C. R. Acad. Sci. Paris, Ser. I 352 (2014) 245–249

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Géométrie différentielle/Probabilités

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

Asymptotic upper bounds for the convex barycenter of probability measure on the homogenous spaces $S^2 \times \mathbb{R}$, $H^2 \times \mathbb{R}$ and the Heisenberg space H^3

Mohamed Gorine, Mohamed Belkhelfa

L.P.Q. 3M, faculté des sciences et de la technologie, université de Mascara, Algérie

INFO ARTICLE

Historique de l'article :
Reçu le 22 octobre 2013
Accepté après révision le 16 décembre 2013
Disponible sur Internet le 14 janvier 2014

Présenté par le Comité de rédaction

RÉSUMÉ

On supposera donnée une mesure de probabilité μ portée par un petit compact dans une variété différentiable M. Notre but est de trouver des majorations du barycentre convexe de μ lorsque M est l'un des espaces homogènes $S^2 \times \mathbb{R}$, $H^2 \times \mathbb{R}$ ou l'espace de Heisenberg H^3. Les majorations sont obtenues par construction de fonctions convexes presque affines.

ABSTRACT

Assume given a probability measure μ carried on a small compact in a differentiable manifold M. Our goal is to find upper bounds for the convex barycenter of μ where M is one of the spaces $S^2 \times \mathbb{R}$, $H^2 \times \mathbb{R}$ or the Heisenberg space H^3. The upper bounds are obtained with the construction of almost affine convex functions.

1. Introduction

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.

On supposera donnée une mesure de probabilité μ à support suffisamment petit.

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 $b(\mu)$.

La définition du barycentre est évidente, ainsi que son unicité, lorsque μ est portée par un ou deux points de M ; mais, pour des mesures plus générales, en raison de l'abscence de linéarité dans l'espace M, aucune définition du barycentre de μ ne sera pleinement satisfaisante. Nous travaillons 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. Des majorations de $b(\mu)$ peuvent être obtenues en construisant des fonctions convexes ; ces approximations sont d'autant plus précises que les fonctions convexes utilisées sont plus près d'être affines.

Adresses e-mail : gorinemoh@gmail.com (M. Gorine), mohamed.belkhelfa@gmail.com, Belkhelfa@univ-mascara.dz (M. Belkhelfa).

1631-073X/$ – see front matter
http://dx.doi.org/10.1016/j.crma.2013.12.013

2. Définitions et rappels

On considére une variété M vérifiant les propriétés citées dans l'introduction.

Définition 2.1. Une fonction f définie sur un ouvert U de M est dite convexe si pour toute géodésique $\gamma : I \to U$, où I est un ouvert de $\mathbb{R}$, la fonction $f \circ \gamma$ est convexe sur I.

Proposition 2.2. *Une fonction f, définie sur un ouvert de M et de classe C^2, est convexe si et seulement si sa hessienne $\mathrm{Hess}\, f = \nabla\, \mathrm{d} f$ est positive.*

Si K est un compact de M, on note $C(K)$ l'ensemble des fonctions convexes sur un voisinage ouvert et convexe de K.
Soit μ une mesure de probabilité à support compact K dans M.

Définition 2.3. (Voir [3].) On dit que $x \in M$, est un point du barycentre convexe de μ si, pour toute fonction f appartenant à $C(K)$, $f(x)$ est définie et $f(x) \leqslant \mu(f)$.

On notera $b(\mu)$ le barycentre convexe de μ.
Pour tous x et y dans K, on note $\overrightarrow{xy} \in T_xM$ le vecteur vitesse de la géodésique passant en x au temps 0 et en y au temps 1, $\overrightarrow{xy} = \exp_x^{-1}(y)$.

Définition 2.4. (Voir [3].) On appelle barycentre exponentiel de μ tout point e de M qui vérifie $\int \overrightarrow{ey}\, \mu(\mathrm{d}y) = 0$.

On montre dans [3] que tout barycentre exponentiel de μ est un point de $b(\mu)$.

Définition 2.5. On dit que 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 $\nabla \mathrm{d}\psi_a$ strictement positive sur $M \backslash \{a\}$.

Un compact K convexe d'une variété avec connexion est à géométrie convexe s'il possède un voisinage ouvert qui est à géométrie convexe.

Définition 2.6. (Voir [5].) B est une boule géodésique fermée de centre p et de rayon r dans une variété riemannienne.
On dira 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.

On montre dans [5] que, pour une variété riemanienne, les boules géodésiques régulières sont à géométrie convexe.
Pour une mesure de probabilité portée par un compact à géométrie convexe dans une variété remplissant seulement les conditions de l'introduction, il est montré dans [5] que le barycentre exponentiel existe et est unique. En outre :

Proposition 2.7. *(Voir [1].) Pour toute distance riemanienne δ sur K, il existe une constante C telle que, pour toute probabilité μ sur K, pour tout $b \in b(\mu)$, on ait l'inégalité :*

$$\left|b(\mu)\right| \leqslant C \int_k \delta^3(b, y)\mu(\mathrm{d}y),$$

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

3. Espace homogène

Une variété riemanienne est dite homogène si son groupe d'isométries agit transitivement dessus, 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 par τ la courbure de fibration $(k - 4\tau^2 \neq 0)$.
On envisage deux cas.

- Lorsque $\tau = 0$, la variété correspond au 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 est le plan hyperbolique $(k < 0)$.

- Quant au cas $\tau \neq 0$, ce sont 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 $k \geqslant 0$ (respectivement $k \leqslant 0$), on considère sur $\mathbb{R}^3$ (respectivement, $\mathbb{D}^2(\frac{2}{\sqrt{-k}}) \times \mathbb{R}$) la métrique riemannienne g définie par :

$$\mathrm{d}s^2 = \lambda^2\big(\mathrm{d}x^2 + \mathrm{d}y^2\big) + \big(\tau\lambda(y\,\mathrm{d}x - x\,\mathrm{d}y) + \mathrm{d}z\big)^2,$$

avec :

$$\lambda = \frac{1}{1 + \frac{k}{4}(x^2 + y^2)}.$$

$\mathbb{D}^2$ désigne le disque de Poincaré.

Lorsque $\tau \neq 0$, la variété a la géométrie des sphères de Berger 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 ce cas, on définit le repère canonique (E_1, E_2, E_3) par :

$$E_1 = \frac{1}{\lambda}\left[\cos(\sigma z)\frac{\partial}{\partial x} + \sin(\sigma z)\frac{\partial}{\partial y}\right] + \tau\big(x\sin(\sigma z) - y\cos(\sigma z)\big)\frac{\partial}{\partial z},$$
$$E_2 = \frac{1}{\lambda}\left[-\sin(\sigma z)\frac{\partial}{\partial x} + \cos(\sigma z)\frac{\partial}{\partial y}\right] + \tau\big(x\cos(\sigma z) + y\sin(\sigma z)\big)\frac{\partial}{\partial z}$$

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 que [2] 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.

4. Résultats

Proposition 4.1. *(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 $T_O M$, 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 $r = r(x) = \|\overrightarrow{Ox}\|$ est la distance riemanienne de O à x.

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 á 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 *Hess* f est positive sur $B\backslash\{O\}$. En effet, cela assure que $f o\gamma$ est convexe pour toute géodésique γ dans $B\backslash\{O\}$; d'autre part, pour les géodésiques γ passant par O, $f o\gamma$ est convexe parceque les fonctions d'une variable réelle $\xi \mapsto a\xi + C|\xi|^3$ sont convexes.

Pour calculer *Hess* f sur $B\backslash\{O\}$, fixons une base orthonormée (e_1, e_2, e_3) 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, x^3) centré en O et associé à cette base, on a donc $\lambda(\overrightarrow{Ox}) = x^1$.

Dans ces coordonnées normales, le développement de Taylor à l'ordre 2 de la métrique g est donné au voisinage de O par :

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

(voir par exemple [6], pp. 72–76), où le terme d'erreur $o(r^2)$ désigne une matrice $\epsilon(x)$, non nécessairement radiale, mais telle que $\epsilon(x)/r^2(x) \to 0$ quand $x \to O$. Il en résulte que les symboles de Christoffel Γ_{ij}^l vérifient $\Gamma_{ij}^l(O) = 0$ et $\frac{\partial \Gamma_{ij}^l}{\partial x^n}(O) = \frac{1}{3}(R_{lijn}(O) + R_{ljin}(O))$.

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, x^3) est donnée par :

$$\frac{1}{3C}\,Hess\,f = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} - hx^2 & \frac{x^1x^3}{r} \\ \frac{x^1x^2}{r} - hx^2 & r + \frac{(x^2)^2}{r} + 2hx^1 & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} \end{pmatrix} + o(r^2).$$

Nous allons vérifier que $Hess\,f(x)$ 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.

Ce qui est assuré sous l'hypothèse $h < \sqrt{2} - 1$. □

Proposition 4.2. *(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 $T_O M$; 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 $r = r(x) = \|\overrightarrow{Ox}\|$ est la distance riemanienne de O à x.

La preuve est analogue à celle de la proposition 4.1, en remplaçant R par $-R$.

Proposition 4.3. *(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 $T_O M$ 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 $T_O M$, la fonction $f(x) = \lambda(\overrightarrow{Ox}) + Cr^3$ est convexe sur un voisinage de O.*

τ est la courbure de fibration et $r = r(x) = \|\overrightarrow{Ox}\|$ est la distance riemanienne de O à x.

Preuve. Dans ce cas, les courbures sectionnelles sont :

$$R_{1212} = -3\tau^2 \quad \text{et} \quad R_{1313} = R_{2323} = \tau^2.$$

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 :

$$\frac{1}{3C}\,Hess\,f = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} + x^2\frac{\tau^2}{3C} & \frac{x^1x^3}{r} - x^3\frac{\tau^2}{9C} \\ \frac{x^1x^2}{r} + x^2\frac{\tau^2}{9C} & r + \frac{(x^2)^2}{r} - 2x^1\frac{\tau^2}{3C} & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} - x^3\frac{\tau^2}{9C} & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} + 2x^1\frac{\tau^2}{9C} \end{pmatrix} + o(r^2).$$

Posons $\frac{\tau^2}{9C} = h$:

$$\frac{1}{3C}\,Hess\,f = \begin{pmatrix} r + \frac{(x^1)^2}{r} & \frac{x^1x^2}{r} + 3x^2h & \frac{x^1x^3}{r} - x^3h \\ \frac{x^1x^2}{r} + 3x^2h & r + \frac{(x^2)^2}{r} - 6x^1h & \frac{x^2x^3}{r} \\ \frac{x^1x^3}{r} - x^3h & \frac{x^2x^3}{r} & r + \frac{(x^3)^2}{r} + 2x^1h \end{pmatrix} + o(r^2).$$

Pour que $Hess\,f$ soit définie positive sur $B\backslash\{O\}$, il suffit que $h < \frac{\sqrt{65}-7}{8}$. □

Corollaire 4.4. *Soient (M, g), O et R comme dans la proposition 4.1.*

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.

La preuve est analogue au corollaire 2.2 dans [4].

On peut énoncer, de même :

Corollaire 4.5. *Soient (M, g), O et R et τ comme dans la propositions 4.2 et 4.3 respectivement.*

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 4.6. *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*

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

- *si M est le produit* $H^2 \times \mathbb{R}$, *alors*

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

- *si M est l'espace de Heisenberg , alors*

$$\limsup_{D\to 0} \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leqslant \frac{7+\sqrt{65}}{36} \tau^2,$$

$E(D)$ *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 4.4, toutes les fonctions $f_{C,\lambda}(x) = \lambda(\overrightarrow{Ox}) + C r^3$, 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 :

$$r(x) = \lambda(\overrightarrow{Ox}) \leqslant \lambda(\overrightarrow{Ox}) + C\, r(x)^3 = f_{C,\lambda}(x) \leqslant \mu(f_{C,\lambda}).$$

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 $r \leqslant D/2$ sur le support de μ. Il reste donc $r(x) \leqslant C\, D^3/8$. Ceci ayant lieu pour tout x dans $b(\mu)$, on a $|b(\mu)| \leqslant C\, D^3/4$. On a ainsi montré que :

$$\forall C > \frac{1+\sqrt{2}}{9} R \quad \exists \varepsilon > 0 \quad \forall D < \varepsilon \qquad \sup_{\mu\in E(D)} \frac{|b(\mu)|}{D^3} \leqslant \frac{C}{4}. \qquad \square$$

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 $\frac{7+\sqrt{65}}{9} \tau^2$.

Références

[1] M. Arnaudon, Barycentres convexes et approximations des martingales continues dans les variétés, in : Séminaire de probabilités XXIX, in : Springer Lecture Notes in Math., vol. 1613, 1995, pp. 70–85.

[2] S. Cartier, Surfaces des espaces homogènes de dimension 3, thèse de doctorat, école doctorale MSTIC, université Paris-Est, février 2012, pp. 15–20.

[3] M. Émery, G. Mokobodzki, Sur le barycentre d'une probabilité dans une variété, in : Séminaire de probabilités XXV, in : Springer Lecture Notes in Math., vol. 1485, 1991, pp. 220–223.

[4] 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. I 350 (23–24) (2012) 1047–1050.

[5] W.S. Kendall, Convexity and the hemisphere, J. Lond. Math. Soc. (2) 43 (1991) 223–261.

[6] J.M. Lee, Riemannian Manifolds. An Introduction to Curvature, Graduate Texts in Mathematics, vol. 176, Springer-Verlag, 1997.