ sur K et soit λ la probabilité sur la diagonale Δ dont les deux marges sur M sont égales à μ.
Le lemme précédent entraine que (x, y) est un barycentre exponentiel de λ sur $K\times K$ et la proposition (3.1.1) implique que (x, y) est dans $b(\lambda)$. Si φ est une fonction de séparation sur un voisinage de $K \times K$, on a $\varphi(x, y) \leq \lambda(\varphi) = 0$, car λ est portée par la diagonale ; donc $x = y$. □

Proposition 3.1.4. *(Associativité des barycentres.) Soient $(A, \mathfrak{A}, \pi)$ un espace probabilisé, $(\mu_\alpha)_{\alpha\in A}$ une famille de probabilités sur K, dépendant mesurablement de α et $(x_\alpha)_{\alpha\in A}$ une v. a dans K telle que $x_\alpha \in b(\mu_\alpha)$ pour chaque α. Si $\nu = \int \varepsilon_{x_\alpha}\pi d(\alpha)$ et $\mu = \int \mu_\alpha\pi(d\alpha)$ Alors $b(\nu) \subset b(\mu)$.*
Autrement dit, tout point du barycentre d'une famille de barycentres de μ est lui même dans le barycentre de μ.

Preuve
C'est très simple : soit $f \in C(K)$. Si $x \in b(\nu)$, alors $f(x) \leq \nu(f) = \int f(x_\alpha)\pi d(\alpha)$, or $x_\alpha \in b(\mu_\alpha)$, donc $f(x_\alpha) \leq \mu_\alpha(f)$, ce qui implique

$$f(x) \leq \int \mu_\alpha(f)\pi d(\alpha) = \mu(f). \qquad \square$$

La réciproque de la proposition d'associativité précédente est fausse. Étant donnés $\mu = \int \mu_\alpha \pi d(\alpha)$ et $x_\alpha \in b(\mu)$, il peut ne pas exister de points $x_\alpha \in b(\mu_\alpha)$ tels que $x \in b(\int \varepsilon_{x_\alpha}\pi d(\alpha))$.
Prenons comme exemple le cas où M est la sphère bidimensionnelle, K une boule fermée incluse dans un hémisphère ouvert, et A, B et C trois points de K formant un triangle non dégénéré, et considérons la mesure somme de masses de Dirac $\mu = \frac{1}{2}\varepsilon_A + \frac{1}{4}\varepsilon_B + \frac{1}{4}\varepsilon_C$. Si A', B' et C' sont les milieux des

côtés du triangle ABC, le milieu M du segment géodésique AA' n'est pas sur la géodésique $B'C'$. On a donc $B' = b(\frac{1}{2}\varepsilon_A + \frac{1}{2}\varepsilon_C)$ et $C' = b(\frac{1}{2}\varepsilon_A + \frac{1}{2}\varepsilon_B)$, mais M, bien que dans le barycentre de la mesure μ, n'est cependant pas celui de $\frac{1}{2}\varepsilon_{B'} + \frac{1}{2}\varepsilon_{C'}$.
Cet exemple montre bien où se situe le problème : l'ensemble $b(\mu)$ prend en compte les barycentres obtenus en regroupant entre eux de toutes les manières possibles les points chargés par μ ; une décomposition $\mu = \int \mu_\alpha \pi(d\alpha)$ est un choix parmi ces regroupements, d'où moins de latitude dans la construction du barycentre.

Lien entre les barycentres et l'interpolation géodésique

La procédure d'interpolation géodésique consiste à remplacer deux points x et y de K, affectés de poids positifs $p(x)$ et $p(y)$ de somme 1, par le point $\gamma_{xy}(p(y)) = \gamma_{yx}(p(x))$ situé sur la géodésique qui joint x et y. Ce point $\gamma_{xy}(p(y))$ est dans le barycentre $b(\mu)$, où μ est la mesure $p(x)\varepsilon_x + p(y)\varepsilon_y$.

Cette procédure peut être itérée de la façon suivante : On se donne un arbre dyadique de longueur n, ayant donc 2^n extrémités, et à chacune de ces 2^n extrémités on associe un point de K et un poids ≥ 0, la somme des poids étant 1. On peut alors, de manière unique, associer de proche en proche à chaque sommet S de l'arbre un point de K et un poids, de telle sorte que le poids de S soit la somme des poids des extrémités de l'arbre qui descendent de S, et que le point associé à S soit l'interpolé géodésique des deux successeurs de S, affectés des poids correspondants. Le point associé à la racine de l'arbre (avec poids 1) sera le barycentre géodésique itéré des points associés aux 2^n extrémités, affectés des poids corespondants. Mais il est important de remarquer, comme le montre l'exemple qui suit la proposition 3.1.4, que ce barycentre géodésique itéré ne dépend pas seulement des 2^n points donnés et de leurs poids, mais aussi de l'arbre, qui dicte l'ordre dans lequel sont faites les interpolations géodésiques.

Voici une façon rigoureuse de décrire cette procédure ; nous suivons les notations utilisées dans [23].

On se donne un entier $n \geq 1$, et on appelle W^n l'ensemble $\{0, 1\}^n$, qui a 2^n éléments ; cet ensemble W^n est une façon commode de coder les 2^n extrémités de l'arbre dyadique de longueur n. Les poids affectés aux extrémités seront codés par une probabilité p sur l'ensemble fini W^n, et les points de K associés aux extrémités seront donnés par une fonction $y : W^n \to K$. Le barycentre géodésique itéré de (y, p) est le point $\beta_n(y, p)$ de K défini comme suit.

Pour $n = 1$, $W^n = \{0, 1\}$, et $\beta_1(y, p)$ est par définition le point $\gamma(p(1))$, où $\gamma : [0, 1] \to K$ est la géodésique telle que $\gamma(0) = y(0)$ et $\gamma(1) = y(1)$.

C'est simplement l'interpolation géodésique.

Pour $n > 1$, à partir de la probabilité p sur W^n et de la fonction $y : W^n \to K$, on définit deux probabilités p_0 et p_1 sur W^{n-1} et deux fonctions y_0 et y_1 sur W^{n-1} par

$$p_\alpha(w_2, \ldots, w_n) = \frac{p(\alpha, w_2, \ldots, w_n)}{p\{w_1 = \alpha\}} = \frac{p(\alpha, w_2, \ldots, w_n)}{\sum_{(v_2,\ldots,v_n)\in W^{n-1}} p(\alpha, v_2, \ldots, v_n)} \ ;$$

$$y_\alpha(w_2, ..., w_n) = y(\alpha, w_2, ..., w_n) \ .$$

On peut alors définir $\beta_n(y, p)$ comme le point $\gamma(p\{w_1 = 1\})$ de K, où $\gamma : [0, 1] \to K$ est la géodésique telle que $\gamma(0) = \beta_{n-1}(y_0, p_0)$ et $\gamma(1) = \beta_{n-1}(y_1, p_1)$.

Si $p\{w_1 = 0\} = 0$ (resp. $p\{w_1 = 1\} = 0$), le dénominateur s'annule et p_0 et $\gamma(0)$ (resp. p_1 et $\gamma(1)$) ne sont pas définis ; mais ce n'est pas gênant, car $p\{w_1 = 1\}$ vaut alors 1 (resp. 0), et seul le point $\gamma(1)$ (resp. $\gamma(0)$) intervient dans la définition de $\beta_n(y, p)$.

Notons que dans [23], les barycentres géodésiques itérés sont des ensembles, alors qu'ici se sont des points ; cette différence provient de notre hypothèse d'unicité de la géodésique reliant deux points.

Il existe une relation étroite entre les itérés de l'opérateur G intervenant dans la construction de l'enveloppe convexe inférieure d'une fonction (proposition 2.5.1) et les barycentres géodésiques itérés :

Lemme 3.1.5. *Si $f : K \to \mathbb{R}$ est une fonction minorée, alors pour tout x de K*

$$G^n f(x) = \inf_{\beta_n(y,p)=x} \int_{W^n} f \circ y \, dp \ ,$$

où l'$\inf$ *est pris sur toutes les fonctions $y : W^n \to K$ et toutes les probabilités p sur W^n telles que $\beta_n(y, p) = x$.*

Preuve :
Remarquer que l'intégrale n'est autre que la somme finie $\sum_{w\in W^n} f(y(w))\, p(w)$.
La formule est vraie pour $n = 1$ puisque les couples (y, p) tels que $\beta_1(y, p) = x$ correspondent aux triplets (u, v, t) tels que $x = \gamma_{uv}(t)$ par les relations $u = y(0)$, $v = y(1)$ et $t = p(1)$.
Supposons que ce soit vrai pour $n - 1$. Alors en écrivant

$$G^n f(x) = \inf_{\substack{u_0,u_1,t \\ \gamma_{u_0u_1}(t)=x}} (1 - t)\, G^{n-1} f(u_0) + t\, G^{n-1} f(u_1),$$

on peut remplacer $G^{n-1} f(u_\alpha)$ par $\inf_{\beta_{n-1}(y_\alpha,p_\alpha)=u_\alpha} \int f \circ y_\alpha \, dp_\alpha$. Il ne reste qu'à poser $y(\alpha, w) = y_\alpha(w)$, $p(0, w) = (1 - t)\, p_0(w)$ et $p(1, w) = t\, p(w)$ et à utiliser la définition de β_n pour achever la démonstration. $\square$

La proposition et le théorème qui suivent explicitent les liens entre barycentres convexes et barycentres géodésiques itérés.

Proposition 3.1.6. *Etant donné un $n \geq 1$, une probabilité p sur W^n et une application y de W^n dans K, le barycentre géodésique itéré $\beta_n(y,p)$ est un point du barycentre de la mesure $y(p) = \sum_{w \in W^n} p(w)\, \varepsilon_{y(w)}$.*

Preuve :
Lorsque $n = 1$, en appelant μ la probabilité $y(p) = p(0)\varepsilon_{y(0)} + p(1)\varepsilon_{y(1)}$ et γ la géodésique $\gamma_{y(0)y(1)}$, il suffit d'écrire, pour toute fonction f convexe sur un voisinage de K,

$$\begin{aligned} f(\beta_1(y,p)) = f[\gamma(p(1))] \ &\leq \ (1-p(1))\, f(\gamma(0)) + p(1)\, f(\gamma(1)) \\ &= p(0)\, f(y(0)) + p(1)\, f(y(1)) = \mu(f) \ . \end{aligned}$$

Pour $n > 1$, on procède par récurrence. En reprenant les p_0, p_1, y_0 et y_1 utilisés dans la définition de $\beta_n(y,p)$, le même calcul que ci-dessus dit que le point $\beta_n(y,p)$ appartient à $b(\nu)$, où ν est la mesure

$$p\{w_1 = 0\}\, \varepsilon_{\beta_{n-1}(y_0,p_0)} + p\{w_1 = 1\}\, \varepsilon_{\beta_{n-1}(y_1,p_1)} \ .$$

L'hypothèse de récurrence dit que pour $\alpha = 0$ ou 1, $\beta_{n-1}(y_\alpha, p_\alpha)$ est dans le barycentre $b(\mu_\alpha)$ de la mesure $\mu_\alpha = y_\alpha(p_\alpha)$. L'associativité des barycentres (proposition 3.1.4) entraîne que $b(\nu)$ est contenu dans $b(\mu)$, où μ est la mesure

$$\begin{aligned} \mu \ &= \ p\{w_1 = 0\}\, \mu_0 + p\{w_1 = 1\}\, \mu_1 \\ &= \ p\{w_1 = 0\}\, y_0(p_0) + p\{w_1 = 1\}\, y_1(p_1) \ . \end{aligned}$$

En écrivant

$$p\{w_1 = \alpha\}\, y_\alpha(p_\alpha) = p\{w_1 = \alpha\} \sum_{w \in W^{n-1}} p_\alpha(w)\, \varepsilon_{y_\alpha(w)} = \sum_{w \in W^{n-1}} p(\alpha, w)\, \varepsilon_{y(\alpha,w)},$$

on obtient

$$\mu = \sum_{w \in W^{n-1}} p(0,w)\, \varepsilon_{y(0,w)} + \sum_{w \in W^{n-1}} p(1,w)\, \varepsilon_{y(1,w)} = \sum_{w \in W^n} p(w)\, \varepsilon_{y(w)},$$

c'est-à-dire que μ n'est autre que $y(p)$; ainsi, $\beta_n(y,p)$ est dans $b(y(p))$. $\square$

Toujours à propos des liens entre barycentres convexes et barycentres géodésiques itérés, voici l'analogue, dans notre cadre, du théorème 1 de [23].

Théorème 3.1.7. *On se donne une suite d'ouverts V_k qui décroît vers K, formant une base de voisinages convexes de K, et tels que $\overline{V_{k+1}}$ soit un compact inclus dans V_k.*

Soient μ une probabilité portée par K, et x un point du barycentre $b(\mu)$. Il existe des entiers $n(k)$, des applications y_k de $W^{n(k)} = \{0,1\}^{n(k)}$ dans V_k, et des probabilités p_k sur $W^{n(k)}$ tels que, pour chaque k,

i) $x = \beta(y_k, p_k)$;

ii) les mesures-images $\mu_\ell = y_\ell(p_\ell)$, qui sont portées par $\overline{V_k}$ pour $\ell \geq k$, tendent vers μ lorsque ℓ tend vers l'infini, au sens de la convergence étroite des mesures sur le compact $\overline{V_k}$.

Démonstration :

On fixe le point $x \in K$ et un entier k.

Remarquons d'abord que si $x = \beta_n(y,p)$ pour une probabilité p sur W^n et une application y de W^n dans $\overline{V_k}$, on a aussi $x = \beta_{n+1}(\bar{y}, \bar{p})$ pour des $\bar{y}$ et $\bar{p}$ tels que la mesure $\bar{y}(\bar{p})$ soit égale à $y(p)$. Il suffit pour cela de choisir (par exemple) $\bar{p}(\alpha, w) = \frac{1}{2}p(w)$ et $\bar{y}(\alpha, w) = y(w)$ pour tous $\alpha \in \{0,1\}$ et $w \in W^n$.

On note $D_k(x)$ l'ensemble des mesures de la forme $y(p)$ pour un entier n, une application y de W^n dans $\overline{V_k}$ et une probabilité p sur W^n tels que $x = \beta_n(y,p)$.

Cet ensemble $D_k(x)$ n'est pas vide, puisqu'il contient la mesure de Dirac ε_x. En outre, $D_k(x)$ est convexe. En effet, si λ_0 et λ_1 sont dans $D_k(x)$, on a $x = \beta_{n_\alpha}(y_\alpha, p_\alpha)$ avec $y_\alpha(p_\alpha) = \lambda_\alpha$ pour $\alpha \in \{0,1\}$. La remarque précédente permet de supposer que $n_0 = n_1$; appelons n ce nombre. Soit π une probabilité définie sur $\{0,1\}$; en définissant y et p sur W^{n+1} par $y(\alpha, w) = y_\alpha(w)$ et $p(\alpha, w) = \pi(\alpha)p(w)$ pour $\alpha \in \{0,1\}$ et $w \in W^n$, on a $x = \beta_{n+1}(y,p)$ par définition des barycentres géodésiques itérés, et il en résulte que $y(p) \in D_k(x)$. Comme $y(p) = \pi(0)\lambda_0 + \pi(1)\lambda_1$, $D_k(x)$ est convexe.

Rappelons que nous avons fixé une probabilité μ sur K telle que $x \in b(\mu)$; nous allons en déduire que μ, qui est aussi une mesure sur le compact plus gros $\overline{V_k}$, est dans l'adhérence de $D_k(x)$ pour la topologie de la convergence étroite des mesures sur $\overline{V_k}$. Si μ n'était pas dans l'adhérence de ce convexe non vide, le théorème de Hahn-Banach appliqué aux mesures sur $\overline{V_k}$ en dualité avec les fonctions continues sur $\overline{V_k}$, fournirait une fonction f continue sur $\overline{V_k}$ et telle que $\mu(f) < \inf_{\lambda \in D(x)} \lambda(f)$. En utilisant (sur le compact $\overline{V_k}$ au lieu de K) le lemme 3.1.5, on aurait

$$\mu(G^\infty f) \leq \mu(f) < \inf_{\lambda \in D(x)} \lambda(f)$$

$$= \inf_n \inf_{(y,p)\in\beta_n^{-1}(x)} \int f\circ y\,dp = \inf_n G^n f(x) = G^\infty f(x).$$

Comme la fonction $G^\infty f$ est convexe sur $\overline{V_k}$, et est donc dans $C(K)$, cette inégalité stricte interdirait à x d'appartenir à $b(\mu)$, ce qui est absurde. Ainsi, $\mu \in \overline{D_k(x)}$.
Soit r une distance sur l'ensemble des probabilités sur $\overline{V_1}$, définissant la topologie étroite. La restriction de r à l'ensemble des probabilités portées par $\overline{V_k}$ définit la convergence étroite des probabilités sur $\overline{V_k}$. Puisque μ est dans l'adhérence de $D_k(x)$ pour cette topologie, il existe dans $D_k(x)$ une suite $(\nu_\ell^k)_{\ell\in\mathbb{N}}$ telle que $r(\nu_\ell^k,\mu) < 2^{-\ell}$. La suite diagonale des $\mu_\ell = \nu_\ell^\ell$ converge vers μ au sens étroit sur chaque $\overline{V_k}$, et vérifie $\mu_\ell \in D_\ell(x)$; ceci établit le théorème.
$\square$

3.2 Majoration du diamètre des barycentres convexes

Soit μ une mesure de probabilité portée par K que l'on suppose à géométrie convexe.
Arnaudon [3] a montré

Proposition 3.2.1. *Sous la conditions écrite ci-dessus, pour toute distance riemannienne δ sur K, il existe une constante C telle que pour toute mesure de probabilité μ sur K, pour tout v dans K, on ait l'inégalité*

$$|b(\mu)| \leq C \int_K \delta^3(v,y)\,\mu(dy),$$

où $|b(\mu)|$ désigne le diamètre de $b(\mu)$ pour la distance δ.

Démonstration de la proposition
Soit g une métrique riemannienne sur K et δ la distance associée. On note $S^\star K$ l'ensemble des formes linéaires sur K de norme 1 pour g. Les applications exp et $\overrightarrow{xy}$ seront calculés avec la connexion (et non pas avec g). En raison de la compacité de K, il existe une constante c_1 strictement positive, telle que pour tous x, y dans K, on ait l'inégalité $c_1\delta(x,y) \leq \sup_{\lambda\in S_x^\star K} \lambda(\overrightarrow{xy})$.
D'après la proposition 2.6.4, il existe une constante $C' > 0$ telle que pour tous $a \in K$ et $\lambda \in S_a^\star K$, les fonctions $x \mapsto \lambda(\overrightarrow{ax}) + C'h_a(x)$ soient convexes.
Démontrons d'abord la proposition lorsque v est égal à e_μ, le barycentre exponentiel de μ, unique d'après la proposition (3.1.2)).
Soient x un élément de $b(\mu)$ et $\lambda \in S_{e_\mu}^\star K$. On a

$$\lambda(\overrightarrow{e_\mu x}) + C'h_{e_\mu}(x) \leq \int_K \mu(dy)(\lambda(\overrightarrow{e_\mu y}) + C'h_{e_\mu}(y)).$$

On a

$$\int_K \mu(dy)(\lambda(\overrightarrow{e_\mu y})) = \lambda(\int_K \mu(dy)\overrightarrow{e_\mu y}) = 0$$

puisque e_μ est le barycentre exponentiel de μ.
D'autre part, $C'h_{e_\mu}(x) \geq 0$,donc on obtient

$$\lambda(\overrightarrow{e_\mu x}) \leq C' \int_K \mu(dy)h_{e_\mu}(y).$$

L'inégalité est encore valable avec à gauche le supremum sur $\lambda \in S^\star_{e_\mu}K$, ce qui permet ensuite d'obtenir $c_1\delta(e_\mu, x) \leq C' \int_K \mu(dy)h_{e_\mu}(y)$. Or h_a est positive, de classe C^∞ sur $K \times K \setminus \Delta$ et conïcide avec la distance au cube $\delta^3(a, x)$ dans un voisinage ouvert de la diagonale Δ et donc, il existe deux constantes c_δ C_δ telles que, pour tous a, x, on a

$$c_\delta\delta^3(a, x) \leq h_a(x) \leq C_\delta\delta^3(a, x).$$

En posant $C = \frac{2C'C_\delta}{c_1}$, on obtient bien l'inégalité

$$|b(\mu)| \leq C \int_K \delta^3(e_\mu, y)\, \mu(dy).$$

Si maintenant v est un élément quelconque de K, il suffit d'établir qu'il existe une constante C' indépendante de μ telle que $\int_K \delta^3(e_\mu, y)\, \mu(dy) \leq C' \int_K \delta^3(v, y)\, \mu(dy)$. Puisque h_v est convexe, on a $h_v(e_\mu) \leq \int_K h_v(y)\, \mu(dy)$, que l'on peut transformer en $\delta^3(v, e_\mu) \leq \frac{C_\delta}{c_\delta} \int_K \delta^3(v, y)\, \mu(dy)$.
De l'inégalité

$$\delta^3(e_\mu, y) \leq 4(\delta^3(e_\mu, v) + \delta^3(b, y))$$

pour tout $y \in K$, on déduit

$$\int_K \delta^3(e_\mu, y)\, \mu(dy) \leq 4\frac{C_\delta}{c_\delta} \int_K \delta^3(v, y)\, \mu(dy) + 4 \int_K \delta^3(v, y)\, \mu(dy).$$

Ceci achève la démonstration.

Chapitre 4

Résultats

On considère toujours une variété différentiable M, sans bord, de classe C^∞, munie d'une connexion affine, et telle que deux points quelconques de M sont toujours reliés par une géodésique et une seule, qui en outre dépend de façon C^∞ des deux points.
On supposera donnée une mesure de probabilité μ à support suffisamment petit, inclus dans un compact convexe K admettant une base de voisinages ouverts convexes. Le but est l'étude de l'influence de la structure de M (courbure ou tous autres éléments géométriquement significatifs) sur la taille du barycentre $b(\mu)$, qui a été défini et étudié au chapitre précédent.
Nous avons vu (proposition 3.1.6 et théorème 3.1.7) que l'on peut obtenir des points de $b(\mu)$ comme limites de barycentres géodésiques itérés de mesures approchant μ ; en construisant ainsi deux points de $b(\mu)$ suffisamment éloignés, on pourra obtenir des minorations de la taille de $b(\mu)$.
Inversement, des majorations de la taille de $b(\mu)$ peuvent être établies, comme cela a été fait dans la proposition 3.2.1, en construisant des fonctions convexes ; ces approximations seront d'autant plus précises que les fonctions convexes utilisées seront plus près d'être affines.

4.1 Encadrement du barycentre convexe d'une loi uniforme portée par une petite sphère S^2 dans une 3-sphère

Dans cette partie, nous donnons un encadrement du barycentre convexe de la loi de probabilité uniforme portée par une petite sphère S^2 dans la

3-sphère. La majoration est obtenue par construction de fonctions convexes presque affines et la minoration par construction d'une suite de mesures finies qui converge étroitement vers la loi uniforme portée par un petit cercle en dimension 2 puis en utilisant l'associativité des barycentres pour le passage à la dimension 3.

Proposition 4.1.1. *Soient* (M, g) *une variété riemannienne de dimension* 2, O *un point de* M, $T_O M$ *l'espace tangent en* O, k *la courbure de* M *en* O. *Pour toute constante réelle* $C > \frac{2}{9}|k|$, *et pour toute forme linéaire* λ, *de norme* 1, *définie sur* $T_O M$, *la fonction* $f(x) = \lambda(\overrightarrow{Ox}) + C\|\overrightarrow{Ox}\|^3$ *est convexe sur un voisinage de* O.

Démonstration :
La constante C et la forme linéaire λ sont fixées.
La fonction f est de classe C^2 sauf en O, en raison de la singularité à l'origine de la fonction r. Pour montrer que f est convexe sur une petite boule B centrée en O, il suffit de montrer que $Hessf$ est positive sur $B\backslash\{O\}$. En effet, cela assure que $fo\gamma$ est convexe pour toute géodésique γ dans $B\backslash\{O\}$; et d'autre, part pour les géodésiques γ passant par O, $fo\gamma$ est convexe parceque les fonctions d'une variable réelle $\xi \mapsto a\xi + C|\xi|^3$ sont convexes.
Pour calculer $Hessf$ sur $B\backslash\{O\}$, fixons une base orthonormée (e_1, e_2) de $T_O M$ telle que $\lambda(\overrightarrow{Ox}) = \langle e_1, \overrightarrow{Ox}\rangle$.
Dans le système de coordonnées normales (x^1, x^2) centré en O et associé à cette base, on a donc $\lambda(\overrightarrow{Ox}) = x^1$. On pose en outre $r^2 = (x^1)^2 + (x^2)^2 = \|\overrightarrow{Ox}\|^2$.
Dans ces coordonnées normales, le développement de Taylor à l'ordre 2 de la métrique g est donné au voisinage de O par

$$g_{ij}(x) = \delta_{ij} - \frac{1}{3}\langle R(e_i, e_m)e_n, e_j\rangle x^m x^n + o(r^2)$$

(voir par exemple ([48] p. 72–76), où le terme d'erreur $o(r^2)$ désigne une matrice $\varepsilon(x)$, non nécessairement radiale, mais telle que $\varepsilon(x)/r^2(x) \to 0$ quand $x \to O$. Il en résulte que les symboles de Christoffel Γ^l_{ij} vérifient $\Gamma^l_{ij}(O) = 0$ et $\frac{\partial \Gamma^l_{ij}}{\partial x^n}(O) = \frac{1}{3}(R_{lijn}(O) + R_{ljin}(O))$.
En utilisant les formules $(Hess\,\varphi)_{ij} = \frac{\partial^2 \varphi}{\partial x^i \partial x^j} - \Gamma^l_{ij}\frac{\partial \varphi}{\partial x^l}$, $\frac{\partial r}{\partial x^i} = \frac{x^i}{r}$, $\frac{\partial (r^3)}{\partial x^i} = 3rx^i$ et $\frac{\partial^2 (r^3)}{\partial x^i \partial x^j} = 3(\delta_{ij} r + \frac{x^i x^j}{r})$, en remplaçant $R_{lijn}(O)$ par $k(\delta_{lj}\delta_{in} - \delta_{ij}\delta_{ln})$ et en posant $h = \frac{k}{9C}$, on trouve que, hors de l'origine O, la matrice hessienne

de f dans les coordonnées (x^1, x^2) est donnée par

$$\frac{1}{3C} Hessf = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} - hx^2 \\ \frac{x^1x^2}{r} - hx^2 & r + \frac{(x^2)^2}{r}) + 2hx^1 \end{pmatrix} + o(r^2).$$

Nous allons vérifier que $Hessf(x)$ est non seulement positive, mais définie positive pour $x \in B\backslash\{O\}$, où B est un voisinage de O.
Puisque le premier terme diagonal $r + \frac{(x^1)^2}{r} + o(r^2)$ est supérieur strictement à zéro au voisinage de O (sauf O), il suffit que le déterminant le soit aussi. Or, en posant $x^1 = r\cos\theta$ et $x^2 = r\sin\theta$, ce déterminant vaut

$$r^2(2 + 4h\cos\theta - h^2\sin^2\theta + o(r)),$$

et l'hypothèse $|h| = \frac{|k|}{9C} < \frac{1}{2}$ assure qu'il est strictement positif au voisinage de O (hors O). $\square$

Corollaire 4.1.2. *Soient (M, g), O et k comme dans la proposition. Pour toute constante $C > \frac{2}{9}|k|$ il existe un voisinage V de O tel que pour toute forme linéaire λ sur T_0M de norme 1, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + C\|\overrightarrow{Ox}\|^3$ soit convexe sur V.*

Démonstration :
Fixons $C > \frac{2}{9}|k|$, et choisissons un $\alpha > 1$ tel que $C > \alpha\,\frac{2}{9}|k|$.

Posons $f_{C,\lambda}(x) = \lambda(\overrightarrow{Ox}) + C\|\overrightarrow{Ox}\|^3$. Par la proposition 4.1.1, pour chaque λ de norme $\|\lambda\| \leq \alpha$, la fonction $f_{C,\lambda} = \alpha\, f_{C/\alpha\,,\,\lambda/\alpha}$ est convexe au voisinage de O. Choisissons dans la boule $B(0, \alpha)$ du plan cotangent T^*_OM un ensemble fini Λ dont l'enveloppe convexe contient la boule unité $B(0, 1)$. Il existe un voisinage de l'origine sur lequel $f_{C,\lambda}$ est convexe pour tout $\lambda \in \Lambda$, donc aussi (l'application $\lambda \mapsto f_{C,\lambda}$ étant affine) pour tout λ dans l'enveloppe convexe de Λ, et a fortiori pour tout λ de norme 1. $\square$

Théorème 4.1.3. (M, g) *est une variété riemanienne de dimension* 2*, O est un point de M et k désigne la courbure de M en O. Alors*

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{|k|}{18},$$

$E(D)$ désignant l'ensemble des mesures de probabilité portées par la boule géodésique $B(O, \frac{D}{2})$ et de barycentre exponentiel O.

Ce théorème montre l'influence de la courbure de M sur la taille de $b(\mu)$ pour des mesures à support compact suffisamment petit.

Preuve du théorème : Soit C une constante supérieure à $\frac{2}{9}|k|$. D'après le corollaire 4.1.2, toutes les fonctions $f_{C,\lambda}(x) = \lambda(\overrightarrow{Ox}) + C\,r^3$, où $\|\lambda\| = 1$ et où $r = r(x) = \|\overrightarrow{Ox}\|$, sont convexes sur une même boule géodésique $B(O, \varepsilon/2)$.

Soit $D < \varepsilon$; soit μ une mesure de probabilité portée par la boule géodésique $B(O, \frac{D}{2})$, et ayant O pour barycentre exponentiel ; soit enfin x un point de $b(\mu)$.

Il existe λ de norme 1 telle que $r(x) = \lambda(\overrightarrow{Ox})$; pour un tel λ, on a

$$r(x) = \lambda(\overrightarrow{Ox}) \leq \lambda(\overrightarrow{Ox}) + C\,r(x)^3 = f_{C,\lambda}(x) \leq \mu(f_{C,\lambda}) \ .$$

Le calcul de $\mu(f_{C,\lambda})$ fait apparaître deux termes. Le premier est la moyenne selon μ de $y \mapsto \lambda(\overrightarrow{Oy})$; ce terme est nul car O est le barycentre exponentiel de μ. Le deuxième est la moyenne de $C\,r^3$; il est majoré par $C\,(D/2)^3$ car $r \leq D/2$ sur le support de μ. Il reste donc $r(x) \leq C\,D^3/8$. Ceci ayant lieu pour tout x dans $b(\mu)$, on a $|b(\mu)| \leq C\,D^3/4$. On a ainsi montré que

$$\forall\, C > \frac{2}{9}|k| \quad \exists\, \varepsilon > 0 \quad \forall\, D < \varepsilon \qquad \sup_{\mu \in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{C}{4} \ . \qquad \square$$

Encadrement du diamètre du barycentre convexe de la loi uniforme portée par un petit cercle sur l'hémisphère

Dans cette partie, la variété M sera une boule géodésique ouverte de centre O et de rayon $R < \frac{\pi}{2}$ dans la sphère unité de dimension 2. Pour $r \in\,]0, R[$, on note μ_r la probabilité uniforme sur le cercle de centre O et de rayon r. L'ensemble $b(\mu_r)$, qui est convexe et invariant par le groupe des rotations de centre O, est une boule géodésique de centre O dont nous allons estimer le rayon.

Théorème 4.1.4. *Si c et C sont des réels positifs tels que $c < \frac{\sqrt{3}}{8\pi}$ et $C > \frac{2}{9}$, il existe $r_0 \in\,]0, R[$ tel que pour tout $r \in\,]0, r_0[$, on ait la double inclusion :*

$$B(O, cr^3) \subset b(\mu_r) \subset B(O, Cr^3),$$

où $B(O, a)$ est la boule géodésique de centre O et de rayon a.

Preuve :
En raison de l'invariance de μ_r par les rotations de centre O, le point O est barycentre exponentiel de μ_r. La deuxième inclusion du théorème est alors une conséquence du théorème 4.1.3 avec $k = 1$.
Pour établir la première inclusion, on va considérer μ_r comme la limite étroite

de mesures uniformes sur les sommets de polygones réguliers à $2^n \times 3$ côtés. Soit ABC un triangle sphérique équilatéral dont les sommets sont sur le support de μ_r. On choisit une orientation sur l'hemisphère, et pour tout $\theta \in \mathbb{R}$, on note ρ_θ la rotation de centre O et d'angle θ, on note de même ρ_θ la rotation de centre O et d'angle θ sur l'espace tangent en O.
On définit par récurrence pour $n \geq 0$ une suite de probabilités ν_n par :

$$\nu_0 = \frac{1}{3}(\delta_A + \delta_B + \delta_C)\,, \qquad \nu_{n+1} = \frac{1}{2}\big(\nu_n + \rho_{\frac{\pi}{3.2^n}}(\nu_n)\big)\,.$$

Soit x_0 le barycentre de I milieu de (A, B) et C affectés des coefficients $\frac{2}{3}$ et $\frac{1}{3}$ respectivement. On définit par récurrence x_n tel que

$$\overrightarrow{Ox_{n+1}} = \frac{1}{2}\,\big(\overrightarrow{Ox_n} + \rho_{\frac{\pi}{3.2^n}}(\overrightarrow{Ox_n})\big).$$

On démontre par récurrence que le point x_n appartient au barycentre $b(\nu_n)$; en effet, si x_{n-1} et $\rho_{\frac{\pi}{3.2^{n-1}}}(x_{n-1})$ sont respectivement dans $b(\nu_{n-1})$ et $b(\rho_{\frac{\pi}{3.2^{n-1}}}(\nu_{n-1}))$, alors leur milieu x'_n appartient à $b(\nu_n)$; mais le point O est aussi dans $b(\nu_n)$, et comme M est incluse dans une sphère, le point x_n est sur le segment géodésique joignant O à x'_n, c'est donc un élément de $b(\nu_n)$.
On identifie isométriquement le plan $T_O M$ et le plan complexe $\mathbb{C}$, de façon que O soit identifié à 0. Pour tout n, le vecteur $\overrightarrow{Ox_n}$ est identifié à un complexe z_n, et on a la relation de récurrence

$$z_{n+1} = \frac{1}{2}(z_n + e^{\frac{i\pi}{3.2^n}} z_n) = \frac{1}{2}\,\frac{1 - e^{\frac{i\pi}{3.2^{n-1}}}}{1 - e^{\frac{i\pi}{3.2^n}}}\, z_n\,,$$

d'où l'on tire

$$z_n = z_0\,\frac{1}{2^n}\,\frac{1 - e^{\frac{2i\pi}{3}}}{1 - e^{\frac{i\pi}{3.2^{n-1}}}}.$$

Puisque $1 - e^{\frac{i\pi}{3.2^{n-1}}} \sim -\frac{i\pi}{3.2^{n-1}}$, on obtient que $z_n \to \frac{3\sqrt{3}}{2\pi} e^{\frac{i\pi}{3}}\, z_0$ lorsque $n \to \infty$, et x_n tend vers un point limite x_∞ tel que $d(O, x_\infty) = \frac{3\sqrt{3}}{2\pi}\, d(O, x_0)$. Comme ν_n converge étroitement vers μ_r, le point x_∞ est dans $b(\mu_r)$, et le rayon de $b(\mu_r)$ est donc au moins $\frac{3\sqrt{3}}{2\pi}\, d(O, x_0)$.
Soit $d(A, C) = s$ la longeur du côté du triangle sphérique ABC. On a d'une part, $\cos s = \cos^2 r + \sin^2 r \cos\frac{2\pi}{3}$ et d'autre part, $\cos d(C, I) \cos\frac{s}{2} = \cos s$; un équivalent de $d(C, I)$ lorsque r tend vers zéro est donc $\frac{3}{2} r(1 + \frac{r^2}{12})$, ce qui donne $d(C, x_0) \sim r + \frac{r^3}{12}$, d'où $d(O, x_0) \sim \frac{r^3}{12}$, et finalement $d(O, x_\infty) \sim \frac{\sqrt{3} r^3}{8\pi}$. $\square$

Voici une généralisation en dimension 3 de la proposition 4.1.1 et du corollaire 4.1.2, sur la convexité des fonctions $f_{C,\lambda}(x) = \lambda(\overrightarrow{Ox}) + C\|\overrightarrow{Ox}\|^3$.

Proposition 4.1.5. *Soient (M, g) une variété riemannienne de dimension 3, et O un point de M tel que toutes les courbures sectionnelles en O sont égales à une même valeur k.*

Pour tout $C > \frac{1+\sqrt{2}}{9}|k|$, il existe un voisinage de O sur lequel toutes les fonctions $f_{C,\lambda}$ sont convexes, où λ décrit la sphère unité de l'espace cotangent T_O^*M.

Démonstration :
Comme en dimension 2, il suffit de montrer séparément que chaque $f_{C,\lambda}$ est convexe au voisinage de O, et cela va se ramener à une minoration de la hessienne.

On fixe C et λ, et on écrit f pour $f_{C,\lambda}$. On choisit une base orthonormée (e_1, e_2, e_3) de T_OM telle que $\lambda(\overrightarrow{Ox}) = \langle e_1, \overrightarrow{Ox}\rangle$.
Dans le système de coordonnées normales (x^1, x^2, x^3) centré en O et associé à cette base, on a donc $\lambda(\overrightarrow{Ox}) = x^1$. Posons $r = r(x) = \|\overrightarrow{Ox}\|$.
En posant $h = \frac{k}{9C}$, on trouve que, hors de l'origine O, la matrice hessienne de f dans les coordonnées (x^1, x^2, x^3) est donnée par

$$\frac{1}{3C} Hessf = \begin{bmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} - hx^2 & \frac{x^1x^3}{r} - hx^3 \\ \frac{x^1x^2}{r} - hx^2 & r + \frac{(x^2)^2}{r} + 2hx^1 & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} - hx^3 & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} + 2hx^1 \end{bmatrix} + o(r^2) \ .$$

Nous allons vérifier que $Hessf(x)$ est définie positive pour $x \in B\backslash\{O\}$, où B est un voisinage de O.
Puisque le premier terme diagonal $r + \frac{(x^1)^2}{r}$ est supérieur strictement à zéro au voisinage de O (sauf O), il suffit que les deux autres mineurs principaux le soient aussi.
En posant $x^1 = r\cos\theta\sin\varphi$, $x^2 = r\sin\theta\sin\varphi$ et $x^3 = r\cos\varphi$ ces mineurs valent

$$M_2 = r^2[2 - \cos^2\varphi + 2\cos\theta\sin\varphi(2 - \cos^2\varphi)h - (\sin^2\varphi\sin^2\theta)h^2 + o(r)]$$

et

$$M_3 = r^3[2+8(\cos\theta\sin\varphi)h-(1-9\cos^2\theta\sin^2\varphi)h^2+2\cos\theta\sin\varphi(\cos^2\theta\sin^2\varphi-1)h^3+o(r)].$$

L'hypothèse $|h| = \frac{|k|}{9C} < \sqrt{2} - 1$ assure que $M_2 > 0$ et $M_3 > 0$ au voisinage de O (hors O).

Exactement comme en dimension 2, la proposition 4.1.5 entraîne :

Théorème 4.1.6. *Soient (M, g) une variété riemannienne de dimension 3, et O un point de M tel que toutes les courbures sectionnelles en O sont égales à une même valeur k.*

On a la majoration :

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{1+\sqrt{2}}{36}\,|k|,$$

où $E(D)$ désigne l'ensemble des probabilités portées par la boule géodésique $B(O, \frac{D}{2})$ et dont le barycentre exponentiel est O.

Encadrement du diamètre du barycentre convexe de la loi uniforme portée par une petite sphère S^2 dans la 3-sphère.

La variété M est une boule géodésique régulière de rayon R dans la 3-sphère unité ; soit O le centre de cette boule. Pour $r \in]0, R[$ on appelle μ_r la probabilité uniforme sur la sphère de centre O et de rayon r.
L'ensemble $b(\mu_r)$ est une boule géodésique centrée en O ; nous allons estimer son rayon.

Théorème 4.1.7. *Si c et C sont des réels positifs tels que $c < \frac{\sqrt{3}}{32}$ et $C > \frac{1+\sqrt{2}}{9}$, alors il existe $r_0 \in]0, R[$ tel que pour tout $r \in]0, r_0[$, on ait la double inclusion :*

$$B(O, cr^3) \subset b(\mu_r) \subset B(O, Cr^3).$$

Preuve du théorème :

La deuxième inclusion résulte du théorème 4.1.6, avec courbure constante $k = 1$; la première est le cas particulier $n = 3$ du théorème 4.1.8 ci-dessous. $\square$

Minoration du diamètre du barycentre convexe de la loi uniforme portée par une petite sphère S^{n-1} dans la n-sphère.

La variété M est maintenant de dimension n et à courbure constante 1 ; plus précisément, M est une boule géodésique régulière de rayon R et de centre O dans la n-sphère unité. Nous allons minorer le rayon du barycentre convexe $b(\mu_r)$, où μ_r est la probabilité uniforme sur la sphère géodésique de petit rayon r et de centre O. Nous utiliserons la suite des nombres

$$I_k = \int_0^{\frac{\pi}{2}} (\cos\alpha)^k \, d\alpha = \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{k}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{k}{2}+1\right)} \, .$$

Théorème 4.1.8. *Pour tout c tel que $0 < c < \frac{\sqrt{3}}{8\pi} I_{n-1}$ il existe $r_0 \in]0, R[$ tel que pour tout $r \in]0, r_0[$, on ait la minoration*

$$B(O, cr^3) \subset b(\mu_r).$$

Démonstration :
Pour $n = 2$, on a $I_{n-1} = 1$, et la minoration a été établie dans le théorème 4.1.4.

Nous allons démontrer la propriété dans le cas $n > 2$ par récurrence sur n, en la supposant vraie pour $n - 1$. Fixons $c < \frac{\sqrt{3}}{8\pi} I_{n-1}$, et posons $c' = \frac{I_{n-2}}{I_{n-1}} c$, de sorte que la propriété au rang $n-1$ a lieu avec cette valeur c'. Pour démontrer le théorème, il suffit de construire, pour tout r assez petit, un point x de $b(\mu_r)$ tel que $d(O, x) \geq cr^3 = \frac{I_{n-1}}{I_{n-2}} c' r^3$, car toute la boule $B(O, cr^3)$ sera alors incluse dans $b(\mu_r)$.

Considérons G_r l'ensemble des grands cercles de la petite sphère support de μ_r ; ce sont les intersections de cette sphère avec les sous-variétés totalement géodésiques de codimension 1 passant par O. Notons α_r la mesure de probabilité uniforme sur G_r et pour tout $g \in G_r$, appelons ν_g la mesure de probabilité uniforme sur g. Étant invariante par rotations, la probabilité $\int_{G_r} \nu_g \, \alpha_r(dg)$ est égale à μ_r.

Chaque $g \in G_r$ est inclus dans une hypersurface totalement géodésique H_g, passant par O, et à courbure constante 1. Appliquons dans H_g l'hypothèse de récurrence : puisque $c' < \frac{\sqrt{3}}{8\pi} I_{n-2}$, il existe $r_0 \in]0, R[$ tel que pour tout $r \in]0, r_0[$ et tout cercle $g \in G_r$, le barycentre convexe $b(\nu_g)$ de ν_g contient la boule géodésique $B(O, c'r^3)$ de H_g, c'est-à-dire l'intersection $\hat{g}$ de H_g avec la boule géodésique n-dimensionnelle $B(O, c'r^3)$. Ce $\hat{g}$ est la boule centrée en O, de rayon $c'r^3$, située dans H_g. Lus dans la carte exponentielle en O, g et $\hat{g}$ deviennent une sphère de rayon r et une boule de rayon $c'r^3$, dans un même hyperplan passant par O.

Fixons dans $T_O M$ une direction arbitraire, dite “verticale”, et appelons $h(g)$ le point de $\hat{g}$ le plus “haut” dans cette direction. (Lorsque g est horizontal, $h(g)$ n'est pas défini.) Quand g parcourt G_r, $h(g)$ décrit un hémisphère supérieur S_r, centré en O et de rayon $c'r^3$. Notons $\hat{\alpha}_r$ la mesure image de α_r par l'application h ; c'est une certaine probabilité sur S. Par associativité des barycentres (proposition 3.1.4) et en remarquant que $h(g) \in \hat{g} \subset b(\nu_g)$, on a $b(\hat{\alpha}_r) \subset b(\mu_r)$; pour achever la preuve, il suffit donc de trouver dans $b(\hat{\alpha}_r)$ un point à distance au moins cr^3 de O.

Le grand cercle g étant choisi uniformément, son axe coupe S_r en un point $p(g)$ de loi uniforme sur S_r, et la latitude θ de $p(g)$ (l'angle de $Op(g)$ avec l'horizontale) suit la loi de densité $(1/I_{n-2}) \, (\cos\theta)^{n-2} d\theta$. La latitude de

$h(g)$ est $\lambda = \frac{\pi}{2} - \theta$, et le centre de la $(n-1)$-sphère formée des points de S_r de latitude λ fixée se trouve au-dessus de O, à distance supérieur à $c'r^3 \sin\lambda$ (il y aurait égalité en courbure nulle ; l'inégalité vient de la courbure positive). La moyenne de ces centres fournit un point x de $b(\hat{\alpha}_r)$ tel que

$$d(O,x) > \int_0^{\frac{\pi}{2}} c'r^3 \sin\left(\tfrac{\pi}{2} - \theta\right) \frac{(\cos\theta)^{n-2}\, d\theta}{I_{n-2}} = \frac{I_{n-1}}{I_{n-2}}\, c'r^3 = cr^3 \ .$$

Ceci achève la démonstration. □

4.2 Minoration du barycentre convexe d'une loi uniforme portée par une petite sphère S^{n-1} dans l'espace hyperbolique H^n

Considérons la mesure uniforme μ porté par un petit cercle C_r de centre o et de rayon r dans l'espace hyperbolique de dimension 2. Cette mesure est invariante sous l'action du groupe orthogonal des isométries fixant o.
Fixons un arc $\mathfrak{C}$ de C_r, de longeur $\frac{1}{4}$ du périmètre de C_r ; appelons λ la probabilité uniforme sur l'arc $\mathfrak{C}$. Chaque carré inscrit dans C_r (c'est-à-dire dont les quatre sommets sont sur C_r) a un sommet x appartenant à $\mathfrak{C}$. Si $ABCx$ est un tel carré, on notera $\nu_x = \frac{1}{4}(\delta_A + \delta_B + \delta_C + \delta_x)$ la mesure uniforme sur les sommets du carré $ABCx$; on a alors la désintégration $\mu = \int_{x\in\mathfrak{C}} \nu_x\, \lambda(dx)$.

Pour $x \in \mathfrak{C}$, nous allons définir un certain barycentre géodésique itéré b_x de la probabilité ν_x. En appelant I (resp. J) le milieu du segment géodésique AB (resp. BC), le milieu L de IJ est un barycentre géodésique itéré de la mesure $\frac{1}{4}\varepsilon_A + \frac{1}{2}\varepsilon_B + \frac{1}{4}\varepsilon_C$. Puisque o est le milieu du segment AC, le milieu N de ox est un barycentre géodésique itéré de la mesure $\frac{1}{4}\varepsilon_A + \frac{1}{4}\varepsilon_C + \frac{1}{2}\varepsilon_x$. Finalement, le milieu b_x de LN est dans le barycentre $b(\nu_x)$ de la mesure uniforme sur les quatre sommets du carré.

Les formules de la géométrie hyperbolique donnent un équivalent de la distance $\rho(r) = ob_x$ au voisinage de zéro : lorsque $r \to 0$, on a $\rho(r) \sim \frac{1}{8}\, r^3$.

Notons $\widehat{\mathfrak{C}}$ le quart de cercle centré en o décrit par les points b_x lorsque x parcourt $\mathfrak{C}$. En appelant $\psi : \mathfrak{C} \to \widehat{\mathfrak{C}}$ l'application $x \mapsto b_x$, on peut définir la mesure image $\widehat{\lambda} = \psi(\lambda)$; c'est la probabilité uniforme sur l'arc de cercle $\widehat{\mathfrak{C}}$, mais nous utiliserons uniquement le fait qu'elle est portée par $\widehat{\mathfrak{C}}$.

Puisque b_x est dans $b(\nu_x)$ et que $\mu = \int_{x\in\mathfrak{C}} \nu_x\, \lambda(dx)$, l'associativité (proposition 3.1.4) entraîne que le barycentre $b(\widehat{\lambda})$ est inclus dans $b(\mu)$. Mais il est aussi inclus dans l'enveloppe convexe du support $\widehat{\mathfrak{C}}$ de $\widehat{\lambda}$; on peut donc minorer le rayon du disque $b(\mu)$ par le distance de o à l'enveloppe convexe de $\widehat{\mathfrak{C}}$. Un calcul de trigonométrie hyperbolique montre que cette distance est $r_1 = \alpha(\rho)$, où $\rho = ob_x$ est le rayon du quart de cercle $\mathfrak{C}$, et où la fonction α est donnée par

$$\mathrm{sh}\,(\alpha(\rho)) = \frac{\mathrm{sh}\,\rho}{\sqrt{2 + \mathrm{sh}^2 \rho}} \ .$$

On obtient ainsi que $b(\mu)$ contient la boule géodésique de rayon $r_1 = \alpha(\rho(r))$. Lorsque $r \to 0$, on a l'équivalent $r_1 \sim \frac{1}{8\sqrt{2}}\, r^3$.

Le passage à la dimension $n > 1$ se fait de façon analogue au cas sphérique ; dans ce cas μ est la mesure uniforme sur une petite sphère S^{n-1} plongée dans l'espace hyperbolique de dimension n.

On montre que le barycentre $b(\mu)$ contient la boule géodésique de centre o et de rayon r_n défini par la relation de récurrence

$$r_{n+1} = \int_0^1 \mathrm{Arg\,th}\,(\sqrt{1-s^2}\ \ \mathrm{th}\, r_n)\, ds \ .$$

Pour n fixé, lorsque r tend vers zéro, on a $r_n \sim \left(\frac{\pi}{4}\right)^{n-1} \frac{1}{8\sqrt{2}}\, r^3$.

4.3 Majorations asymptotiques du barycentre convexe d'une mesure de probabilité sur les espaces homogènes $S^2 \times \mathbb{R}$, $H^2 \times \mathbb{R}$ et l'espace de Heisenberg H^3

Dans cette partie, nous donons des majorations asymptotiques du barycentre convexe d'une mesure de probabilité à support compact dans les espaces homogènes à courbure non constante $S^2 \times \mathbb{R}$, $H^2 \times \mathbb{R}$ et l'espace de Heisenberg H^3.

Une variété riemanienne est dite homogène si son groupe d'isométries agit transitivement, c'est-à-dire si, pour tout couple (p, q) de points, il existe une isométrie qui envoie p sur q.

Un espace homogène de dimension 3 avec groupe d'isométries de dimension 4 est une fibration riemanienne sur une variété simplement connexe de dimension 2 à courbure constante k dont les fibres sont géodésiques ; nous noterons τ la courbure de fibration ($k - 4\tau^2 \neq 0$).

On envisage deux cas :

Lorsque $\tau = 0$, la variété est l'un des deux espaces produits $S^2 \times \mathbb{R}$ et $H^2 \times \mathbb{R}$,

où S^2 est la sphère ronde ($k > 0$) et H^2 le plan hyperbolique ($k < 0$).
Quant au cas $\tau \neq 0$, il regroupe les trois types suivants : Les variétés ayant le groupe d'isométries des sphères de Berger, l'espace de Heisenberg et le revêtement universel de $PSL_2(\mathbb{R})$.
Si $k \geq 0$ (respectivement $k \leq 0$), on considère sur $\mathbb{R}^3$ (respectivement $\mathbb{D}^2(\frac{2}{\sqrt{-k}}) \times \mathbb{R}$) la métrique riemannienne g définie par :

$$ds^2 = \lambda^2(dx^2 + dy^2) + (\tau\lambda(ydx - xdy) + dz)^2,$$

où

$$\lambda = \frac{1}{1 + \frac{k}{4}(x^2 + y^2)}$$

et $\mathbb{D}^2$ désigne le disque de Poincaré.
Lorsque $\tau \neq 0$ la variété a la géométrie des sphères de Bergers si $k > 0$, celle du groupe de Heisenberg si $k = 0$ et celle du recouvrement universel de $PSL_2(\mathbb{R})$ si $k < 0$. Dans ces cas on définit le repère canonique (E_1, E_2, E_3) par :

$$E_1 = \frac{1}{\lambda}[cos(\sigma z)\frac{\partial}{\partial x} + sin(\sigma z)\frac{\partial}{\partial y}] + \tau(xsin(\sigma z) - ycos(\sigma z))\frac{\partial}{\partial z}$$

$$E_2 = \frac{1}{\lambda}[-sin(\sigma z)\frac{\partial}{\partial x} + cos(\sigma z)\frac{\partial}{\partial y}] + \tau(xcos(\sigma z) + ysin(\sigma z))\frac{\partial}{\partial z}$$

et

$$E_3 = \xi = \frac{\partial}{\partial z}$$

avec

$$\sigma = \frac{k}{2\tau}\ .$$

Ce repère est orthonormé et on a :
$[E_1, E_2] = 2\tau E_3$; $[E_2, E_3] = \frac{k}{2\tau}E_1$ et $[E_3, E_1] = \frac{k}{2\tau}E_2$.
On vérifie [12] que les courbures sectionnelles des plans (E_2, E_3), (E_1, E_3) et (E_1, E_2) sont respectivement : τ^2, τ^2 et $k - 3\tau^2$.
Lorsque $\tau = 0$ et $k \neq 0$ (cas des espaces produits), on prend comme repère orthonormé $E_1 = \frac{1}{\lambda}\frac{\partial}{\partial x}$, $E_2 = \frac{1}{\lambda}\frac{\partial}{\partial y}$ et $E_3 = \frac{\partial}{\partial z}$.
On vérifie de même que les courbures sectionnelles des plans (E_2, E_3), (E_1, E_3) et (E_1, E_2) sont respectivement : 0, 0 et k.

Proposition 4.3.1. *(M, g) est l'espace produit $S^2 \times \mathbb{R}$ muni de la métrique g citée ci-dessus, O est un point de M et $T_O M$ est l'espace tangent en O.*

Pour toute constante $C > \frac{1+\sqrt{2}}{9}R$ et pour toute forme linéaire λ, de norme 1, définie sur T_OM ; la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ est convexe sur un voisinage de O.
R est la courbure sectionnelle du plan (E_1, E_2) défini ci-dessus et $r = r(x) = \|Ox\|$ est la distance riemanienne de O à x.

Démonstration :
La preuve est analogue à celle de la proposition (4.1.1) ; on trouve que, hors de l'origine O, la matrice hessienne de f dans les coordonnées (x^1, x^2, x^3) est donnée par

$$\frac{1}{3C}Hessf = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} - hx^2 & \frac{x^1x^3}{r} \\ \frac{x^1x^2}{r} - hx^2 & r + \frac{(x^2)^2}{r} + 2hx^1 & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} \end{pmatrix} + o(r^2).$$

où $h = \frac{k}{9C}$
Nous allons vérifier que $Hessf(x)$ est non seulement positive, mais définie positive pour $x \in B\backslash\{O\}$, où B est un voisinage de O.
Puisque le premier terme diagonal $r + \frac{(x^1)^2}{r} + o(r^2)$ est supérieur strictement à zéro au voisinage de O (sauf O), il suffit que les deux autres mineurs principaux dominants le soient aussi.
Ceci est assuré sous l'hypothèse $h < \sqrt{2} - 1$. □

Proposition 4.3.2. *(M, g) est l'espace produit $H^2 \times \mathbb{R}$ muni de la métrique g citée ci-dessus, O est un point de M et T_OM est l'espace tangent en O. Pour toute constante $C > -\frac{1+\sqrt{2}}{9}R$ et pour toute forme linéaire λ, de norme 1, définie sur T_OM, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ est convexe sur un voisinage de O.*
R est la courbure sectionnelle du plan (E_1, E_2) défini ci-dessus et $r = r(x) = \|Ox\|$ est la distance riemanienne de O à x.

La preuve est analogue à celle de la proposition (4.3.1) en remplaçant R par $-R$.

Proposition 4.3.3. *(M, g) est l'espace de Heisenberg de dimension 3 muni de la métrique g citée ci-dessus, O est un point de M et T_OM est l'espace tangent en O. Pour toute constante $C > \frac{7+\sqrt{65}}{18}\tau^2$ et pour toute forme linéaire λ, de norme 1, définie sur T_OM ; la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ est convexe sur un voisinage de O.*

τ est la courbure de fibration et $r = r(x) = \|Ox\|$ est la distance riemanienne de O à x.

Preuve :
Dans ce cas les courbures sectionnelles sont $R_{1212} = -3\tau^2$ et $R_{1313} = R_{2323} = \tau^2$, où R_{ijij} est la courbure sectionnelle du plan (E_i, E_j), $(i, j = 1, 2, 3)$. Hors de l'origine O, la matrice hessienne de f dans les coordonnées (x^1, x^2, x^3) est donnée par

$$\frac{1}{3C}Hessf = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} + x^2\frac{\tau^2}{3C} & \frac{x^1x^3}{r} - x^3\frac{\tau^2}{9C} \\ \frac{x^1x^2}{r} + x^2\frac{\tau^2}{9C} & r + \frac{(x^2)^2}{r} - 2x^1\frac{\tau^2}{3C} & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} - x^3\frac{\tau^2}{9C} & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} + 2x^1\frac{\tau^2}{9C} \end{pmatrix} + o(r^2)$$

Posons $\frac{\tau^2}{9C} = h$

$$\frac{1}{3C}Hessf = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} + 3x^2h & \frac{x^1x^3}{r} - x^3h \\ \frac{x^1x^2}{r} + 3x^2h & r + \frac{(x^2)^2}{r} - 6x^1h & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} - x^3h & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} + 2x^1h \end{pmatrix} + o(r^2)$$

Pour que $Hessf$ soit définie positive sur $B\backslash\{O\}$ il suffit que $h < \frac{\sqrt{65}-7}{8}$.

Corollaire 4.3.4. *Soient (M, g), O et R comme dans la proposition 4.3.1. Pour toute constante $C > \frac{1+\sqrt{2}}{9}R$ il existe un voisinage V de O tel que pour toute forme linéaire λ sur T_0M de norme 1, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ soit convexe sur V.*

Preuve : analogue au corollaire (4.1.2).

On peut énoncer de même :

Corollaire 4.3.5. *Soient (M, g), O et R et τ comme dans la proposition 4.3.1 (respectivement 4.3.3).
Pour toute constante $C > -\frac{1+\sqrt{2}}{9}R$ (respectivement $C > \frac{7+\sqrt{65}}{18}\tau^2$), il existe un voisinage V de O tel que pour toute forme linéaire λ sur T_0M de norme 1, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ soit convexe sur V.*

Théorème 4.3.6. *Avec les notations précédentes (M, g), O, R et τ, on a les majorations asymptotiques suivantes : Si M est le produit $S^2 \times \mathbb{R}$ alors*

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{1+\sqrt{2}}{36}R$$

Si M est le produit $H^2 \times \mathbb{R}$ alors

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq -\frac{1+\sqrt{2}}{36} R$$

Si M est l'espace de Heisenberg , alors

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{7+\sqrt{65}}{36} \tau^2$$

$E(D)$ désignant l'ensemble des mesures de probabilité portées par la boule géodésique $B(O, \frac{D}{2})$ et de barycentre exponentiel O.

Preuve du théorème : Soit C une constante supérieure à $\frac{1+\sqrt{2}}{9}R$. D'après le corollaire 4.3.4, toutes les fonctions $f_{C,\lambda}(x) = \lambda(\overrightarrow{Ox}) + Cr^3$, où $\|\lambda\| = 1$, sont convexes sur une même boule géodésique $B(O, \varepsilon/2)$.

Soit $D < \varepsilon$; soit μ une mesure de probabilité portée par la boule géodésique $B(O, \frac{D}{2})$, et ayant O pour barycentre exponentiel ; soit enfin x un point de $b(\mu)$.

Il existe λ de norme 1 telle que $r(x) = \lambda(\overrightarrow{Ox})$; pour un tel λ, on a

$$r(x) = \lambda(\overrightarrow{Ox}) \leq \lambda(\overrightarrow{Ox}) + C\, r(x)^3 = f_{C,\lambda}(x) \leq \mu(f_{C,\lambda}) \ .$$

Le calcul de $\mu(f_{C,\lambda})$ fait apparaître deux termes. Le premier est la moyenne selon μ de $y \mapsto \lambda(\overrightarrow{Oy})$; ce terme est nul car O est le barycentre exponentiel de μ. Le deuxième est la moyenne de $C\, r^3$; il est majoré par $C\,(D/2)^3$ car $r \leq D/2$ sur le support de μ. Il reste donc $r(x) \leq C\, D^3/8$. Ceci ayant lieu pour tout x dans $b(\mu)$, on a $|b(\mu)| \leq C\, D^3/4$. On a ainsi montré que

$$\forall\, C > \frac{1+\sqrt{2}}{9} R \quad \exists\, \varepsilon > 0 \quad \forall\, D < \varepsilon \qquad \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leq \frac{C}{4} \ . \qquad \square$$

Pour les espaces $H^2 \times \mathbb{R}$ et celui de Heisenberg la preuve est analogue en remplaçant le nombre $\frac{1+\sqrt{2}}{9}R$ par $-\frac{1+\sqrt{2}}{9}R$ puis par $\frac{7+\sqrt{65}}{9}\tau^2$.

Chapitre 5

Conclusion et perspectives

5.1 Majorations, minorations et bord du barycentre convexe

Les majorations établies dans le chapitre précédent reposent toutes sur le même principe. On sait (voir la démonstration de la proposition 3.2.1) qu'au voisinage d'un point a sur une variété riemannienne, les fonctions $\lambda(\overrightarrow{ax}) + C\|\overrightarrow{ax}\|^3$ sont convexes pour toute forme linéaire λ de norme 1 sur T_aM et toute constante C assez grande. Nous avons cherché, dans les différents cas étudiés, à expliciter en fonction de la géométrie de la variété une valeur de C aussi petite que possible. Les résultats obtenus ne sont pas optimaux, d'une part parce que nous avons souvent dû procéder par conditions suffisantes uniquement, donc sans parvenir à la meilleure constante C possible, mais d'autre part aussi parce nous nous sommes ainsi restreint à ne travailler qu'avec une famille très pauvre, à un paramètre (la constante C), de fonctions convexes.

Pour chaque minoration, nous avons d'abord construit deux barycentres géodésiques itérés d'une même mesure auxiliaire à support fini, de façon qu'ils soient aussi éloignés que possible l'un de l'autre ; la minoration ainsi obtenue a ensuite été transférée à la mesure donnée par associativité et intégration (parfois itérée pour changer de dimension) ou par passage à la limite. Ici aussi, chaque étape de l'argument éloigne un peu plus de l'optimalité et on ne sait pas contrôler la perte de précision introduite dans le raisonnement.

Mesurer exactement la taille du barycentre convexe $b(\mu)$ d'une mesure μ revient à savoir localiser les points du bord $\partial b(\mu)$. La proposition suivante donne une condition suffisante pour qu'un point de $b(\mu)$ soit sur la frontière.

Proposition 5.1.1. *Soit M une variété convexe, K un compact convexe de M admettant une base de voisinages convexes, et μ une mesure de probabilité portée par K. Soit x un point de $b(\mu)$ et f une fonction de $C(K)$, non constante au voisinage de x, et telle que $f(x) = \mu(f)$. Alors $x \in \partial b(\mu)$.*

Si de plus μ est à support fini et x un barycentre géodésique itéré de μ, alors la fonction f a une restriction affine à chacun des segments géodésiques intervenant dans la représentation de x comme barycentre géodésique itéré de μ.

Preuve :
Pour établir que $x \in \partial b(\mu)$, il suffit de montrer que tout un voisinage ouvert convexe V de x contient un point hors de $b(\mu)$. Par hypothèse, la fonction convexe f n'est pas constante sur V, donc (théorème 2.4.2) f n'atteint pas en x son maximum sur V ; il existe ainsi un $y \in V$ tel que $f(y) > f(x)$. Puisque $f(x) = \mu(f)$ par hypothèse, on a $f(y) > \mu(f)$, ce qui montre que y n'est pas dans $b(\mu)$.

Si maintenant x est un barycentre géodésique itéré de μ, l'inégalité de convexité $f(\gamma_{uv}(t)) \leq (1-t)f(u) + tf(v)$ écrite pour chaque segment géodésique de ce barycentre géodésique itéré donne finalement $f(x) \leq \mu(f)$. Puisque cette inégalité est par hypothèse une égalité, chacune des inégalités partielles est aussi une égalité, et pour chaque segment géodésique γ_{uv} il existe un $t \in [0,1]$ tel que $f(\gamma_{uv}(t)) = (1-t)f(u) + tf(v)$. Comme $f \circ \gamma_{uv}$ est convexe, ceci implique que $f \circ \gamma_{uv}$ est affine sur l'intervalle $[0,1]$. $\square$

Réciproquement, si x est un point du bord de $b(\mu)$, on peut trouver arbitrairement près de x des points y qui ne sont pas dans $b(\mu)$; pour un tel y, il existe une fonction g convexe sur un voisinage de K et telle que $g(y) > \mu(f) \geq g(x)$. En faisant tendre y vers x, et en renormalisant g, on peut espérer faire tendre g vers une fonction convexe f, non constante au voisinage de x, telle que $\mu(f) = f(x)$. Cette fonction limite devrait être définie sur K mais il n'y a pas de raison de croire qu'elle puisse toujours être prolongée en une fonction convexe au voisinage de K.

Conjecture 5.1.2. *Soit dans une variété convexe M un compact convexe K admettant une base de voisinages convexes, et x un point de la frontière de $b(\mu)$. Il existe une fonction réelle f définie et convexe sur K, non constante au voisinage de x, et telle que $f(x) = \mu(f)$.*

5.2 Une configuration bien particulière

Devant la difficulté de décrire explicitement $b(\mu)$ même dans les cas simples envisagés dans le chapitre précédent, nous allons dans cette partie ébaucher la recherche du bord $\partial b(\mu)$, et illustrer la proposition 5.1.1 et la conjecture 5.1.2, dans un cas aussi simple que possible : la variété sera de dimension 2, à courbure constante, et μ ne sera portée que par 3 points.

Même dans ce cas, nous n'avons que peu de résultats ; cette partie vise seulement à présenter des idées qui pourront peut-être s'avérer utiles dans des recherches ultérieures.

Pour fixer les choses, la variété sera un plan hyperbolique (courbure -1) ; on aurait aussi pu se placer sur un hémisphère ouvert (courbure $+1$), ou étendre certains arguments à des situations plus générales, mais nous ne cherchons pas ici les hypothèses minimales.

La variété ambiante est donc le plan hyperbolique H. On fixe dans H trois points A, B et C formant les sommets d'un triangle équilatéral. On appelle K l'enveloppe convexe de ces trois points ; c'est le triangle dont les côtés sont les segments géodésiques BC, CA et AB. La mesure μ est la probabilité uniforme $\frac{1}{3}(\varepsilon_A + \varepsilon_B + \varepsilon_C)$.

Appelons A', B' et C' les milieux respectifs des segments géodésiques BC, CA et AB. Par symétrie, le centre O du triangle, intersection des trois médianes AA', BB' et CC', est le barycentre exponentiel de μ ; c'est donc un point de $b(\mu)$ (proposition 3.1.1).

Le point $M = \gamma_{AA'}(\frac{2}{3})$, situé sur AA' aux deux tiers en partant de A, est lui aussi dans $b(\mu)$, par la proposition 3.1.6, car c'est un barycentre géodésique itéré de μ. (Formellement, c'est $\beta_2(y,p)$, où

$$y(0,0) = y(0,1) = A, \quad y(1,0) = B, \quad y(1,1) = C,$$

$$p(0,0) = p(0,1) = \tfrac{1}{6}, \quad p(1,0) = p(1,1) = \tfrac{1}{3} \text{ .)}$$

Ce point M est plus précisément un point du bord $\partial b(\mu)$, comme on peut le vérfier en appliquant la proposition 5.1.1 à la fonction $f(x) =$ distance de x à la droite BC. Cette fonction est convexe sur H d'après la proposition 2.2.3. En outre, f n'est nulle part localement constante, car $\|\nabla f\| = 1$ hors de la droite BC ; enfin, par construction de M, on a

$$f(M) = \tfrac{1}{3}\, f(A) = \tfrac{1}{3}\,[f(A) + f(B) + f(C)] = \mu(f).$$

On peut dire un peu plus : f étant convexe au voisinage de K, tout point x de $b(\mu)$ vérifie $f(x) \leq \mu(f)$, c'est-à-dire que la distance de x à BC est au maximum MA' ; en particulier, l'intersection de $b(\mu)$ avec le segment AA' est incluse dans le segment MA'.

Quel est le point de $b(\mu) \cap AA'$ le plus proche de A' ? Posons $I = \gamma_{AB}(\frac{2}{3})$ et $J = \gamma_{AC}(\frac{2}{3})$, et appelons N le milieu $\gamma_{IJ}(\frac{1}{2})$ de IJ. Ce point N est un barycentre géodésique itéré de μ ; les théorémes de comparaison en géométrie hyperbolique permettent de voir que les points A', N, O, M et A sont distincts et situés dans cet ordre sur le segment $A'A$. Mais des simulations numériques indiquent que N n'est pas sur le bord de $b(\mu)$; ce n'est donc pas le point de $b(\mu) \cap AA'$ le plus proche de A'. Pour rechercher ce point, nous allons nous intéresser à une classe particulière de fonctions convexes.

5.3 Des fonctions convexes presque affines

Nous restons dans le cadre où K est un triangle équilatéral de sommets A, B et C dans le plan hyperbolique H, et où la probabilité μ est portée par les trois points A, B et C ; mais nous ne supposons plus μ uniforme sur ces trois points. Il s'agit de rechercher les points du bord de $b(\mu)$.

Soit $x \in \partial b(\mu)$. Si la conjecture 5.1.2 est exacte, il existe une fonction convexe $f : K \to \mathbb{R}$, non constante au voisinage de x, telle que $f(x) = \mu(f)$. Si g est n'importe quelle autre fonction convexe au voisinage de K qui coïncide avec f sur les trois sommets A, B et C, on a d'une part $g(x) \leq \mu(g)$ parce que $x \in b(\mu)$, et d'autre part $\mu(g) = \mu(f) = f(x)$ parce que $g = f$ sur le support de μ ; donc $g(x) \leq f(x)$. Ainsi, f est la plus grande fonction convexe sur K prenant en A, B et C les valeurs $f(A)$, $f(B)$ et $f(C)$. Ceci suggère d'introduire une classe restreinte de fonctions convexes sur K, caractérisées par leurs valeurs sur les sommets de K.

Pour $(a, b, c) \in \mathbb{R}^3$, appelons φ_{abc} la plus grande fonction convexe φ définie sur K telle que $\varphi(A) = a$, $\varphi(B) = b$ et $\varphi(C) = c$. Cette fonction existe, car c'est aussi l'enveloppe convexe inférieure sur K de n'importe quelle fonction prenant les valeurs respectives a, b et c en A, B et C et des valeurs supérieures à $\max(a, b, c)$ sur $K \setminus \{A, B, C\}$. (Du moment qu'elles excèdent $\max(a, b, c)$, les valeurs sur $K \setminus \{A, B, C\}$ ne jouent aucun rôle, puisque tout point de K est un barycentre géodésique itéré des trois sommets A, B et C ; par exemple, tout point de K est sur le segment géodésique joignant A à un point du segment géodésique BC.) La valeur maximale de φ_{abc} sur K est $\max(a, b, c)$ et sa valeur minimale $\min(a, b, c)$; ces valeurs extrémales sont atteintes sur des sommets de K.

Le triangle K restant fixé, on désignera par Φ l'ensemble de toutes les fonctions φ_{abc} où (a, b, c) décrit $\mathbb{R}^3$.

Les fonctions $\varphi \in \Phi$ sont, par définition, convexes sur K. Elles sont donc aussi continues dans l'intérieur K° de K (corollaire 2.3.6). Nous ne savons pas si ces fonctions sont continues sur K ; a priori, à la frontière de K, elles sont seulement semi-continues supérieurement. Parmi toutes les fonctions convexes qui valent a, b et c en A, B et C, la fonction φ_{abc} est la plus grande, donc la moins convexe, ou encore la plus proche d'être affine.

Pour les raisons expliquées plus haut, l'appartenance à $b(\mu)$ devrait pouvoir être testée avec uniquement les fonctions de Φ, et les points du bord de $b(\mu)$ devraient être ceux pour lesquels on a de surcroît l'égalité $\varphi(x) = \mu(\varphi)$ pour au moins une fonction non constante $\varphi \in \Phi$.

Par ailleurs, il est clair que, si $u : \mathbb{R} \to \mathbb{R}$ est une fonction affine croissante, $u \circ \varphi_{abc} = \varphi_{u(a)\,u(b)\,u(c)}$. Ceci permet de se restreindre à la classe Φ' formée des fonctions $\varphi \in \Phi$ telles que $\min \varphi = 0$ et $\max \varphi = 1$, c'est-à-dire les fonctions φ_{abc} telles que $\min(a, b, c) = 0$ et $\max(a, b, c) = 1$. Topologiquement, l'ensemble E des triplets $(a, b, c) \in \mathbb{R}^3$ tels que $\min(a, b, c) = 0$ et $\max(a, b, c) = 1$ est un cercle. En outre, en raison de l'inégalité

$$a \leq a', \quad b \leq b', \quad c \leq c' \qquad \Longrightarrow \qquad \varphi_{abc} \;\leq\; \varphi_{a'b'c'} \;,$$

l'application $(a, b, c) \mapsto \varphi_{abc}$ est uniformément continue de $\mathbb{R}^3$ vers les fonctions continues à l'intérieur de K avec la norme uniforme . Lorsque (a, b, c) décrit le cercle E, φ_{abc} parcourt l'ensemble Φ' qui est donc une courbe fermée, et cela laisse penser que le point x de $b(\mu)$ tel que $\varphi_{abc}(x) = \mu(\varphi_{abc})$ décrit une courbe de Jordan qui entoure $b(\mu)$ en parcourant sa frontière. Il est intéressant de remarquer que c'est la même famille de fonctions test Φ' qui intervient à la fois pour toutes les mesures μ portées par les trois sommets.

Pour illustrer ceci, revenons à l'exemple où μ est la mesure uniforme sur les trois sommets du triangle. Nous avons vu que le point de $b(\mu) \cap AA'$ le plus proche de A est le point M tel que $f(M) = \frac{1}{3}\, f(A)$; ce devrait être aussi le point de AA' tel que $\varphi_{100}(M) = \mu(\varphi_{100}) = \frac{1}{3}$. Effectivement, puisque l'on a $\varphi = f/f(A)$ sur les deux segments géodésiques AA' et BC, on peut, sans rien changer, substituer φ_{100} à f dans la construction de M et dans la démonstration de l'optimalité de M. Mais on observera que $f < f(A)\, \varphi_{100}$ sur tout l'ensemble $K \backslash (AA' \cup BC)$; c'est donc bien φ_{100} et non f qui sera utile pour connaître le barycentre d'une autre mesure ν telle que $\nu(B) \neq \nu(C)$.

De même, le point de $b(\mu) \cap AA'$ le plus proche de A' devrait être obtenu comme la solution sur AA' de l'équation $\varphi_{011}(x) = \frac{2}{3}$.

Pour terminer, voici une dernière propriété que devraient vérifier les fonctions φ_{abc}, suggérée à la fois par leur définition et par la seconde partie de la proposition 5.1.1. Selon cette proposition, si un point x de $\partial b(\mu)$ se trouve aussi être un barycentre géodésique itéré de μ, la fonction $\varphi \in \Phi'$ telle que $\varphi(x) = \mu(\varphi)$ est affine sur un segment de géodésique contenant x. Dans le cas général où x est seulement approché par des barycentres géodésiques itérés de μ, on devrait avoir à la limite un vecteur non nul $v \in T_x H$ tel que $Hess\,\varphi(v, v) = 0$. En faisant varier la mesure μ et le point x de toutes les façons possibles, ceci donnerait, pour chaque $\varphi \in \Phi$, un champ de vecteurs non nuls défini sur l'intérieur de K et annulant $Hess\,\varphi$. Ceci rejoint la remarque faite pus haut sur le caractère presque affine de φ.

Bibliographie

[1] B. Afsari, Riemannian L^p center of mass : existence, uniqueness and convexity, Proceding of the Mathematical Society , S 0002-9939(2010)10541-5, article electronically published on August 27, 2010.

[2] T. Ando, C.K. Lie and R. Mathias, Geometric Means, Linear Algebra And Its Applications 385(2004)305-334.

[3] M. Arnaudon. Barycentres convexes et approximations des martingales continues dans les variétés, Séminaire de Probabilités XXIX, Springer Lecture Notes in Math. 1613 (1995), p. 70-85.

[4] M. Arnaudon. Espérances conditionnelles et C-martingales dans les variétés. Seminaire de probabilité (Strasbourg), t, 28, 1994, 300-311.

[5] M. Arnaudon et X.M. Li. Barycenters of measures transported by stochastic flow. The annals of mathematical statistics, 2005, t,33, No 4, 1509-1549.

[6] M. Belkhelfa, R. Deszcz, L. Vertraelen, Symmetric Properties of 3-dimensional D'Atri spaces, Kjungpook Mathematical Journal, 2006, Vol 46 Issue 3, p 367-376. 10p

[7] V. Benès, Martingales on metric spaces, Theor. Veroyatnost. i ego Premnen, 7,81-82, 1969.

[8] L. Bianchi, Lezioni sulla teoria dei gruppi continui finiti di trasformasioni, Zanichelli 1928.

[9] W.M. Boothby, An introduction to differentiable manifolds and Riemannian geometry, Second edition, Academic Press, Orlando, Florida, 1986. ISBN 0-12-116052-1 ;0-12-116053-X.

[10] S.R. Buss and J.P. Fillmore, Spherical averages and applications to spherical splines and interpolation, ACM transaction on Graphics, vol. 20(2001), pp.95-126.

[11] É. Cartan, leçons sur la géométrie des espaces de Riemann, Gauthier-Villars 1951.

[12] S. Cartier. Surfaces des espaces homogènes de dimension 3, thèse de doctorat, école doctorale MSTIC université Paris-est, février 2012, 15-20.

[13] I. Chavel, Riemannian geometry : a modern introduction, Cambriddge Tracts in mathematics 108, Cambridge University Press, Cambridge, Melbourne, New York, 1993. ISBN 0-521-43201-4.

[14] J. Cheeger and D. G. Ebin, Comparison theorems in Riemannian geometry, (North Hollande, Amsterdam, 1975).

[15] L. Cruz et D. Meijia, Strongly hyperboically convex functions, J. Math. Anal. Appl, 335(2007) 1403-1415

[16] R.W.R. Darling, Martingales in Manifolds-Definition, Examples and Behaviour under Maps, Seminaires de Probabilité XV, Supplément Géométrie différentielle stochastique, Lectures Notes in Mathematics 921, Springer 1982.

[17] M.P. Do Carmo, Riemannian geometry, Mathematics : Theory and applications. Birkhauser, 1962.

[18] S. Doss. Sur la moyenne d'un élément aléatoire dans un espace distancié. Bull. Sc. Math. 73, 1949,48-72.

[19] S. Doss. Moyennes conditionnelles et martingales dans un espace métrique, C. R. Ac. Sc Paris, Série I, t,254. 3630-3632,1962.

[20] T.E. Duncan, Stochastic Integrals in Riemannian Manifolds. J. Multivarite Anal. 6,397-413, 1976.

[21] L.P. Eisenhart, Riemannian geometry, Princeton University Press, Princeton N.J. (1925)

[22] M. Émery et P.A. Meyer, Stochastic calculus in manlfolds (Springer, Berlin, 1989)

[23] M. Émery et G. Mokobodzki. Sur le barycentre d'une probabilité dans une variété, Séminaire de Probabilités XXV, Springer Lecture Notes in Math. 1485 (1991), p. 220–223.

[24] M. Émery et W.A. Zheng. Fonctions convexes et semi-martingales dans une variété. Séminaire de probabilités (Strasbourg), t 18, 1984, 501-518.

[25] M. Fréchet, l'integrale abstraite d'une fonction abstraite d'une variable abstraite et son application à la moyenne d'un élément aléatoire de nature quelconque. Revue scientifique, 1944, 483-512.

[26] M. Fréchet, Les éléments aléatoire de nature quelconque dans un espace distancié, Annales de L'I.H.P., t 10, no 4(1948), p. 215-310.

[27] M. Fréchet, Positions typiques d'un élément aléatoire de nature quelconque, Ann. Sc. ENS, S3, t65(1948), p.211-237.

[28] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry, Universitext, Springer-Verlag, 1980.

[29] W.B. Gordon, Convex functions and harmonic maps, Proceeding of the American Mathematical Society, V 33, no 2, June 1272

[30] M. Gorine, M. Belkhelfa. Encadrement du barycentre convexe d'une loi uniforme portée par une petite sphère S^2 dans une 3-sphère. C. R. Acad. Sci. Paris, Ser. 1350, 2012, 1047-1050.

[31] M. Gorine, M. Belkhelfa. Encadrement du barycentre convexe d'une loi uniforme portée par une petite sphère S^2 dans $S^2 \times \mathbb{R}$, $H^2 \times \mathbb{R}$ et l'espace de Heisenberg H^3. C. R. Acad. Sci. Paris, Ser. 1352, 2014, 245-249.

[32] A. Gray, The volume of a small godesic ball of a Riemannian manifold, Michigan Math. J. 20, 1973, 329-344.

[33] R.E. Greene, H. Wu, Function theory on manifolds wich posses a pole, lecture notes in mathematics, 699, Springer-Verlag, Berlin-New York-Heidelberg (1979).

[34] R.E. Greene, K. Shiohama, Riemannian manifolds having a nowhere constant convex function, Notices Amer. Math. Soc. 26, 2-a-223 (1979).

[35] S. Gudmundsson, An intruduction to Riemannian Geometry, `http://www.matematik.lu.se/matematiklu/personal/sigma/index.html`

[36] E. Hebey, Intruduction à l'analyse non linéaire sur les variétés, Arts et Sciences, Diderot, Éditeur, Paris, 1997. ISBN 2-84134-031-7.

[37] S. Helgasson, Differential geometry and symmetric spaces, Acad Press, New York (1962).

[38] W. Herer. Espérance mathématique au sens de Doss d'une variable aléatoire à valeurs dans un espace métrique. C. R. Acad. Sc. Paris, t. 302, 131-134, 1986.

[39] W. Herer. Martingales à valeurs fermées bornées d'un espace métrique. C. R. Acad. Sc. Paris, t. 305, 275-278, 1987.

[40] W. Herer. Espérance mathématique d'une variable aléatoire à valeurs dans un espace métrique à courbure négative. C. R. Acad. Sc. Paris, t, 306, 681-684, 1988.

[41] W. Herer. Mathematical expectation and martingales of random subsets of a metric space Prob. And Math. Stat 11,291-304, 1991.

[42] W. Herer. Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature, Prob. And Math. Stat 13, 59-70, 1992.

[43] L. Huilier, Trigonometrie Sphérique. Analogie entre les triangles rectilignes et sphériques, Ann. Math. Appl. T1, (1810-1811), 197–201.

[44] W.S. Kendall. Convexity and the hemisphere, J. London Math. Soc. (2) t. 43, 1991, p. 223-261.

[45] W.S. Kendall, Probability, convexity and harmonic maps with small image I : uniqueness and fine existence, Proc. London Math. Soc., (2) 43 (1991), no. 3, 567-576.

[46] O. Kowalski, Spaces with volume-preserving symmetries and related classes of Riemannian manifolds. Rend.Sev. univ. Politecnico speciale, 1983, 131-157.

[47] S. Kronwith, Convex Manifolds of Nonnegative Curvature, J. Differential Geometry 14 (1979) 621-628.

[48] J.M. Lee. Riemannian manifolds. An introduction to curvature. Graduate Texts in Mathematics, 176. Springer-Verlag, 1997.

[49] A. Loi, P. Sitzia. Explicit formulas for the geodesics of homogeneous $SO(2)$-isotropic three-dimensionnal manifolds. Dipartimento di matematica- Università di Sassari, università di Cagliari- Italy.

[50] P.A. Meyer, Géométrie stochastique sans larmes, Seminaires de Probabilité XV, Lectures Notes in Mathematics 850, Springer 1981.

[51] J. Milnor, Curvature of left invariant metrics on Lie groups, Adv.in Math. 21, 1976, 293-329.

[52] B. O'Neill, Semi-Riemannian Geometry With applications to Relativity, Pure and applied Mathematics, Academic Press, 1983.

[53] J. Picard, Barycentre et martingales sur une variété, Ann, Inst, H, Poincaré Probab. Statist, 30(1994), no, 4, 647-702.

[54] J. Picard, Martingales on Riemannian Manifolds with precribed limits. J. Func. Anal.99,223-261, 1991.

[55] A. Ranjan and H. Shah, Convexity of spheres in a manofold without conjugate points, Proc, Andian. Acad. Sc.(Math. Sci) Vol.112,n4(2002)p. 595-599.

[56] T. Rockafeller, Convex Analysis, Princeton Landmarks in Mathematics and Physics.

[57] J.P. Sha, p-convex Riemannian manifolds, invent, math. 83, 437-447, Springer-Verlag, 1986.

[58] W. Tou, Geodesic Triangles and Expansion of Metrics in Normal Coordinates, Chinese Journal of Physic, Vol 16, no 4, 1978.

[59] C. Udrist, Convex functions and optimization methods on Riemannian manifolds, Kluwer Academic Publishers (1994).

[60] A. Ungar, The Hyperbolic Centroid, Comment. Math. Univ. Carolinne. 45, 2 (2004) 335-369.

[61] G. Vranceanu, Leçons de géométrie différentielle, vol. I Ed. Acad. Rp. Pop. Roumaine 1957.

[62] L. Yang, Riemannian median and its estimation, LMS J. Comput. Math. vol 13 (2010), 461-479.

Résumé. On supposera donnée une mesure de probabilité μ portée par un petit compact dans une variété différentiable M. Notre but est de trouver un encadrement du barycentre convexe de μ. La majoration est obtenue par construction de fonctions convexes presque affines et la minoration par construction d'une suite de mesures finies qui converge étroitement vers μ, puis en utilisant l'associativité des barycentres pour le passage à la dimension n.

Mots clés : boule géodésique, barycentre convexe, barycentre géodésique, convexité Riemannienne, mesure de probabilité sur les variétés Riemanniennes.

Abstract. Assume given a probability measure μ carried on a small compact in a differentiable manifold M. Our goal is to find upper and lower bounds for the convex barycenter of μ. The upper bound is obtained with the construction of almost affine convex functions. The lower bound rests on the construction of a sequence of finite measures which converges narrowly to μ, then we use the associativity of barycenters in order to go over to dimension n.

Key words : geodesic ball, convex barycenter, geodesic barycenter, Riemannian convexity, probability measure on Riemannian manifolds.