ANNALES DE L'I. H. P., SECTION B

JEAN PICARD

Barycentres et martingales sur une variété

Annales de l'I. H. P., section B, tome 30, nº 4 (1994), p. 647-702
<http://www.numdam.org/item?id=AIHPB_1994__30_4_647_0>

Ann. Inst. Henri Poincaré,

Vol. 30, n° 4, 1994, p. 647-702. *Probabilités et Statistiques*

Barycentres et martingales sur une variété

by

Jean PICARD

Laboratoire de Mathématiques Appliquées
C.N.R.S. (URA 1501)
Université Blaise Pascal (Clermont-Ferrand II)
63177 Aubière Cedex, France

RÉSUMÉ. – Étant donnée une variété V, nous considérons une famille d'applications, que nous appelons barycentres, qui à toute probabilité sur V font correspondre un point de V. Pour chaque barycentre nous définissons la classe des martingales sur V comme l'ensemble des semimartingales avec sauts à valeurs dans V et possédant une propriété liée au barycentre; cette notion généralise la notion de martingale continue définie par la géométrie différentielle d'ordre 2. Nous en donnons plusieurs caractérisations équivalentes, et montrons que sous certaines hypothèses, il existe une et une seule martingale de valeur finale donnée; ce résultat peut être appliqué à l'étude probabiliste des applications harmoniques à valeurs dans V.

Mots clés : calcul stochastique sur les variétés; barycentres; variétés à géométrie convexe; construction de martingales.

ABSTRACT. – On a manifold V, we consider a family of V-valued maps, called barycentres, defined on the set of probability measures on V. For each barycentre, we define the class of martingales on V as the set of V-valued semimartingales with jumps which satisfy a condition involving the barycentre; this notion extends the notion of continuous martingales defined by means of the second order differential geometry. We give several equivalent characterizations, and prove that under some assumptions, there exists one and only one martingale with prescribed final value; this result can be applied to the probabilistic study of V-valued harmonic maps.

Classification A.M.S.: 58G32, 60G44.

0. INTRODUCTION

Le but de ce travail est d'introduire une notion de barycentre pour une probabilité portée par une variété, et d'étudier la notion de martingale associée. Dans le cas d'un espace euclidien, les martingales permettent de donner une interprétation probabiliste à certaines équations aux dérivées partielles, ou plus généralement aux équations associées à des générateurs de semi-groupes markoviens; la notion de martingale à valeurs dans une variété va permettre de représenter les solutions d'équations soumises à une contrainte non linéaire. Les équations associées à des opérateurs différentiels vont pouvoir se représenter à l'aide de processus continus, mais pour des opérateurs non locaux, nous aurons besoin de processus discontinus. Dans ce cas, les outils de la géométrie différentielle locale tels que les connexions ne suffisent plus, alors que la notion de barycentre va nous fournir un outil global permettant de traiter le problème; lorsque nous nous restreindrons à des processus continus, nous retrouverons la notion classique de martingale continue à valeurs dans une variété.

Dans une première étape, commençons par rappeler la caractérisation probabiliste d'une fonction harmonique réelle. Soit E une variété régulière, soit $\mathcal{L}$ un opérateur différentiel d'ordre 2 sans terme d'ordre 0, agissant sur les fonctions de classe C^2 de E dans $\mathbb{R}$, et soit h une fonction de E dans $\mathbb{R}$. On dit que h est harmonique sur E si $\mathcal{L}h = 0$. Si Z_t est le processus de diffusion associé à $\mathcal{L}$, il est bien connu qu'une fonction régulière h est harmonique si et seulement si $h(Z_t)$ est une martingale locale réelle continue pour toute condition initiale $Z_0 = z$. De plus, cette caractérisation peut se généraliser à des opérateurs non locaux, plus précisément aux générateurs infinitésimaux de diffusions avec sauts Z_t; dans ce cas la martingale locale $h(Z_t)$ pourra présenter des sauts. Une conséquence de ce résultat est que la solution de certaines équations peut s'obtenir en étudiant l'ensemble des martingales locales réelles convergeant vers une variable fixée. Ainsi, si ∂E est une partie de E, si u est une fonction réelle définie sur ∂E, et si le temps d'atteinte τ de ∂E par Z_t est presque sûrement fini, une fonction régulière h est solution du problème de Dirichlet

$$\mathcal{L}h(z) = 0, \quad z \in E\backslash\partial E; \quad h(z) = u(z), \quad z \in \partial E$$

si et seulement si $h(Z_{t\wedge r})$ est une martingale locale prenant la valeur $u(Z_r)$ pour $t \geqq \tau$. De même, si u est une fonction réelle définie sur E, on vérifie en considérant la diffusion (t, Z_t) qu'une fonction régulière h

est solution de l'équation de la chaleur

$$\left(\frac{\partial}{\partial t} + \mathcal{L}\right) h(t, \cdot) = 0, \quad 0 \leqq t \leqq 1; \quad h(1, \cdot) = u$$

si et seulement si $h(t, Z_t)$, $0 \leqq t \leqq 1$, est une martingale locale prenant la valeur $u(Z_1)$ en $t = 1$.

Dans une deuxième étape, considérons un espace euclidien $\mathbb{R}^d$; en le faisant agir composante par composante, l'opérateur $\mathcal{L}$ s'étend aux fonctions de E dans $\mathbb{R}^d$, et la caractérisation probabiliste des fonctions harmoniques se généralise immédiatement en considérant des martingales locales euclidiennes. La troisième étape consiste à remplacer l'espace euclidien $\mathbb{R}^d$ par une variété V. Dans ce cadre, au moins deux questions se posent. D'une part, il faut savoir comment faire agir $\mathcal{L}$ sur les fonctions h de E dans V afin que la notion d'application harmonique garde un sens; d'autre part, il faut savoir s'il est possible de définir une notion de martingale sur V telle que h soit harmonique si et seulement si $h(Z_t)$ est une martingale. Une fois ces deux définitions données, la résolution probabiliste d'équations associées à $\mathcal{L}$ s'effectue en étudiant l'existence et l'unicité d'une martingale sur V prenant une valeur finale fixée.

Nous commençons par rappeler comment, dans le cadre classique des opérateurs différentiels d'ordre 2, ces deux questions sont résolues par la géométrie différentielle d'ordre 2 de [22] et [27]. Supposons d'abord que $\mathcal{L}$ est d'ordre 1, c'est-à-dire un champ de vecteurs sur E; si h est une fonction régulière de E dans V, l'application

$$L_z : g \mapsto \mathcal{L}(g \circ h)(z)$$

définie sur les fonctions régulières g de V dans $\mathbb{R}$ ne dépend que du comportement à l'ordre 1 de g en $h(z)$, donc L_z est un vecteur tangent à V en $h(z)$ que nous noterons $\mathcal{L}_V\, h(z)$. Supposons ensuite que $\mathcal{L}$ est un opérateur différentiel d'ordre 2 sans terme d'ordre 0; dans le vocabulaire de la géométrie différentielle d'ordre 2, $\mathcal{L}$ est un champ de vecteurs tangents d'ordre 2 sur E; dans ce cas, l'application L_z dépend du comportement de g à l'ordre 2, donc est un vecteur tangent d'ordre 2 à V en $h(z)$; la donnée d'une connexion sur V permet de lui associer un vecteur d'ordre 1 que nous noterons à nouveau $\mathcal{L}_V\, h(z)$; nous dirons alors que h est harmonique si $\mathcal{L}_V\, h = 0$. En ce qui concerne la caractérisation probabiliste, la donnée d'une connexion permet également de définir une notion de martingale continue à valeurs dans V; cette notion a été introduite dans [7], puis a été formalisée et étudiée dans [3] et [22]; diverses caractérisations équivalentes sont maintenant connues : voir le livre [11] auquel nous ferons généralement référence pour ces questions; en particulier, nous savons qu'une application régulière h est harmonique si et seulement si $h(Z_t)$ est une martingale pour toute condition initiale Z_0. Le cas qui a été essentiellement étudié dans la littérature non probabiliste est le cas du laplacien (une bibliographie pourra être trouvée dans [8] et [9]); plus précisément, si E et V sont des variétés riemanniennes et si $\mathcal{L}$ est le laplacien Δ sur E, $\mathcal{L}_V$ est appelé opérateur de tension et nous obtenons la notion classique d'application harmonique entre variétés; h est alors harmonique si et seulement si elle transforme le mouvement browien sur E en une martingale sur V; de même, si $\mathcal{L} = \Delta + \partial/\partial t$, nous obtenons l'équation de la chaleur non linéaire qui est un outil pour rechercher les applications harmoniques.

Si $\mathcal{L}$ est le générateur d'une diffusion avec sauts, L_z dépend de tout le comportement de g donc la notion de connexion n'est plus suffisante pour lui associer un vecteur tangent et nous avons besoin d'un objet plus général. Nous allons utiliser les connecteurs de [26]. Un connecteur consiste en la donnée d'une famille d'applications γ_x, $x \in V$, de V dans l'espace tangent $T_x(V)$; en gros, le vecteur $\gamma_x(y)$ exprime comment est vu le point y depuis le point x. L'application L_z peut être étendue aux fonctions à valeurs dans $T_{h(z)}(V)$, et nous pouvons définir $\mathcal{L}_V\, h(z)$ comme étant $L_z(\gamma_{h(z)})$; comme précédemment, l'application h est dite harmonique si $\mathcal{L}_V\, h = 0$. De plus, le connecteur permet de définir l'intégrale de Ito d'une forme d'ordre 1 par rapport à une semimartingale avec sauts sur V (la construction de [26] est en fait un cas particulier du calcul différentiel de [2]); on peut en déduire une notion de martingale avec sauts sur V (définition 4.1 de [26]), et nous verrons que la caractérisation des applications harmoniques est encore valable. Ces notions sont compatibles avec la géométrie différentielle d'ordre 2: à chaque connecteur est associée une connexion qui fournit le même $\mathcal{L}_V$ lorsque $\mathcal{L}$ est un opérateur différentiel, et les deux notions de martingales coïncident pour les processus continus.

Si maintenant nous voulons résoudre

$$\left(\frac{\partial}{\partial t} + \mathcal{L}_V\right) h(t, \cdot) = 0, \quad 0 \leqq t \leqq 1; \quad h(1, \cdot) = u \tag{0.1}$$

ou

$$\mathcal{L}_V\, h(z) = 0, \quad z \in E \backslash \partial E; \quad h(z) = u(z), \quad z \in \partial E, \tag{0.2}$$

nous sommes amenés, comme dans le cas scalaire, à étudier l'ensemble des martingales sur V convergeant vers une variable fixée. Le but de ce travail est de proposer une construction approchée de ces martingales par discrétisation du temps. Pour cela, nous allons introduire une notion de barycentre; plus précisément, un barycentre est une application qui à une probabilité sur V fait correspondre un point, et qui satisfait une propriété de compatibilité avec le connecteur. Les barycentres vont nous permettre de définir toute une famille de problèmes discrétisés approchés pour la recherche d'applications harmoniques. De plus, les barycentres fournissent une nouvelle notion de martingale qui coïncide avec la notion obtenue à partir du connecteur dans le cas quasi-continu à gauche. Les résultats que nous allons obtenir ont un intérêt même dans le cas continu : ils fournissent aussi bien des nouvelles caractérisations de la notion de martingale que des algorithmes de calcul. On sait que sans hypothèse sur V, dans le cas du laplacien, l'équation de la chaleur (0.1) n'admet pas toujours une solution classique; plus précisément la solution peut exploser en temps fini dans l'espace de Sobolev $W^{1,\infty}$: voir un exemple dans [1]; notre méthode utilisant également des estimations de type lipschitzien, nous retrouverons la même difficulté dans la construction des martingales, et nous serons amenés à faire une hypothèse de convexité sur la géométrie de V (comme en [19] pour le cas continu).

Détaillons la construction du barycentre dans le cas où V est une sous-variété de $\mathbb{R}^n$, c'est-à-dire représente un ensemble de contraintes non linéaires pour un système évoluant dans $\mathbb{R}^n$. Nous supposons donnée une projection π sur V définie sur l'enveloppe convexe de V dans $\mathbb{R}^n$; π représente la façon dont le système est contraint à rester sur V. Si h est à valeurs dans V, nous pouvons également la considérer comme étant à valeurs dans $\mathbb{R}^n$ et noter $\mathcal{L}\, h$ l'action de $\mathcal{L}$ sur chacune des composantes de h; nous définissons alors

$$\mathcal{L}_V\, h(z) = \pi'(h(z))\, \mathcal{L}\, h(z)$$

où π' désigne la dérivée de π, c'est-à-dire que $\pi'(y)$ est une projection linéaire de $\mathbb{R}^n$ sur l'espace tangent $T_y(V)$. Nous voulons discrétiser l'équation de la chaleur (0.1) associée à $\mathcal{L}_V$. Soit donc $\sigma = (t_i,\, 0 \leqq i \leqq k)$ une subdivision de l'intervalle [0, 1]. Nous supposons qu'entre les instants de la subdivision, le système discrétisé diffuse dans $\mathbb{R}^n$ selon $\mathcal{L}$, alors qu'aux instants t_i il est ramené sur V au moyen de π; plus précisément, si $P(t)$ est le semi-groupe associé à $\mathcal{L}$, nous considérons la fonction $h_\sigma(t, z)$ de $\{t_i,\, 0 \leqq i \leqq k\} \times E$ dans V définie par

$$h_\sigma(t_i, .) = \pi\,[P(t_{i+1} - t_i)\, h_\sigma(t_{i+1}, .)], \quad h_\sigma(1, .) = u. \tag{0.3}$$

Si nous faisons tendre le pas de σ vers 0, nous retrouvons, au moins formellement, l'équation en temps continu (0.1) (une telle procédure n'est évidemment valable que si l'équation limite a une solution, ce qui imposera des hypothèses sur la variété); cette équation apparaît donc comme une équation de réaction-diffusion avec réaction instantanée donnée par π. Quant à l'interprétation probabiliste de cet algorithme, le passage de l'instant t_{i+1} à l'instant t_i permet d'écrire

$$h_\sigma(t_i, Z_{t_i}) = \pi(\mathbb{E}[h_\sigma(t_{i+1}, Z_{t_{i+1}}) \mid Z_{t_i}]).$$

Nous sommes donc amenés à définir le barycentre d'une variable aléatoire X à valeurs dans V comme étant

$$\mathbb{B}[X] = \pi(\mathbb{E}[X]).$$

Nous pouvons alors dire que $h_\sigma(t_i, Z_{t_i})$ est le barycentre conditionnel de $h_\sigma(t_{i+1}, Z_{t_{i+1}})$ sachant Z_{t_i}, ou que le processus à temps discret $h_\sigma(t_i, Z_{t_i})$ est une martingale sur V. Cette notion de barycentre ou de martingale à temps discret sur V dépend évidemment de la projection π. Nous donnerons également une définition de la notion de martingale à temps continu qui permettra de dire que, après passage à la limite, $h(t, Z_t)$ est une martingale sur V avec valeur terminale $u(Z_1)$. L'algorithme de discrétisation apparaît ainsi comme l'approximation d'une martingale à temps continu par une martingale à temps discret. De plus, cette procédure de construction par discrétisation peut être employée en dehors du cadre markovien; étant données une filtration $\mathcal{F}_t$, $0 \leqq t \leqq 1$, une variable $L\,\mathcal{F}_1$-mesurable à valeurs dans V et une subdivision σ de [0, 1], nous pouvons considérer la martingale à temps discret

$$X_1^\sigma = L, \quad X_{t_i}^\sigma = \pi(\mathbb{E}[X_{t_{i+1}}^\sigma \mid \mathcal{F}_{t_i}])$$

et étudier la limite de cette famille de processus lorsque le pas de la subdivision tend vers 0.

Dans l'exemple précédent, V est une sous-variété de $\mathbb{R}^n$, mais nous pouvons aussi considérer une variété abstraite. Supposons que V est riemannienne et que deux points de V sont joints par une géodésique et une seule, la vitesse à l'instant initial de la géodésique allant de x à y étant égale à $\exp_x^{-1}(y)$. Dans ce cas, nous pouvons définir l'opérateur de tension $\mathcal{L}_V$ par

$$\mathcal{L}_V h(z) = \mathcal{L}(\exp_{h(z)}^{-1} \circ h)(z). \tag{0.4}$$

Si E est une variété riemannienne et $\mathcal{L}$ le laplacien, nous pouvons vérifier que l'équation $\mathcal{L}_V h = 0$ caractérise les applications harmoniques classiques. Le barycentre associé à cet opérateur $\mathcal{L}_V$ est défini, pour toute variable X, comme le point x tel que

$$\mathbb{E}[\exp_x^{-1} X] = 0 \tag{0.5}$$

(il faut une hypothèse supplémentaire pour avoir l'existence et l'unicité d'un tel point). Comme précédemment, nous en déduisons une notion de martingale à temps discret définie par

$$\mathbb{E}\,[\exp_{X_{t_i}^\sigma}^{-1} X_{t_{i+1}}^\sigma \,|\, \mathcal{F}_{t_i}] = 0$$

et nous définirons une notion de martingale à temps continu permettant de passer à la limite. En analyse classique, les applications harmoniques sont les points extrémaux d'une fonction d'énergie; une telle caractérisation variationnelle semble ne pas exister pour les martingales, mais plutôt pour les barycentres (*voir* [10]); en effet, le barycentre d'une variable X est un point extrémal pour la distance quadratique moyenne à X et peut donc être défini par

$$\mathbb{B}\,[X] = \arg\min_x \; \mathbb{E}\,[\delta^2\,(x,\, X)],$$

où δ est la distance riemannienne. Comme cela a été remarqué depuis longtemps (*voir* [15]), ce type de barycentre peut être défini pour une distance non riemannienne; en utilisant la définition (0.5), on peut également étendre la notion de barycentre à des variétés non riemanniennes mais munies d'une connexion. Pour chacune de ces deux généralisations, nous définissons un opérateur $\mathcal{L}_V$ et une notion de martingale.

Les barycentres permettent donc de résoudre des problèmes associés à des diffusions avec sauts. De telles diffusions interviennent par exemple dans l'étude des processus d'excursion. Donnons-en un exemple plus simple; supposons que l'espace d'état E est fini, que α est une application symétrique positive sur $E \times E$, que $(V,\, \delta)$ est une variété riemannienne et que nous cherchons une application h de E dans V minimisant la fonctionnelle d'énergie

$$\mathcal{E}\,(h) = \sum_{(y,\,z)\in E^2} \alpha\,(y,\, z)\,\delta^2\,(h\,(y),\, h\,(z))$$

avec une contrainte fixant les valeurs de h sur une partie ∂E de E; cela correspond à la construction sur V d'une grille indexée par E minimisant la distance quadratique moyenne entre points. Si $\mathcal{L}$ est le générateur du processus de Markov Z_t sur E pour lequel l'intensité du saut de y à z est $\alpha\,(y,\, z)$ et si $\mathcal{L}_V$ est défini par (0.4), alors les applications $\mathcal{L}_V$-harmoniques sont les points extrémaux de $\mathcal{E}$, et sont telles que $h\,(Z_t)$ est une martingale jusqu'au temps d'atteinte de ∂E.

Terminons cette introduction en donnant le plan de l'article et les principales hypothèses. Au § 1, nous allons donner une définition du barycentre unifiant et généralisant les exemples donnés ci-dessus (variété plongée dans un espace euclidien et variété munie d'une connexion ou d'une distance); nous décrirons les liens avec d'autres notions de barycentres décrits dans [6] et [12]. Au § 2, nous montrerons quelques propriétés. Au § 3, nous introduirons la notion de convexité relative à un barycentre. Au § 4, nous associerons à chaque barycentre une notion de martingale; nous en donnerons plusieurs caractérisations et vérifierons en particulier la compatibilité avec d'autres définitions, en particulier celle classique du cas continu; ces résultats seront démontrés au § 5. Au § 6, nous décrirons les liens entre convexité et martingales, et en donnerons les conséquences sur la discrétisation en temps des martingales; cela fournira par exemple un schéma de discrétisation pour l'équation de la chaleur avec contrainte non linéaire (0.1). Finalement, aux §§ 7 et 8, nous aborderons le problème de l'existence et de l'unicité d'une martingale avec valeur terminale donnée; la condition suffisante que nous donnerons portera sur la variété, le barycentre, mais aussi sur l'espace de probabilité filtré. En particulier, nous obtiendrons ainsi sous certaines hypothèses sur la variété l'existence probabiliste de solutions à l'équation de la chaleur (0.1) et au problème de Dirichlet (0.2), ce qui généralisera les résultats du cas continu de [19] (*voir* aussi [25]). Les principaux objets qui seront définis sont les notions de connecteur (définition 1.1), hessien (formule (1.1)), barycentre (définition 1.2), fonction convexe (définition 3.1), variété à géométrie convexe (définition 3.3), γ-martingale (définition 4.2), $\mathbb{B}$-martingale (définition 4.7) et processus dominant (définition 5.1). le mot «régulier» signifiera C^∞, ou du moins différentiable jusqu'à un certain ordre que nous ne chercherons pas à préciser.

Hypothèse de base sur la variété. – Dans tout l'article, nous considérons une variété V régulière compacte avec bord.

La compacité sera utile pour l'uniformité des estimations. Les espaces tangent et cotangent seront notés $T(V)$ et $T^*(V)$, les fibres en un point x étant notées $T_x(V)$ et $T_x^*(V)$. Les métriques riemanniennes que nous considérerons seront supposées régulières; elles fournissent des distances sur V qui sont équivalentes entre elles (car V est compacte), et qui sont aussi équivalentes à la distance induite par un plongement régulier quelconque de V dans un espace euclidien $\mathbb{R}^n$. Dans la plupart de nos estimations, la distance entre deux points de V pourra donc être calculée pour l'une quelconque de ces distances, sans qu'il soit besoin de la préciser. L'aspect stochastique sera modélisé par un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ satisfaisant les conditions habituelles; l'intervalle de temps sera en général $[0, 1]$.

1. BARYCENTRES

Nous désirons définir une notion de barycentre pour une probabilité sur V. Pour cela, nous introduisons tout d'abord la notion de connecteur sur V de [26]; il s'agit d'une fonction qui à deux points x et y de la variété fait correspondre un vecteur qui remplace la différence $y - x$ du cas euclidien.

DÉFINITION 1.1. – *Un connecteur est une application régulière γ de $V \times V$ dans le fibré tangent $T(V)$ telle que*

(*a*) *pour tout x de V, l'application $\gamma_x : y \mapsto \gamma(x, y)$ envoie V dans $T_x(V)$ et s'annule en $y = x$;*

(*b*) *la dérivée de γ_x en $y = x$ est égale à l'identité de $T_x(V)$.*

Exemple 1. – Si V est une sous-variété de $\mathbb{R}^n$, l'espace tangent $T_x(V)$ peut être identifié à un sous-espace vectoriel de $\mathbb{R}^n$, et si pour x dans V, $A(x)$ est une projection linéaire de $\mathbb{R}^n$ sur $T_x(V)$, l'application $\gamma(x, y) = A(x)(y - x)$ est un connecteur.

Exemple 2. – Supposons que V est munie d'une distance δ dont le carré est régulier sur $V \times V$; nous pouvons considérer

$$g(x, y) = \frac{\partial}{\partial y} \frac{\partial}{\partial x} \delta^2(x, y)$$

comme une application linéaire de $T_y(V)$ dans $T_x^*(V)$, et nous supposons que pour $y = x$, cette application est non dégénérée (cela signifie que la distance est équivalente aux distances induites par les métriques riemanniennes). Alors

$$\gamma(x, y) = (g(x, x))^{-1} \frac{\partial}{\partial x} \delta^2(x, y)$$

est un connecteur que nous appellerons connecteur métrique. Si δ est la distance euclidienne calculée dans $\mathbb{R}^n$ pour un plongement de V dans $\mathbb{R}^n$, nous retrouvons l'exemple précédent avec $A(x)$ égal à la projection orthogonale sur $T_x(V)$.

Exemple 3. – Si V est munie d'une connexion et si deux points de V sont reliés par une géodésique et une seule, l'application $\gamma(x, y) = \exp_x^{-1}(y)$ est un connecteur que nous appellerons connecteur géodésique. S'il s'agit de la connexion de Levi-Civita d'une métrique riemannienne, nous retrouvons l'exemple précédent du connecteur métrique. Un autre cas particulier est le cas des parties convexes de groupes de Lie.

Soit x un point de V qui n'est pas sur le bord. Si γ est un connecteur, l'application γ_x est un difféomorphisme d'un voisinage de x dans V sur un voisinage de 0 dans $T_x(V)$; notons γ_x^{-1} l'inverse. Si f est une application réelle régulière sur V, nous pouvons définir Hess $f(x)$ comme étant le hessien classique de $f \circ \gamma_x^{-1}$ en 0; c'est une forme bilinéaire symétrique sur $T_x(V)$ satisfaisant

$$\begin{aligned} f(y) = f(x) + f'(x)\gamma(x, y) \\ + \frac{1}{2}\operatorname{Hess} f(x)\langle \gamma(x, y), \gamma(x, y)\rangle + O(\delta^3(x, y)) \end{aligned} \tag{1.1}$$

si $y \to x$, pour toute distance riemannienne δ; cette définition s'étend aux points situés sur le bord de V. L'opérateur Hess possède les propriétés d'un hessien au sens de [11] (définition 4.1), c'est-à-dire que nous pouvons associer à tout connecteur une connexion sans torsion. Si on applique cette opération à un connecteur géodésique, on retrouve la partie sans torsion de la connexion de départ. Deux connecteurs γ_1 et γ_2 peuvent fournir la même connexion; il faut et il suffit pour cela que

$$\gamma_2(x, y) - \gamma_1(x, y) = O(\delta^3(x, y))$$

si $y \to x$.

Nous allons maintenant définir la notion de barycentre qui à toute variable aléatoire X à valeurs dans V fait correspondre un point $\mathbb{B}[X]$ de V. Remarquons que si δ est une distance riemannienne (ou une distance équivalente), on peut définir la variance de X par

$$\operatorname{var} X = \inf_{x \in V} \mathbb{E}\,\delta^2(x, X).$$

La variance dépend de δ mais toutes les variances sont équivalentes, donc une propriété du type $o(\operatorname{var} X)$ est intrinsèque. Une autre définition conduisant à une expression équivalente est

$$\operatorname{var} X = \mathbb{E}\,\delta^2(X, X')$$

où X' est indépendant et de même loi que X.

DÉFINITION 1.2. – *Soit $\mathbb{B}$ une application définie sur l'espace $\mathcal{P}$ des probabilités sur V et à valeurs dans V; si X est une variable aléatoire à valeurs dans V, nous noterons $\mathbb{B}[X]$ l'image par $\mathbb{B}$ de la loi de X. Nous dirons que $\mathbb{B}$ est un barycentre si*

(*a*) *$\mathbb{B}$ vérifie une condition de Lipschitz*

$$\delta(\mathbb{B}[X], \mathbb{B}[Y]) \leqq C\,\mathbb{E}\,\delta(X, Y);$$

(*b*) *$\mathbb{B}[X] = X$ si X est déterministe;*

(*c*) *il existe un connecteur* γ *tel que*

$$\mathbb{E}\,\gamma(\mathbb{B}[X],\,X) = o(\operatorname{var} X) \qquad si \quad \operatorname{var} X \to 0. \tag{1.2}$$

La condition (a) implique que $\mathbb{B}$ est continu sur $\mathcal{P}$ pour la topologie de la convergence étroite; en effet, si une suite de probabilités converge étroitement, on peut trouver sur un espace de probabilité une suite de variables ayant pour lois ces probabilités et convergeant presque sûrement. Dans la condition (c), le membre de gauche est un vecteur tangent à V en $\mathbb{B}[X]$; il s'agit donc d'une estimation sur la norme de ce vecteur calculée pour une métrique riemannienne quelconque sur V. Pour $0 \leqq t \leqq 1$, si X_t est une variable prenant les valeurs x et y avec probabilités respectives $1-t$ et t, alors $\mathbb{B}[X_t]$ est une courbe continue allant de x à y, et (c) implique que sa vitesse à l'instant initial est égale à $\gamma(x,\,y)$. Cela implique que γ est déterminé de façon unique par $\mathbb{B}$; ainsi à tout barycentre on associe un connecteur, et à tout connecteur on associe une connexion sans torsion. Un cas important est le cas du γ-barycentre pour lequel l'expression (1.2) est toujours nulle.

DÉFINITION 1.3. – *Soit* γ *un connecteur; le* γ*-barycentre d'une variable* X *à valeurs dans* V *est l'ensemble des* $x \in V$ *tels que* $\mathbb{E}\,\gamma(x,\,X) = 0$. *Si cet ensemble est un singleton qui vérifie la condition (a) de la définition* 1.2, *nous dirons que le* γ*-barycentre est bien défini.*

Par exemple, si V est un compact de $\mathbb{R}^n$ avec $\gamma(x,\,y) = y - x$, le γ-barycentre est bien défini si et seulement si V est convexe. Dans le cas général, même si $X = x$, son γ-barycentre n'est pas nécessairement réduit à $\{x\}$; il faut pour cela que γ ne s'annule que sur la diagonale. Si c'est le cas, on peut montrer que le γ-barycentre est bien défini pour les variables à support suffisamment petit (on utilise la condition (b) de la définition 1.1).

Remarque. – La variété V se plonge canoniquement dans $\mathcal{P}$ en faisant correspondre à tout point la masse de Dirac en ce point. Ainsi nous demandons à $\mathbb{B}$ d'être une projection continue de $\mathcal{P}$ sur V. Une telle hypothèse impose des restrictions topologiques sur V; ainsi V doit être contractable sur un point car $\mathcal{P}$ l'est; de plus, le bord de V doit être non vide. Pour étudier d'autres variétés, comme par exemple la sphère, il faudrait une définition plus générale; on peut considérer des barycentres qui ne sont définis que sur un voisinage de V dans $\mathcal{P}$, par exemple sur les variables de variance suffisamment petite; notons toutefois que cette restriction ne dispense pas de la donnée d'un connecteur sur toute la variété, sauf si on se limite à des variables dont le support est suffisamment petit.

Exemple 1. – Soit V une sous-variété compacte à bord de $\mathbb{R}^n$; soit π une projection régulière de l'enveloppe convexe de V sur V. Alors $\mathbb{B}[X] = \pi(\mathbb{E}[X])$ est un barycentre associé au connecteur

$$\gamma(x, y) = \pi'(x)(y - x). \tag{1.3}$$

En effet la condition (*a*) se déduit immédiatement du caractère lipschitzien de π; quant à la condition (*c*), on remarque que

$$\pi'(\pi(y))(y - \pi(y)) = O(|y - \pi(y)|^2)$$

et que la distance entre $\mathbb{E}[X]$ et V est $O(\operatorname{var} X)$ donc

$$\begin{aligned}\mathbb{E}\,\gamma(\mathbb{B}[X], X) &= \pi'(\mathbb{B}[X])(\mathbb{E}[X] - \mathbb{B}[X]) \\ &= O((\mathbb{E}[X] - \mathbb{B}[X])^2) = O((\operatorname{var} X)^2).\end{aligned}$$

D'après (1.3), deux projections ayant les mêmes dérivées sur V fournissent des barycentres associés au même connecteur. Dans le cas particulier où les fibres $\pi^{-1}(\{x\})$ sont des parties d'espaces affines, $\mathbb{B}$ est le γ-barycentre. Plus généralement, si π n'est pas défini sur toute l'enveloppe convexe de V mais seulement sur un voisinage de V, on obtient un barycentre qui n'est défini que pour les variables de variance suffisamment petite; c'est le cas par exemple si V est une sphère et π la projection orthogonale.

Exemple 2. – Le γ-barycentre d'un connecteur métrique pour une distance δ est l'ensemble des points critiques de la fonction $x \mapsto \mathbb{E}\,\delta^2(x, X)$. Si pour tout X, cette fonction admet un et un seul point critique, et si ce point critique est un minimum non dégénéré, alors le γ-barycentre est bien défini (le caractère lipschitzen du barycentre provient du caractère non dégénéré du minimum). Nous obtenons ainsi le barycentre métrique

$$\mathbb{B}[X] = \arg\min_x \mathbb{E}\,\delta^2(x, X). \tag{1.4}$$

Cette définition est très naturelle et a été considérée depuis longtemps (*voir* [15]). Si δ est la distance induite par un plongement de V dans $\mathbb{R}^n$, nous retrouvons le cas de l'exemple 1 avec π égal à la projection orthogonale sur V. Remarquons que deux distances dont le quotient tend vers 1 sur la diagonale (comme par exemple la distance calculée dans $\mathbb{R}^n$ et la distance riemannienne sur V induite par le même plongement de Vdans $\mathbb{R}^n$) fournissent des barycentres associés à des connecteurs différents mais à la même connexion.

Exemple 3. – Le γ-barycentre d'un connecteur géodésique (s'il est bien défini) sera appelé barycentre géodésique (ou exponentiel dans [12]). Si V est riemannienne, nous retrouvons un barycentre métrique que nous pouvons appeler barycentre riemannien dans ce cas; il est montré dans [19] que ce barycentre est bien défini lorsque V est de rayon suffisamment petit.

D'autres types de barycentres non linéaires, définis non plus comme des points mais comme des ensembles ont été considérés dans la littérature; nous allons rapidement en décrire deux. Le premier, étudié dans [12], sera appelé barycentre convexe. Si nous nous donnons une connexion sur V, nous pouvons en déduire une notion de fonction convexe: ce sont les fonctions qui sont convexes sur les géodésiques. Le barycentre convexe $C[X]$ d'une variable X est alors défini comme l'ensemble des x tels que $h(x) \leqq \mathbb{E}\, h(X)$ pour toute fonction convexe bornée h; ce barycentre est constitué de la variété V entière si les fonctions convexes bornées sont constantes; il contient le barycentre géodésique de l'exemple 3 (proposition 2 de [12]); il n'y a cependant pas de raison pour que les autres barycentres répondant à notre définition y appartiennent. Avec une hypothèse de convexité sur la variété, il est égal à l'ensemble des valeurs initiales des martingales continues dont la valeur finale a même loi que X (théorème 2 de [12]). On peut associer une notion de martingale à ce barycentre: ce sont les processus X_t tels que $h(X_t)$ soit une sous-martingale pour toute fonction convexe bornée h. Pour que ces martingales aient les mêmes propriétés que les martingales que nous allons décrire dans cet article, il faudrait (*voir* condition (1.2)) que

$$\sup_{x \in C[X]} \mathbb{E}\, \gamma(x, X) = o(\operatorname{var} X)$$

pour un connecteur γ. Or ceci n'est possible que si $C[X]$ est inclus dans le γ-barycentre (ce qui n'est en général pas le cas); soit en effet X une variable et x un point dans $C[X]$ mais pas dans le γ-barycentre; soit A_n une suite d'événements indépendants de X et de probabilité $1/n$; soit Y_n égal à X sur A_n et à x sur son complémentaire; alors x est dans le barycentre convexe de Y_n, la variance de Y_n est d'ordre $1/n$, mais

$$\mathbb{E}\, \gamma(x, Y_n) = \frac{1}{n}\, \mathbb{E}\, \gamma(x, X) \neq o(1/n).$$

Ainsi les fonctions convexes, qui peuvent servir de base à la définition des martingales continues (*voir* [3]), semblent mal adaptées au cas avec sauts; nous verrons cependant qu'elles jouent un rôle important dans la construction et l'approximation des martingales. Un autre barycentre, étudié dans [6] et qui ne dépend que de la donnée d'une distance δ sur V, est défini comme l'ensemble des x tels que

$$\delta(a, x) \leqq \mathbb{E}\, \delta(a, X)$$

pour tout $a \in V$. les fonctions convexes de l'exemple précédent sont donc renplacées par les fonctions distance à un point; en particulier, si δ est une distance riemannienne à courbures sectionnelles négatives ou nulles et si V est simplement connexe, ce barycentre contient le barycentre convexe; il ne correspond donc pas mieux à notre notion; de plus il peut être vide. Remarquons toutefois que les martingales à temps discret correspondant à ce barycentre vérifient certaines propriété (*voir* [6]); de plus, cette notion peut être généralisée à des variables à valeurs ensemblistes, et conduit à une notion de martingale à valeur ensemblistes : voir [16].

2. PROPRIÉTÉS DES BARYCENTRES

Dans ce qui suit, δ sera une distance riemannienne quelconque sur V, les propriétés étant indépendantes du choix de cette distance. Le champ de normes sur $T(V)$ associé à δ sera noté $|\cdot|$ et sera utilisé pour estimer les vecteurs tangents.

PROPOSITION 2.1. – *Supposons que V est plongé dans un espace euclidien. Si la variance de X tend vers* 0,

$$\mathbb{B}[X] - \mathbb{E}[X] = O(\operatorname{var} X).$$

Remarque. – Cette proposition montre en particulier que deux barycentres ne peuvent différer que de $O(\operatorname{var} X)$.

Démonstration. – En appliquant

$$\gamma(x, y) = y - x + O(|y - x|^2)$$

à $\mathbb{B}[X]$ et X, et en prenant l'espérance, nous obtenons

$$\mathbb{E}\gamma(\mathbb{B}[X], X) = \mathbb{E}[X] - \mathbb{B}[X] + O(\mathbb{E}[|X - \mathbb{B}[X]|^2]).$$

D'après (1.2), le membre de gauche est $o(\operatorname{var} X)$; de plus

$$\mathbb{E}[|X - \mathbb{B}[X]|^2] = |\mathbb{E}[X] - \mathbb{B}[X]|^2 + \operatorname{var} X$$

donc

$$\mathbb{E}[X] - \mathbb{B}[X] = O(|\mathbb{E}[X] - \mathbb{B}[X]|^2) + O(\operatorname{var} X).$$

Comme d'autre part, les conditions (*a*) et (*b*) de la définition 1.2 impliquent que $\mathbb{E}[X] - \mathbb{B}[X]$ tend vers 0 lorsque $\operatorname{var} X \to 0$, nous en déduisons la proposition. □

PROPOSITION 2.2. – *Si la variance de X tend vers* 0,

$$\mathbb{E}\,\delta^2(\mathbb{B}[X],\, X) = O(\operatorname{var} X).$$

Démonstration. – En plongeant V dans un espace euclidien, il est clair que cette propriété est satisfaite si on remplace $\mathbb{B}[X]$ par $\mathbb{E}[X]$. Il suffit donc d'appliquer l'inégalité du triangle et la proposition 2.1 pour conclure. □

PROPOSITION 2.3. – *Si x et X décrivent respectivement V et l'ensemble des variables à valeurs dans V, alors*

$$\mathbb{E}\,\gamma(x,\, X) - \gamma(x,\, \mathbb{B}[X]) = \delta(x,\, \mathbb{B}[X])\, O(\operatorname{var} X) + o(\operatorname{var} X) \tag{2.1}$$

lorsque la variance de X tend vers 0.

Démonstration. – Notons $\phi(x)$ le membre de gauche de (2.1). Nous allons le dériver. En notant D_1 et D_2 les opérateurs de différentiation par rapport aux première et deuxième variables, nous obtenons

$$\begin{aligned}\phi'(x) &= \mathbb{E}\,[D_1\,\gamma(x,\, X) - D_1\,\gamma(x,\, \mathbb{B}[X])]\\ &= \mathbb{E}\,[D_2\, D_1\,\gamma(x,\, \mathbb{B}[X])\,\langle\,\gamma(\mathbb{B}[X],\, X)\rangle + O(\delta^2(\mathbb{B}[X],\, X))].\end{aligned}$$

Les espérances de $\delta^2(\mathbb{B}[X],\, X)$ et $\gamma(\mathbb{B}[X],\, X)$ sont respectivement $O(\operatorname{var} X)$ et $o(\operatorname{var} X)$, donc $\phi'(x)$ est $O(\operatorname{var} X)$; d'autre part, $\phi(\mathbb{B}[X])$ est $o(\operatorname{var} X)$ et le théorème des accroissements finis permet de conclure. □

PROPOSITION 2.4. – *Si x et X décrivent respectivement V et l'ensemble des variables à valeurs dans V et si $\mathbb{E}\,\delta^2(x,\, X)$ est suffisamment petit, alors il existe une constante $C > 0$ telle que*

$$|\,\mathbb{E}\,\gamma(x,\, X)\,| \leqq C\,\delta(x,\, \mathbb{B}[X]) + o(\operatorname{var} X)$$

et

$$\delta(x,\, \mathbb{B}[X]) \leqq C\,|\,\mathbb{E}\,\gamma(x,\, X)\,| + o(\operatorname{var} X)$$

lorsque la variance de X tend vers 0.

Démonstration. – Remarquons que par régularité de γ,

$$|\,\gamma(x,\, \mathbb{B}[X])\,| \leqq C\,\delta(x,\, \mathbb{B}[X]),$$

donc la proposition 2.3 fournit immédiatement la première estimation. Pour la seconde, remarquons que si $\mathbb{E}\,\delta^2(x,\, X)$ tend vers 0, alors la distance entre x et $\mathbb{B}[X]$ tend vers 0; lorsque celle-ci est suffisamment petite, elle est dominée par $|\gamma(x, \mathbb{B}[X])|$. D'autre part, d'après la proposition 2.3,

$$\gamma(x, \mathbb{B}[X]) = \mathbb{E}\gamma(x, X) + |\gamma(x, \mathbb{B}[X])|\, O(\operatorname{var} X) + o(\operatorname{var} X),$$

d'où

$$|\gamma(x, \mathbb{B}[X])| \leqq C\,|\mathbb{E}\gamma(x, X)| + o(\operatorname{var} X). \quad \square$$

PROPOSITION 2.5. – *Supposons que le γ-barycentre $\mathbb{B}$ soit bien défini. Alors :*

(i) *les restes $o(\operatorname{var} X)$ peuvent être supprimés dans les estimations de la proposition* 2.4;

(ii) *si x et X décrivent respectivement V et l'ensemble des variables à valeurs dans V, la convergence vers* 0 *de $\delta(x, \mathbb{B}[X])$ est équivalente à la convergence vers* 0 *de $\mathbb{E}\gamma(x, X)$.*

Démonstration. – La propriété (i) se montre en reprenant les démonstrations des propositions 2.3 et 2.4 dans le cas du γ-barycentre. Montrons (ii). Si $x = \mathbb{B}[X]$, l'espérance de $\gamma(x, X)$ est nulle, donc par continuité uniforme de γ, on en déduit aisément que la première convergence implique la seconde. Pour la réciproque, raisonnons par l'absurde et supposons qu'il existe une suite (x_n, X_n) telle que $\mathbb{E}\gamma(x_n, X_n)$ converge vers 0 et $\delta(x_n, \mathbb{B}[X_n])$ reste minoré; quitte à prendre une sous-suite, on peut supposer que x_n converge vers un point x et que X_n converge en loi vers une variable X; alors $\mathbb{E}\gamma(x, X) = 0$ et $x \neq \mathbb{B}[X]$, ce qui est contradictoire. $\square$

3. BARYCENTRES ET FONCTIONS CONVEXES

Dans ce paragraphe, nous définissons une notion de fonction $\mathbb{B}$-convexe, étudions les rapports avec la notion de fonction convexe pour la connexion associée à $\mathbb{B}$, et donnons deux exemples pour lesquels nous pouvons démontrer une propriété de convexité de la variété qui sera utile par la suite. Rappelons qu'étant donnée une connexion, une fonction f est dite convexe si sa restriction à tout chemin géodésique est convexe, c'est-à-dire si $t \mapsto f(\exp_x(tu))$ est convexe sur son domaine de définition pour tout $x \in V$ et tout $u \in T_x(V)$; si f est régulière, cela équivaut à dire que son hessien est semi-défini positif en tout point.

DÉFINITION 3.1. – *Soit f une fonction réelle borélienne bornée définie sur V. Nous dirons que f est $\mathbb{B}$-convexe si pour toute variable aléatoire X à valeurs dans V,*

$$f(\mathbb{B}[X]) \leqq \mathbb{E} f(X).$$

Si V est un compact convexe de $\mathbb{R}^n$ avec $\mathbb{B}[X] = \mathbb{E}[X]$, cette notion correspond à la notion classique de fonction convexe. Nous pouvons généraliser cette comparaison.

PROPOSITION 3.2. – *Soit $\mathbb{B}$ un barycentre et f une fonction $\mathbb{B}$-convexe régulière; alors f est convexe pour la connexion associée à $\mathbb{B}$. Réciproquement, si f est une fonction régulière convexe pour une connexion et si $\mathbb{B}$ est le barycentre géodésique (voir exemple 3 du § 1) associé à cette connexion (en supposant qu'il est bien défini), alors f est $\mathbb{B}$-convexe.*

Remarques. – Pour la seconde partie de la proposition, nous ne pouvons rien dire pour les barycentres autres que le barycentre géodésique. Les barycentres géodésiques sont donc ceux qui ont le plus de fonctions convexes : si une fonction est convexe pour un barycentre $\mathbb{B}$, elle l'est aussi pour le barycentre géodésique défini à partir de la connexion associée à $\mathbb{B}$. En ce qui concerne la régularité, une petite modification de la démonstration montre qu'il suffit que f soit lipschitzienne; pour la réciproque, l'hypothèse de régularité peut être omise pourvu qu'on se limite à des barycentres dans l'intérieur de V : on utilise pour cela les propriétés des fonctions convexes de (13) (*voir* la proposition 2 de [12]).

Démonstration. – Soit f une fonction $\mathbb{B}$-convexe régulière, $x \in V \backslash \partial V$ et $u \in T_x(V)$; alors d'après (1.1),

$$\operatorname{Hess} f(x)\langle u, u\rangle = \lim_{t \to 0} \frac{1}{t^2} \left(f(\gamma_x^{-1}(tu)) + f(\gamma_x^{-1}(-tu)) - 2 f(x)\right)$$

et il nous faut montrer que cette expression est positive ou nulle. Soit X une variable prenant les valeurs $\gamma_x^{-1}(tu)$ et $\gamma_x^{-1}(-tu)$ avec probabilité 1/2; alors l'espérance de $\gamma(x, X)$ est nulle. En appliquant la proposition 2.4, nous obtenons

$$\delta(x, \mathbb{B}[X]) = o(\operatorname{var} X) = o(t^2).$$

D'autre part, comme f est $\mathbb{B}$-convexe,

$$f(\mathbb{B}[X]) \leqq \frac{1}{2} \left(f(\gamma_x^{-1}(tu)) + f(\gamma_x^{-1}(-tu))\right)$$

donc

$$f(x) \leqq \frac{1}{2} \left(f(\gamma_x^{-1}(tu)) + f(\gamma_x^{-1}(-tu))\right) + o(t^2)$$

et l'inégalité recherchée est démontrée. Pour la réciproque, en notant γ le connecteur géodésique, la convexité de f implique que

$$f(y) \geqq f(x) + f'(x)\gamma(x, y).$$

En appliquant cette inégalité à $x = \mathbb{B}[X]$, $y = X$, et en prenant l'espérance, nous obtenons le résultat. □

Remarque. – Hormis le cas du barycentre géodésique, la $\mathbb{B}$-convexité n'est pas une notion locale; pour pouvoir parler d'une fonction $\mathbb{B}$-convexe sur une partie B de V, il faut que le barycentre d'une variable portée par B soit dans B; nous verrons cependant dans la démonstration du lemme 8.1 que en un certain sens, tout point de V a un voisinage sur lequel il existe une fonction convexe non constante.

Considérons maintenant la variété-produit $V \times V$; si V est munie d'un barycentre $\mathbb{B}$, on peut définir sur $V \times V$ le barycentre

$$(\mathbb{B} \times \mathbb{B})[(X, Y)] = (\mathbb{B}[X], \mathbb{B}[Y]).$$

La connexion associée à ce barycentre est la connexion produit, et si $\mathbb{B}$ est le barycentre géodésique, $\mathbb{B} \times \mathbb{B}$ l'est aussi. L'existence de fonctions convexes sur $V \times V$ qui sont nulles sur la diagonale et positives ailleurs est une propriété qui sera très importante pour la construction des martingales; nous demandons de plus à ces fonctions d'être des puissances de fonctions distances.

DÉFINITION 3.3. – *Soit $p \geqq 1$; nous dirons que $(V, \mathbb{B})$ est à géométrie p-convexe s'il existe une distance δ sur V équivalente aux distances riemanniennes, et telle que pour tout couple (X, Y) de variables,*

$$\delta^p(\mathbb{B}[X], \mathbb{B}[Y]) \leqq \mathbb{E}\,\delta^p(X, Y).$$

Cela signifie que la fonction δ^p est convexe sur $V \times V$ pour le barycentre produit. D'après les propriétés précédentes, si une variété est à géométrie p-convexe pour $\mathbb{B}$, elle l'est aussi pour le barycentre géodésique correspondant à $\mathbb{B}$. Cette définition est plus forte que la définition de [19] d'une variété à géométrie convexe dans laquelle on suppose seulement l'existence d'une fonction f convexe bornée sur $V \times V$, nulle sur la diagonale et strictement positive ailleurs. Nous allons maintenant donner deux exemples dans lesquels on peut montrer la p-convexité de la géométrie; en fait nous ne connaissons pas d'exemple de géométrie qui soit convexe sans être p-convexe pour un $p \geqq 1$.

PROPOSITION 3.4. – *Pour $\eta > 0$, considérons la calotte sphérique*

$$V = \{(x_1, \ldots, x_d, z) \in \mathbb{R}^{d+1};\ \sum x_i^2 + z^2 = 1,\ z \geqq \eta\}$$

plongée dans $\mathbb{R}^{d+1}$ *et le barycentre défini par la projection orthogonale de* $\mathbb{E}[X]$ *sur* V. *Pour tout* $p > 1/\eta$, V *est à géométrie p-convexe.*

Cette proposition sera démontrée en utilisant une forme infinitésimale de la convexité qu'il est souvent plus facile de vérifier que la définition. Soit $L^p(\Omega, T_X(V))$ l'ensemble des variables aléatoires Y à valeurs dans $T(V)$ telles que $Y(\omega)$ est tangent à V en $X(\omega)$, et pour une (ou toute) métrique régulière sur V, la variable $|Y|$ est dans $L^p(\Omega, \mathbb{R})$. Nous pouvons alors définir la norme $\|Y\|_p$ de Y dans $L^p(\Omega, T_X(V))$ comme étant la norme de $|Y|$ dans $L^p(\Omega, \mathbb{R})$ (cela dépend du choix de la métrique). Nous dirons que le barycentre est différentiable si pour toute variable X telle que $\mathbb{B}[X] = x$, il existe une application linéaire $J(X)$ de $L^\infty(\Omega, T_X(V))$ dans $L^1(\Omega, T_x(V))$ satisfaisant la condition suivante : si X^α est une famille de variables telle que $\alpha \mapsto X^\alpha$ est régulière à dérivées bornées,

$$X_0 = X \qquad \text{et} \qquad \left.\frac{\partial X^\alpha}{\partial \alpha}\right|_{\alpha=0} = U$$

alors

$$\left.\frac{\partial}{\partial \alpha}\mathbb{B}[X^\alpha]\right|_{\alpha=0} = \mathbb{E}[J(X)U].$$

Nous allons montrer comment l'estimation de $J(X)$ permet de montrer la convexité de la géométrie et en déduire la proposition.

LEMME 3.5. – *Supposons que* $\mathbb{B}$ *est différentiable et qu'il existe sur* V *une métrique riemannienne pour laquelle*

$$|\mathbb{E}[J(X)U]| \leqq \|U\|_p$$

pour tout X *et tout* U *de* $L^\infty(\Omega, T_X(V))$. *Alors la géométrie est p-convexe.*

Démonstration. – Soit δ la distance associée à la métrique; soit X et Y deux variables de barycentres x et y; en utilisant un chemin presque minimisant de X à Y, on en déduit, pour tout $\varepsilon > 0$ fixé, l'existence d'une famille de variables X^α, $0 \leqq \alpha \leqq 1$, régulière par rapport à α et telle que

$$|\partial X^\alpha / \partial \alpha| \leqq \delta(X, Y) + \varepsilon.$$

Si $x(\alpha)$ est le barycentre de X^α, on en déduit que

$$\left|\frac{dx(\alpha)}{d\alpha}\right| = \left|\mathbb{E}\left[J(X^\alpha)\frac{\partial X^\alpha}{\partial \alpha}\right]\right| \leqq \left\|\frac{\partial X^\alpha}{\partial x}\right\|_p \leqq \|\delta(X, Y)\|_p + \varepsilon,$$

donc

$$\delta(x, y) \leqq \|\delta(X, Y)\|_p + \varepsilon.$$

Il suffit alors de faire tendre ε vers 0. □

Démonstration de la proposition 3.4. – Nous allons montrer que $\mathbb{B}$ est différentiable et appliquer le lemme 3.5. Soit X^α un chemin régulier de variables dont la vitesse en $\alpha = 0$ est $U \in L^\infty(\Omega, T_X(V))$. En utilisant

$$\mathbb{B}[X^\alpha] = \mathbb{E}[X^\alpha]/|\mathbb{E}[X^\alpha]|.$$

et

$$X^\alpha = X + \alpha U + o(\alpha)$$

on en déduit facilement que $\mathbb{B}[X^\alpha]$ est différentiable par rapport à α en $\alpha = 0$, et que, si $\mathbb{B}[X] = x$, $J(X)U$ est la projection orthogonale de U sur $T_x(V)$ divisée par $|\mathbb{E}[X]|$. En particulier, si $|\cdot|$ désigne la norme riemannienne standard, nous avons

$$|\mathbb{E}[J(X)U]| \leqq \frac{\mathbb{E}|U|}{|\mathbb{E}[X]|}. \tag{3.1}$$

Posons $\varepsilon = 1/p < \eta$, soit q l'exposant conjugué de p, et considérons la métrique

$$|\cdot|_\varepsilon = (z-\varepsilon)^{-1/q}|\cdot|.$$

Alors, on notant Z et z les dernières coordonnées de X et x, l'estimation (3.1) s'écrit

$$\begin{aligned} |\mathbb{E}[J(X)U]|_\varepsilon &\leqq \frac{1}{(z-\varepsilon)^{1/q}} \frac{\mathbb{E}[|U|_\varepsilon (Z-\varepsilon)^{1/q}]}{|\mathbb{E}[X]|} \\ &= \frac{z}{(z-\varepsilon)^{1/q}} \frac{\mathbb{E}[|U|_\varepsilon (Z-\varepsilon)^{1/q}]}{\mathbb{E}[Z]}. \end{aligned}$$

Par l'inégalité de Hölder,

$$|\mathbb{E}[J(X)U]|_\varepsilon \leqq \frac{z}{(z-\varepsilon)^{1/q}} \frac{\mathbb{E}[Z-\varepsilon]^{1/q}}{\mathbb{E}[Z]} \mathbb{E}[|U|_\varepsilon^p]^{1/p}.$$

Comme la fonction $z \mapsto z^{-1}(z-\varepsilon)^{1/q}$ est croissante sur $(\varepsilon, 1)$ et $\mathbb{E}[Z] \leqq z$, on en déduit

$$|\mathbb{E}[J(X)U]|_\varepsilon \leqq \mathbb{E}[|U|_\varepsilon^p]^{1/p},$$

donc l'hypothèse du lemme 3.5 est vérifiée pour la métrique $|\cdot|_\varepsilon$. □

Nous allons maintenant considérer le cas d'une variété riemannienne munie du barycentre riemannien défini par (1.4). Pour montrer la p-convexité de la géométrie, il faut trouver une distance δ_0 dont la puissance p-ième soit $\mathbb{B}$-convexe sur $V \times V$. D'après la proposition 3.2, il suffit que δ_0^p soit convexe le long des géodésiques de $V \times V$; de plus, l'existence d'une telle distance implique aussi que le barycentre riemannien ne peut pas contenir plus d'un point (théorème 7.3 de [19] ou proposition 5 de [12]). Sur une variété simplement connexe à courbures sectionnelles négatives ou nulles, il est bien connu que la distance riemannienne δ est convexe; cependant, dès que la courbure devient positive, aucune puissance de δ n'est convexe, et il faut donc chercher une autre distance.

PROPOSITION 3.6. – *Soit V une boule fermée de rayon R dans une variété riemannienne complète sans bord. Nous supposons que deux points de V sont reliés par une géodésique de V et une seule, et que les courbures sectionnelles sont strictement inférieures à $\pi^2/(4R^2)$. Alors le barycentre riemannien est bien défini et est à géométrie p-convexe pour un $p \geqq 1$.*

Remarque. – Nous retrouvons le cas des boules géodésiques régulières de [19]. Cependant la fonction convexe que nous allons construire sur $V \times V$ sera différente car elle doit être une puissance de fonction distance. De plus, bien qu'il s'agisse d'un résultat de nature géométrique, nous allons en donner une démonstration probabiliste utilisant la théorie des martingales continues sur V (en fait les seules martingales que nous utiliserons seront des mouvements browniens sur des géodésiques).

Démonstration. – Dans cette démonstration, l'espace de probabilité sera l'espace de Wiener standard; δ désignera la distance riemannienne. L'hypothèse sur la courbure implique qu'il n'y a pas de couple de points conjugués sur V; en particulier les géodésiques dépendent de façon régulière de leurs extrémités. Soit O le centre de la boule V; le fait que le barycentre riemannien est bien défini est démontré dans [19] au moins pour une boule ouverte (le point le plus délicat est la convexité de la géométrie que nous allons redémontrer); pour obtenir le cas d'une boule fermée, il suffit d'agrandir un peu V en une boule ouverte, et de remarquer que si X est dans V, la convexité de $\delta(O, x)$ implique que son barycentre riemannien y est aussi. Soit K une constante strictement positive dominant les courbures sectionnelles de V, et telle que $K < \pi^2/(4R^2)$. Considérons sur $V \times V$ et sur V les fonctions

$$f(x, y) = \frac{2}{\sqrt{K}} \sin\left(\frac{\sqrt{K}}{2}\,\delta(x, y)\right)$$

et

$$g(x) = \cos\left(\sqrt{Kq}\,\delta(O, x)\right)$$

où la constante $q > 1$ est choisie pour que $g > 0$ sur V. L'hypothèse sur l'existence et l'unicité des géodésiques implique que f et g sont régulières (hors de la diagonale pour f); de plus des estimations sur les dérivées de la fonction distance (lemmes 1.1.1 et 1.2.1 de [25], voir aussi la démonstration du théorème 2.2.1 dans ce même article) montrent que les hessiens vérifient les estimations

$$\operatorname{Hess} f + \frac{K}{2}\, f \geqq 0 \qquad \text{et} \qquad \operatorname{Hess} g + K\, qg \leqq 0.$$

On remarque également que f est équivalent à δ au voisinage de la diagonale. Nous voulons trouver une distance riemannienne δ_0 telle que δ_0^p soit convexe le long de toute géodésique (x_0^α, x_1^α), $0 \leqq \alpha \leqq 1$, de $V \times V$; pour cela il suffit que

$$\delta_0^p\,(x_0^{1/2}, x_1^{1/2}) \leqq \frac{1}{2}\,(\delta_0^p\,(x_0^0, x_1^0) + \delta_0^p\,(x_0^1, x_1^1)). \tag{3.2}$$

Fixons une telle géodésique, et soient x_β^0 et x_β^1, $0 \leqq \beta \leqq 1$, des chemins réguliers allant respectivement de x_0^0 à x_1^0 et de x_0^1 à x_1^1 (ces chemins seront choisis ultérieurement); soit x_β^α la géodésique de x_β^0 à x_β^1. Soit B_t un mouvement brownien réel issu de 1/2 et arrêté au temps d'atteinte τ de 0 ou 1, et posons $M_t^\beta = x_\beta^{B_t}$; alors M_t^β est une famille de martingales continues sur V. Cela implique en particulier (c'est même la définition des martingales continues dans [11]) que le processus

$$g\,(M_t^\beta) - \frac{1}{2}\int_0^t \operatorname{Hess} g\,(M_s^\beta)\,(dM_s^\beta, dM_s^\beta)$$

est une martingale locale. Or

$$|\, dM_t^\beta\,|^2 = \delta^2\,(x_\beta^0, x_\beta^1)\, 1_{\{t \leqq \tau\}}\, dt$$

et, en utilisant l'estimation sur le hessien de g, nous en déduisons que

$$g\,(M_t^\beta) + \frac{K\,q}{2}\,\delta^2\,(x_\beta^0, x_\beta^1)\int_0^{t\wedge\tau} g\,(M_s^\beta)\,ds$$

est une surmartingale locale. La formule de Ito montre alors que

$$g\,(M_t^\beta)\exp\left(\frac{K\,q}{2}\,\delta^2\,(x_\beta^0, x_\beta^1)\,(t\wedge\tau)\right)$$

est une surmartingale locale. Comme elle est positive, c'est une surmartingale et

$$\mathbb{E}\left[g\,(M_\tau^\beta)\exp\left(\frac{K\,q\,\tau}{2}\,\delta^2\,(x_\beta^0, x_\beta^1)\right)\right] \leqq g\,(M_0^\beta). \tag{3.3}$$

En procédant de la même façon, on montre que le processus

$$\frac{1}{\beta' - \beta}\, f\,(M_t^\beta, M_t^{\beta'})\exp\left[\frac{K}{4}\,(\delta^2\,(x_\beta^0, x_\beta^1) + \delta^2\,(x_{\beta'}^0, x_{\beta'}^1))\,(t\wedge\tau)\right]$$

est une sous-martingale locale. D'après (3.3), comme g est minorée et majorée par des constantes positives, l'exponentielle est majorée par une variable bornée dans L^q uniformément en (β, β'), donc en particulier notre processus est une sous-martingale de la classe (D). De plus, M_t^β étant régulier par rapport à β, le terme devant l'exponentielle est majoré par une constante uniformément en (β, β'), donc, en faisant tendre β' vers β, on peut passer à la limite et on montre que le processus

$$S_t^\beta = \left| \frac{\partial M_t^\beta}{\partial \beta} \right| \exp\left(\frac{K}{2} \, \delta^2 (x_\beta^0, x_\beta^1) (t \wedge \tau) \right)$$

est une sous-martingale de la classe (D). Donc

$$\left| \frac{\partial M_0^\beta}{\partial \beta} \right| \leqq \mathbb{E} \left[\left| \frac{\partial M_\tau^\beta}{\partial \beta} \right| \exp\left(\frac{K \tau}{2} \, \delta^2 (x_\beta^0, x_\beta^1) \right) \right].$$

Soit p l'exposant conjugué de q; par l'inégalité de Hölder,

$$\begin{aligned} \left| \frac{\partial M_0^\beta}{\partial \beta} \right| &\leqq \left\| \frac{1}{g^{1/q} (M_\tau^\beta)} \frac{\partial M_\tau^\beta}{\partial \beta} \right\|_p \left\| g^{1/q} (M_\tau^\beta) \exp\left(\frac{K \tau}{2} \, \delta^2 (x_\beta^0, x_\beta^1) \right) \right\|_q \\ &\leqq g^{1/q} (M_0^\beta) \left\| \frac{1}{g^{1/q} (M_\tau^\beta)} \frac{\partial M_\tau^\beta}{\partial \beta} \right\|_p \end{aligned} \tag{3.4}$$

d'après (3.3). Divisons la métrique riemannienne par $g^{1/q}$; notons δ_0 la nouvelle distance riemannienne et $| \cdot |_0$ le champ de normes sur $T(V)$ qui lui est associé. Comme M_τ^β vaut x_β^0 sur $\{B_\tau = 0\}$, x_β^1 sur $\{B_\tau = 1\}$, et comme ces deux événements sont de probabilité 1/2, l'estimation (3.4) s'écrit

$$\left| \frac{\partial x_\beta^{1/2}}{\partial \beta} \right|_0^p \leqq \frac{1}{2} \left(\left| \frac{\partial x_\beta^0}{\partial \beta} \right|_0^p + \left| \frac{\partial x_\beta^1}{\partial \beta} \right|_0^p \right). \tag{3.5}$$

C'est maintenant que nous choisissons x_β^0 et x_β^1; nous considérons des chemins qui sont presque minimisants pour la nouvelle métrique, c'est-à-dire que pour $\varepsilon > 0$ fixé nous les prenons tels que

$$\left| \frac{\partial x_\beta^0}{\partial \beta} \right|_0 \leqq \delta_0 (x_0^0, x_1^0) (1 + \varepsilon), \qquad \left| \frac{\partial x_\beta^1}{\partial \beta} \right|_0 \leqq \delta_0 (x_0^1, x_1^1) (1 + \varepsilon).$$

Alors nous déduisons de (3.5) que

$$\delta_0^p (x_0^{1/2}, x_1^{1/2}) \leqq \left(\int_0^1 \left| \frac{\partial x_\beta^{1/2}}{\partial \beta} \right|_0 d\beta \right)^p$$
$$\leqq \frac{(1+\varepsilon)^p}{2} (\delta_0^p (x_0^0, x_1^0) + \delta_0^p (x_0^1, x_1^1)).$$

Nous pouvons conclure en faisant tendre ε vers 0 de façon à obtenir (3.2). □

4. MARTINGALES

A la notion de barycentre nous allons associer une notion de martingale sur un espace de probabilité muni d'une filtration. Nous supposons que pour toute sous-tribu $\mathcal{F}$ de notre espace de probabilité et toute variable X à valeurs dans V, il existe une version régulière de la loi conditionnelle de X sachant $\mathcal{F}$; le barycentre de cette loi conditionnelle sera appelé barycentre conditionnel de X sachant $\mathcal{F}$ et noté $\mathbb{B}[X \mid \mathcal{F}]$; le choix de la version régulière ne modifie $\mathbb{B}[X \mid \mathcal{F}]$ que sur un ensemble négligeable. Nous obtenons alors immédiatement une notion de martingale à temps discret.

DÉFINITION 4.1. – *Un processus* $(X_t,\ t \in \mathbb{N})$ *est appelé martingale si pour tout* t,

$$X_t = \mathbb{B}[X_{t+1} \mid \mathcal{F}_t].$$

La notion de barycentre n'étant pas associative, X_t n'est en général pas le barycentre conditionnel de X_{t+2}; ceci implique que en temps continu, la définition est plus délicate; nous ne pouvons pas imposer à X_s d'être le barycentre conditionnel de X_t pour tout $t \geqq s$. Nous allons tout d'abord définir une notion de γ-martingale pour un connecteur γ. Remarquons que la notion de semimartingale peut se définir de façon naturelle; nous dirons qu'un processus X_t continu à droite limité à gauche (càdlàg) à valeurs dans V est une semimartingale si pour toute fonction régulière f sur V, $f(X_t)$ est une semimartingale réelle.

DÉFINITION 4.2. – *Soit* $(X_t,\ 0 \leqq t \leqq 1)$ *une semimartingale à valeurs dans* V. *Nous dirons que* X_t *est une* γ*-martingale si pour toute métrique riemannienne sur* V,

$$\lim \sum_i |\mathbb{E}[\gamma(X_{t_i}, X_{t_{i+1}}) \mid \mathcal{F}_{t_i}]| = 0$$

en probabilité lorsque le pas $|\sigma|$ *de la subdivision* $\sigma = (t_i)$ *de* $[0, 1]$ *tend vers* 0.

Remarque. – Nous avons défini la notion de martingale sur [0, 1]; en utilisant par exemple le changement de temps $t' = \tan(\pi t/2)$, nous en déduisons les martingales sur $[0, \infty]$; les martingales sur $[0, \infty)$ peuvent être définies comme les processus M_t tels que $M_{t\wedge T}$ est une martingale pour tout T; une telle définition a un sens car il résulte aisément de la définition qu'une γ-martingale arrêtée en un temps fixe est une γ-martingale. En fait les γ-martingales peuvent être arrêtées en des temps optionnels, ainsi que le précise la proposition suivante.

PROPOSITION 4.3. – *Toute γ-martingale arrêtée est une γ-martingale. Plus généralement, si τ_n est une suite de temps d'arrêt croissant de façon stationnaire vers* 1, *la semimartingale X_t est une γ-martingale si et seulement si pour tout $\tau = \tau_n$, le processus arrêté $\bar{X}_t = X_{t\wedge\tau}$ est une γ-martingale.*

Démonstration. – D'après le comportement de τ_n, il est clair que X_t est une γ-martingale si et seulement si pour tout $\tau = \tau_n$,

$$\lim_{|\sigma|\to 0} \sum_i 1_{\{t_i<\tau\}} \,|\, \mathbb{E}\,[\gamma\,(X_{t_i},\, X_{t_{i+1}}) \,|\, \mathcal{F}_{t_i}] \,| = 0.$$

Pour n fixé et $\tau = \tau_n$, $\bar{X}_t$ est une γ-martingale si et seulement si

$$\lim_{|\sigma|\to 0} \sum_i 1_{\{t_i<\tau\}} \,|\, \mathbb{E}\,[\gamma\,(X_{t_i},\, \bar{X}_{t_{i+1}}) \,|\, \mathcal{F}_{t_i}] \,| = 0.$$

D'après la propriété de Lipschitz de γ, il est donc suffisant de montrer que

$$\Delta = \sum_i 1_{\{t_i<\tau\}}\, \mathbb{E}\,[\delta\,(X_{t_{i+1}},\, \bar{X}_{t_{i+1}}) \,|\, \mathcal{F}_{t_i}] \tag{4.1}$$

converge en probabilité vers 0. Or

$$\mathbb{E}\,[\Delta] = \mathbb{E}\,[\sum_i \delta\,(X_{t_{i+1}},\, X_\tau)\, 1_{\{t_i<\tau\leqq t_{i+1}\}}]$$

tend vers 0 par continuité à droite de X et par le théorème de convergence dominée. □

Nous allons maintenant donner deux conditions nécessaires et suffisantes portant sur les caractéristiques prévisibles pour qu'une semimartingale soit une γ-martingale. Avant d'énoncer le résultat, nous avons besoin de quelques notations. Si X_t, $0 \leqq t \leqq 1$, est une semimartingale sur V, nous utiliserons une mesure aléatoire μ sur $[0, 1] \times V \times V$ caractérisant les sauts de X_t; $\mu\,((s, t] \times H_1 \times H_2)$ est le nombre de sauts non nuls du processus X sur l'intervalle de temps $(s, t]$ d'un point de H_1 à un point de H_2. Remarquons qu'en calcul stochastique réel, la mesure des sauts utilisée classiquement donne une masse à l'amplitude du saut $X_u - X_{u-}$

plutôt qu'à (X_{u-}, X_u), mais sur une variété, l'amplitude du saut n'a pas de sens intrinsèque. Nous noterons ν le compensateur prévisible de μ; il est porté par l'ensemble des (t, x, y) tels que $X_{t-} = x$. D'autre part, si ϕ est une fonction régulière qui à tout x de V associe une forme bilinéaire sur $T_x(V)$, nous pouvons construire l'intégrale

$$\int_0^t \phi(X_{s-})(dX_s^c, dX_s^c)$$

de $\phi(X_{s-})$ par rapport à la partie continue de la covariation quadratique de X (*voir* [26]); cette intégrale peut se calculer en plongeant V dans un espace euclidien mais le résultat ne dépend pas du plongement; elle peut aussi se calculer par discrétisation du temps en utilisant un connecteur, mais là aussi, le résultat ne dépend pas du connecteur; cela est lié au fait (*voir* proposition 6.31 de [11]) que $\phi(x)$ est associé à une forme d'ordre 2 en x. Si nous plongeons V dans un espace euclidien $\mathbb{R}^n$, γ peut être vue comme une application de $V \times V$ dans $\mathbb{R}^n$. En la développant au voisinage de la diagonale, on peut associer à ce plongement une et une seule application ρ définie sur V et telle que pour tout x, $\rho(x)$ est une application bilinéaire symétrique sur $T_x(V)$ à valeurs dans $\mathbb{R}^n$ satisfaisant

$$\gamma(x, y) = y - x + \rho(x)\langle\gamma(x, y), \gamma(x, y)\rangle + O(|y - x|^3) \tag{4.2}$$

si $y \to x$. Si V est de dimension d et se représente à l'aide d'une seule carte, cette carte représente un plongement de V dans $\mathbb{R}^d$, et dans ce cas, $2\rho(x)$ est une application bilinéaire symétrique de $\mathbb{R}^d \times \mathbb{R}^d$ dans $\mathbb{R}^d$ dont les coefficients dans la base canonique sont les symboles de Christoffel de la connexion associée à γ. Plus généralement, en comparant avec la relation (1.1) définissant le hessien, $-2\rho(x)$ est le hessien du plongement de V dans $\mathbb{R}^n$.

THÉORÈME 4.4. – *Soit X_t, $0 \leqq t \leqq 1$, une semimartingale sur V. Alors les conditions suivantes sont équivalentes:*

(*a*) *X_t est une γ-martingale;*

(*b*) *si V est plongé dans $\mathbb{R}^n$ et ρ est défini par (4.2), le processus*

$$\begin{aligned} M_t = X_t &+ \int_0^t \rho(X_{s-})(dX_s^c, dX_s^c) \\ &+ \int_0^t \int (\gamma(x, y) - (y - x))\,\nu(ds, dx, dy) \end{aligned} \tag{4.3}$$

est une martingale locale euclidienne;

(*c*) *pour toute fonction régulière définie sur* V *et à valeurs réelles, le processus*

$$M_t^f = f(X_t) - \frac{1}{2}\int_0^t \operatorname{Hess} f(X_{s-})(dX_s^c, dX_s^c)$$
$$- \int_0^t \int (f(y) - f(x) - f'(x)\gamma(x, y))\,\nu(ds, dx, dy) \quad (4.4)$$

est une martingale locale.

Ce résultat (ainsi que les suivants) sera démontré en § 5. La caractérisation (*c*) montre que notre notion de γ-martingale coïncide avec la notion définie au § 4 de [26]; plus précisément, dans [26], nous avons utilisé le connecteur γ pour construire l'intégrale de Ito par rapport à X d'une forme prévisible d'ordre 1 au-dessus de X_{t-}; une semimartingale X_t est alors une γ-martingale si et seulement si toutes les intégrales de Ito par rapport à X sont des martingales locales réelles. Dans le cas d'une partie convexe d'un groupe de Lie, on peut aussi vérifier que X est une martingale si et seulement si son logarithme stochastique (défini dans [14]) est une martingale locale dans l'algèbre de Lie. Si maintenant $\mathbb{B}$ est un barycentre compatible avec le connecteur γ, nous pouvons nous demander si les γ-martingales peuvent se caractériser à l'aide de $\mathbb{B}$. Si X_t est un processus adapté càdlàg, X_t est dit quasi-continu à gauche (qcàg) si $X_\tau = X_{\tau-}$ en tout temps prévisible τ.

THÉORÈME 4.5. – *Soit* $\mathbb{B}$ *un barycentre compatible avec* γ *et soit* X_t, $0 \leqq t \leqq 1$, *une semimartingale sur* V. *Supposons soit que* X_t *est quasi-continu à gauche, soit que* $\mathbb{B}$ *est le* γ*-barycentre (s'il est bien défini). Alors les conditions (a), (b), (c) du théorème 4.4 sont équivalentes à la condition suivante :*

(*d*) *si* $\sigma = (t_i)$ *décrit l'ensemble des subdivisions de* $[0, 1]$,

$$\lim \sum_i \delta(X_{t_i}, \mathbb{B}[X_{t_{i+1}} \mid \mathcal{F}_{t_i}]) = 0$$

en probabilité quand $|\sigma| \to 0$.

Étant donné un barycentre $\mathbb{B}$, nous désirons introduire une notion de $\mathbb{B}$-martingale X_t. D'après le théorème 4.5, lorsque X_t est quasi-continue à gauche ou lorsque $\mathbb{B}$ est le γ-barycentre, il est naturel de demander à X_t d'être une γ-martingale; de façon plus générale, nous demanderons à X_t de se comporter comme une γ-martingale en dehors des temps de saut prévisibles, et le barycentre n'interviendra donc qu'aux temps de saut prévisibles τ. Si X_t est une γ-martingale, la caractérisation (b) montre que

$$\mathbb{E}\,[\gamma\,(X_{\tau-},\,X_\tau)\,|\,\mathcal{F}_{\tau-}] = 0,$$

ce qui signifie que $X_{\tau-}$ est dans le γ-barycentre conditionnel de X_τ sachant $\mathcal{F}_{\tau-}$; dans le cas général, nous demanderons donc à X_t de satisfaire

$$X_{\tau-} = \mathbb{B}\,[X_\tau\,|\,\mathcal{F}_{\tau-}]. \tag{4.5}$$

Cela fournira un cadre unifiant le cas qcàg et le cas du temps discret. Rappelons que μ désigne la mesure des sauts de X et que ν est son compensateur prévisible; ν peut se décomposer en une partie continue ν^c et une partie purement discontinue ν^d représentant respectivement les sauts totalement inaccessibles et prévisibles de X_t; ν^d vérifie la condition d'intégrabilité

$$\sum_t \left| \int h\,(x,\,y)\,\nu^d\,(\{t\},\,dx,\,dy) \right| < \infty$$

pour toute fonction régulière h s'annulant sur la diagonale. En notant $D\,(x,\,\cdot)$ la masse de Dirac en un point x, la masse de ν^d en un temps τ prévisible est donnée par

$$\nu^d\,(\{\tau\},\,dx,\,dy) = 1_{\{y \neq x\}}\,D\,(X_{\tau-},\,dx) \otimes Q_\tau \tag{4.6}$$

où Q_τ est la loi conditionnelle de X_τ sachant $\mathcal{F}_{\tau-}$.

Proposition 4.6. – *Soit* $\mathbb{B}$ *un barycentre associé à un connecteur* γ *et soit* X_t, $0 \leqq t \leqq 1$, *une semimartingale sur* V. *Les conditions suivantes sont équivalentes :*

(b') *si* V *est plongé dans* $\mathbb{R}^n$ *et* ρ *est défini par (4.2), le processus*

$$\begin{aligned} \bar{M}_t = X_t &+ \int_0^t \rho\,(X_{s-})\,(dX_s^c,\,dX_s^c) \\ &+ \int_0^t \int (\gamma\,(x,\,y) - (y - x))\,\nu\,(ds,\,dx,\,dy) \\ &- \sum_{s \leqq t} \int \gamma\,(x,\,y)\,\nu^d\,(\{s\},\,dx,\,dy) \end{aligned} \tag{4.7}$$

est une martingale locale euclidienne;

(c') pour toute fonction régulière définie sur V et à valeurs réelles, le processus

$$\bar{M}_t^f = f(X_t) - \frac{1}{2}\int_0^t \operatorname{Hess} f(X_{s-})(dX_s^c, dX_s^c)$$
$$- \int_0^t \int (f(y) - f(x) - f'(x)\gamma(x, y))\,\nu(ds, dx, dy)$$
$$- \sum_{s \leqq t} f'(X_{s-}) \int \gamma(x, y)\,\nu^d(\{s\}, dx, dy) \tag{4.8}$$

est une martingale locale.

DÉFINITION 4.7. – *Soit X_t, $0 \leqq t \leqq 1$, une semimartingale. Nous dirons que X_t est une $\mathbb{B}$-martingale (ou une martingale) si elle vérifie (4.5) en tout temps prévisible τ et si de plus elle vérifie les conditions (b'), (c') de la proposition 4.6.*

Cas du γ-barycentre ou d'un processus qcàg. – Si X_t est qcàg, la condition (4.5) est toujours vérifiée, et (b'), (c') se résument à (b) et (c); si $\mathbb{B}$ est le γ-barycentre, (b), (c) sont équivalents aux conditions (b'), (c') jointes à (4.5). Donc dans chacun des deux cas, X_t est une $\mathbb{B}$-martingale si et seulement si c'est une γ-martingale, et cette propriété peut se caractériser par la condition (d) du théorème 4.5.

Cas d'une filtration quasi-continue à gauche. – Si $\mathcal{F}_t$ est qcàg, la condition (4.5) implique que toutes les martingales sont qcàg, donc dans ce cas, tous les barycentres associés à un même connecteur ont les mêmes martingales.

Cas continu. – Si X_t est continu, $\nu = 0$ donc les conditions (b') et (c') peuvent se simplifier. En particulier, seule la connexion associée au barycentre intervient (à travers son hessien) dans la caractérisation (c') simplifiée, et nous retrouvons la notion classique de martingale continue (définition 4.2 de [11]). Si toutes les $\mathcal{F}_t$ martingales réelles sont continues, les martingales à valeurs dans V le sont aussi; en effet, on sait déjà par le paragraphe précédent que X_t est qcàg, et si X_t avait un saut totalement inaccessible, le processus $\bar{M}_t$ de la condition (b') aurait le même saut.

Remarque. – On peut se demander si la condition (d) du théorème 4.5 peut être utilisée pour caractériser les $\mathbb{B}$-martingales dans le cas général. C'est vrai si par exemple les temps de sauts sont déterministes, mais faux en général; donnons-en un contre-exemple (*voir* cependant la proposition 5.10 ci-dessous). Soit W_t un processus de Wiener standard arrêté au premier temps d'atteinte τ de 1, et soit X une variable aléatoire sur V indépendante de W et de barycentre x. Soit X_t, $0 \leqq t \leqq \infty$, le processus prenant la valeur x pour $t < \tau$ et X pour $t \geqq \tau$, et soit $\mathcal{F}_t$ la filtration engendrée par (W_t, X_t). Alors X_t est une martingale. D'autre part, soit U une variable de loi uniforme sur [0, 1] indépendante de X et soit $Y_u = x$ si $u < U$, et $Y_u = X$ sinon; soit $y(u)$ le barycentre de Y_u; alors $y(u)$, $0 \leqq u \leqq 1$, est un lacet issu de x, mais si $\mathbb{B}$ n'est pas le γ-barycentre, $y(u)$ n'a aucune raison d'être égal à x pour $0 < u < 1$. De plus

$$\mathbb{B}[X_{t_{i+1}} \mid \mathcal{F}_{t_i}] = y(\mathbb{P}[t_i < \tau \leqq t_{i+1} \mid \mathcal{F}_{t_i}])$$

sur $\{t_i < \tau\}$. Cette probabilité peut être calculée, et si F désigne la fonction de répartition de la loi normale centrée réduite, on obtient

$$\sum_i \delta(X_{t_i}, \mathbb{B}[X_{t_{i+1}} \mid \mathcal{F}_{t_i}]) = \sum_i 1_{\{t_i < \tau\}}\, \delta\left(x,\, y\left(2F\left(\frac{W_{t_i} - 1}{\sqrt{t_{i+1} - t_i}}\right)\right)\right).$$

Si $y(u) \neq x$ pour $0 < u < 1$, cette expression ne converge pas vers 0 quand $|\sigma| \to 0$; en effet, au dernier instant de la subdivision avant τ, le point où est calculée la fonction y n'est ni proche de 0, ni proche de 1. A propos de cet exemple, nous pouvons aussi noter que si nous considérons la filtration engendrée par X_t, alors τ devient totalement inaccessible, et si x n'est pas dans le γ-barycentre de X, X_t n'est plus une martingale; cela signifie qu'une $\mathbb{B}$-martingale n'est pas nécessairement une $\mathbb{B}$-martingale pour sa filtration naturelle; une telle propriété est en revanche satisfaite pour les γ-martingales.

Avant de passer à la démonstration des résultats et à titre d'application, nous allons étudier le cas où X_t est une fonction d'un processus de markov.

PROPOSITION 4.8. – *Supposons que $X_t = h(Z_t)$, où Z_t est un processus de Feller à valeurs dans E de générateur $\mathcal{L}$, et où h est une fonction telle que pour tout y, l'application*

$$\psi_y : z \mapsto \gamma(h(y),\, h(z))$$

est dans le domaine de $\mathcal{L}$, c'est-à-dire que

$$\psi_y(Z_t) - \int_0^t \mathcal{L}\,\psi_y(Z_s)\,ds$$

est une martingale locale dans $T_{h(y)}(V)$; nous supposons de plus que l'application

$$(y,\, z) \mapsto \mathcal{L}\,\psi_y(z)$$

est mesurable bornée et continue en tout point de la diagonale; nous notons $\mathcal{L}_V\, h(z)$ *sa valeur en* $(z,\, z)$*. Alors* X_t *est une martingale pour la filtration engendrée par* Z *si et seulement si* $\mathcal{L}_V\, h(Z_t)$ *est nul pour tout* t*.*

Remarque. – L'opérateur $\mathcal{L}_V$ agit sur les fonctions de E dans V; il coïncide avec l'opérateur construit dans les deux exemples de l'introduction. On obtient ainsi une caractérisation probabiliste des applications harmoniques comme applications transformant le processus de Markov en une martingale sur V. Seul le connecteur intervient car la filtration est quasi-continue à gauche (théorème XIV.46 de [5]) donc les $\mathbb{B}$-martingales et γ-martingales coïncident. Si E est riemannienne et si $\mathcal{L}$ est le laplacien sur E, on retrouve la caractérisation des applications harmoniques de classe C^2 comme transformant le mouvement brownien sur E en une martingale continue sur V. Si $\mathcal{L}$ est le générateur d'une diffusion avec sauts (au sens de [17]) à coefficients continus sur une variété compacte, on caractérise également les applications harmoniques de classe C^2.

Démonstration. – Soit $\sigma = (t_i)$ une subdivision de [0, 1]. Pour tout t_i et $t \geqq t_i$, le processus

$$\psi_{Z_{t_i}}(Z_t) - \int_{t_i}^{t} \mathcal{L}\,\psi_{Z_{t_i}}(Z_s)\, ds$$

est une martingale bornée, donc en $t = t_{i+1}$ nous obtenons

$$\mathbb{E}\left[\gamma(X_{t_i},\, X_{t_{i+1}}) \,\middle|\, \mathcal{F}_{t_i}\right] = \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \mathcal{L}\,\psi_{Z_{t_i}}(Z_t)\, dt \,\middle|\, \mathcal{F}_{t_i}\right].$$

Le théorème de convergence dominée permet alors de montrer que

$$\begin{aligned}\sum_i |\,\mathbb{E}[\gamma(X_{t_i},\, X_{t_{i+1}}) \,|\, \mathcal{F}_{t_i}]\,| &= \sum_i \Big|\int_{t_i}^{t_{i+1}} \mathbb{E}[\mathcal{L}\,\psi_{Z_{t_i}}(Z_t) \,|\, \mathcal{F}_{t_i}]\, dt\,\Big| \\ &= \sum_i (t_{i+1} - t_i)\,|\,\mathcal{L}_V\, h(Z_{t_i})\,| + o(1) \\ &= \int_0^1 |\,\mathcal{L}_V\; h\; (Z_t)\,|\, dt + o(1)\end{aligned}$$

dans L^1 quand $|\,\sigma\,| \to 0$. Nous concluons en appliquant la définition 4.2. □

Remarque sur les barycentres généralisés. – La notion de $\mathbb{B}$-martingale peut être étendue à des barycentres qui ne sont définis que pour des variables à variance suffisamment petite; il suffit que $\mathbb{B}\,[X_\tau \,|\, \mathcal{F}_{\tau-}]$ soit défini. Remarquons tout de même qu'il nous faut toujours définir le connecteur γ sur toute la variété, sinon on ne peut parler que de martingales continues ou à sauts suffisamment petits; il est donc délicat de définir une notion de martingale discontinue sur la sphère avec barycentre riemannien lorsque le processus peut sauter d'un point à un point voisin du point antipodal.

5. DÉMONSTRATION DES RÉSULTATS DE CARACTÉRISATION

Dans notre étude, nous aurons besoin de la notion de processus dominant qui sera utilisée pour localiser les processus.

DÉFINITION 5.1. – *Soit X_t, $0 \leqq t \leqq 1$, une semimartingale réelle bornée (ou plus généralement spéciale) de décomposition canonique $X_t = V_t + M_t$, où V_t est un processus prévisible à variation finie, et M_t est une martingale locale nulle en* 0; *soit A_t un processus croissant prévisible nul en* 0; *nous dirons que A_t est un processus dominant pour X_t s'il existe une constante $C_X > 0$ telle que*

$$|dV_t| + d\langle X, X\rangle_t \leqq C_X\, dA_t.$$

Si X_t est une semimartingale à valeurs dans V, nous dirons que A_t est un processus dominant pour X_t si pour toute fonction régulière f sur V, A_t est dominant pour $f(X_t)$.

Ces deux définitions sont cohérentes car si A_t est dominant pour une semimartingale réelle X_t, il est dominant pour $f(X_t)$ pour toute fonction f régulière à dérivées bornées. Pour trouver un processus dominant pour une semimartingale X_t sur V, il suffit de plonger V dans un espace euclidien et d'additionner des processus dominants pour chacune des composantes de X_t. Réciproquement, un processus dominant pour X_t l'est pour chacune des composantes. Étant donnée une semimartingale X_t, $0 \leqq t \leqq 1$, de processus dominant A_t, il existe une suite de temps d'arrêt τ_n croissant de façon stationnaire vers 1 telle que $A_{t\wedge\tau_n}$ soit borné; alors d'après la proposition 4.3, X_t est une γ-martingale si et seulement si chacune des $X_{t\wedge\tau_n}$ en est une; de même, nous allons vérifier que la condition (*d*) du théorème 4.5 peut être localisée, ce qui nous permettra de limiter notre étude aux semimartingales de processus dominant borné.

LEMME 5.2. – *Soit τ_n une suite de temps d'arrêt croissant de façon stationnaire vers* 1. *Alors X_t vérifie la condition (d) du théorème* 4.5 *si et seulement si pour tout n, le processus $\bar{X}_t = X_{t\wedge\tau_n}$ la vérifie.*

Démonstration. – En suivant le schéma de démonstration de la proposition 4.3, nous sommes amenés à vérifier que

$$\lim_{|\sigma|\to 0} \sum_i 1_{\{t_i<\tau\}}\, \delta\,(\mathbb{B}\,[X_{t_{i+1}}\,|\,\mathcal{F}_{t_i}],\ \mathbb{B}\,[\bar{X}_{t_{i+1}}\,|\,\mathcal{F}_{t_i}]) = 0$$

converge en probabilité vers 0. Or d'après la propriété de Lipschitz des barycentres, cette expression est dominée par la variable Δ de (4.1). □

LEMME 5.3. – *Si A_t est un processus dominant borné pour X_t, alors pour $s \leqq t$,*

$$\mathbb{E}\,[\delta^2\,(X_s,\,X_t)\,|\,\mathcal{F}_s] + \mathrm{var}\,(X_t\,|\,\mathcal{F}_s) \leqq C\,\mathbb{E}\,[A_t - A_s\,|\,\mathcal{F}_s].$$

Pour démontrer ce lemme, en plongeant V dans un espace euclidien, il suffit de démontrer l'estimation pour une semimartingale X_t réelle bornée, et dans ce cas, le résultat se déduit d'inégalités élémentaires utilisant les décompositions de X_t et X_t^2. La proposition 2.4 permet d'obtenir les estimations suivantes.

LEMME 5.4. – *Si A_t est dominant borné pour X_t et si $s \leqq t$, alors*

$$\begin{aligned}\delta\,(X_s,\,\mathbb{B}\,[X_t\,|\,\mathcal{F}_s]) &\leqq C\,|\,\mathbb{E}\,[\gamma\,(X_s,\,X_t)\,|\,\mathcal{F}_s]\,| + o\,(\mathbb{E}\,[A_t - A_s\,|\,\mathcal{F}_s])\\ &\leqq C\,\mathbb{E}\,[A_t - A_s\,|\,\mathcal{F}_s]\end{aligned}$$

et

$$|\,\mathbb{E}\,[\gamma\,(X_s,\,X_t)\,|\,\mathcal{F}_s]\,| \leqq C\,\delta\,(X_s,\,\mathbb{B}\,[X_t\,|\,\mathcal{F}_s\,]) + o\,(\mathbb{E}\,[A_t - A_s\,|\,\mathcal{F}_s])$$

lorsque $\mathbb{E}\,[A_t - A_s\,|\,\mathcal{F}_s]$ tend vers 0.

Pour étudier la notion de martingale, nous allons commencer par ramener la convergence en probabilité à une convergence en moyenne.

LEMME 5.5. – *Soit X_t une semimartingale de processus dominant A_t borné. Alors X_t est une γ-martingale si et seulement si la convergence de la définition 4.2 a lieu en moyenne. De même, dans la condition (d) du théorème* 4.5, *la convergence en probabilité est équivalente à une convergence en moyenne.*

Démonstration. – Il suffit de vérifier que les familles de variables dont on étudie la convergence sont uniformément intégrables; or pour toute subdivision σ, d'après le lemme 5.4, ces variables sont dominées par

$$U_\sigma = \sum_i \mathbb{E}\,[A_{t_{i+1}} - A_{t_i}\,|\,\mathcal{F}_{t_i}].$$

Pour tout B de $\mathcal{F}_1$, si M_t désigne la martingale de valeur finale 1_B, on a

$$\mathbb{E}[U_\sigma 1_B] = \mathbb{E}\sum_i M_{t_i}(A_{t_{i+1}} - A_{t_i}) \leqq \mathbb{E}[A_1 \sup_t M_t] \leqq 2 \| A_1 \|_2 \mathbb{P}[B]^{1/2}$$

par l'inégalité de Doob; la famille U_σ est donc bien uniformément intégrable. □

Fixons maintenant une semimartingale X_t de processus dominant A_t borné, et posons

$$S = \mathbb{E}\int_0^1 \xi_t \, dA_t \tag{5.1}$$

avec

$$\xi_0 = 0, \quad \xi_t = \frac{|\mathbb{E}[\gamma(X_{t_i}, X_{t_{i+1}}) \mid \mathcal{F}_{t_i}]|}{\mathbb{E}[A_{t_{i+1}} - A_{t_i} \mid \mathcal{F}_{t_i}]} \quad \text{si} \quad t_i < t \leqq t_{i+1} \tag{5.2}$$

le processus ξ_t étant borné d'après le lemme 5.4. Alors S est exactement l'espérance de l'expression de la définition 4.2, donc X_t est une γ-martingale si et seulement si S tend vers 0. De même, X_t vérifie la condition (*d*) si et seulement si

$$S' = \mathbb{E}\int_0^1 \xi'_t \, dA_t \tag{5.3}$$

tend vers 0, avec

$$\xi'_0 = 0, \quad \xi'_t = \frac{\delta(X_{t_i}, \mathbb{B}[X_{t_{i+1}} \mid \mathcal{F}_{t_i}])}{\mathbb{E}[A_{t_{i+1}} - A_{t_i} \mid \mathcal{F}_{t_i}]} \quad \text{si} \quad t_i < t \leqq t_{i+1}. \tag{5.4}$$

LEMME 5.6. – *Si* $\mathbb{B}$ *est le* γ*-barycentre, les convergences vers* 0 *de* S *et* S' *vers* 0 *sont équivalentes.*

Démonstration. – Soit λ la mesure sur $[0, 1] \times \Omega$ définie par

$$\lambda(]s, t] \times B) = \mathbb{E}[1_B(A_t - A_s)].$$

Alors S et S' sont les intégrales de ξ et ξ' pour la mesure λ, donc les convergences vers 0 de S et S' sont respectivement équivalentes aux convergences en λ-mesure vers 0 de ξ et ξ'. Soit a_t le dénominateur dans la définition de ξ_t et ξ'_t, c'est-à-dire

$$a_0 = 0, \quad a_t = \mathbb{E}[A_{t_{i+1}} - A_{t_i} \mid \mathcal{F}_{t_i}] \quad \text{si} \quad t_i < t \leqq t_{i+1}. \tag{5.5}$$

Nous allons appliquer la proposition 2.5; la propriété (i) implique (en utilisant aussi le lemme 5.3) que si a est inférieur à une certaine valeur, alors les numérateurs de ξ et ξ' sont équivalents, donc ξ et ξ' eux-mêmes sont équivalents; la propriété (ii) dit que la convergence vers 0 du numérateur de ξ est équivalente à la convergence vers 0 du numérateur de ξ', donc si a est minoré, les convergences vers 0 de ξ et ξ' sont équivalentes; l'addition de ces deux propriétés permet d'obtenir l'équivalence recherchée. □

LEMME 5.7. – *Supposons que A_t est continu. Alors les convergences vers 0 de S et S' sont équivalentes.*

Démonstration. – Reprenons les notations de la démonstration précédente. Le lemme 5.4 montre que

$$\xi'_t \leqq C\,\xi_t + \varepsilon\,(a_t) \qquad \text{et} \qquad \xi_t \leqq C\,\xi'_t + \varepsilon\,(a_t) \tag{5.6}$$

où ε est une fonction tendant vers 0 en 0, donc

$$S' \leqq CS + \int \varepsilon\,(a_t\,(\omega))\,d\lambda\,(t,\,\omega)$$

et

$$S \leqq CS' + \int \varepsilon\,(a_t\,(\omega))\,d\lambda\,(t,\,\omega).$$

On veut donc montrer que l'intégrale de $\varepsilon \circ a$ par rapport à λ converge vers 0; il suffit pour cela que $a_t\,(\omega)$ converge en λ-mesure vers 0. Soit $\alpha\,(u)$ l'oscillation du processus A_t définie par

$$\alpha\,(u) = \sup\,\{A_t - A_s;\; 0 \leqq t - s \leqq u\}.$$

On vérifie facilement que

$$a_t \leqq \sup_s\, \mathbb{E}\,[\alpha\,(\,|\,\sigma\,|\,)\,|\,\mathcal{F}_s],$$

donc par l'inégalité de Doob,

$$\mathbb{E}\,\sup_t\, a_t^2 \leqq 4\,\mathbb{E}\,\alpha^2\,(\,|\,\sigma\,|\,).$$

Comme A est continu, cette expression tend vers 0, donc en particulier, la λ-intégrale de $a_t\,(\omega)$ tend vers 0. □

LEMME 5.8. – *Soit F_t un processus càdlàg prévisible à valeurs dans $\mathbb{R}^n$ et à variation intégrable. Alors*

$$\limsup_{|\,\sigma\,|\to 0}\, \mathbb{E} \sum_i \,|\,\mathbb{E}\,[F_{t_{i+1}} - F_{t_i}\,|\,\mathcal{F}_{t_i}]\,|\, = \mathbb{E} \int_0^1 |\,dF_t\,|.$$

Démonstration. – Ce résultat est une propriété classique des quasimartingales (*voir* [4]); en effet, le membre de droite s'écrit

$$V = \mathbb{E} \int_0^1 |\, dF_t \,| = \sup_H \mathbb{E} \int_0^1 H_t \, dF_t$$

où la borne supérieure porte sur les processus prévisibles H bornés en valeur absolue par 1; l'espérance V_σ du membre de gauche s'exprime de la même façon en se limitant aux processus prévisibles élementaires associés à la subdivision σ; on en déduit que $V_\sigma \leqq V$ et que V_{σ_n} croît vers V le long de toute suite croissante σ_n dont le pas tend vers 0. □

LEMME 5.9. – *Soit F_t la partie à la variation finie prévisible de X_t dans $\mathbb{R}^n$ (X étant de processus dominant borné); soit μ la mesure des sauts de X_t, ν son compensateur prévisible, et ρ défini par (4.2). Alors*

$$\limsup_{|\sigma| \to 0} \mathbb{E} \sum_i |\, \mathbb{E} [\gamma (X_{t_i}, X_{t_{i+1}}) \,|\, \mathcal{F}_{t_i}] \,|$$

$$= \mathbb{E} \int_0^1 \left| dF_t + \rho (X_{t-}) (dX_t^c, dX_t^c) + \int (\gamma (x, y) - (y - x)) \, \nu (dt, dx, dy) \right|.$$

Démonstration. – Par la formule de Ito, en désignant par γ' et γ'' les dérivées de γ par rapport à la seconde variable calculées dans $\mathbb{R}^n$, nous obtenons

$$\begin{aligned} \gamma (X_{t_i}, X_{t_{i+1}}) = & \int_{t_i}^{t_{i+1}} \gamma' (X_{t_i}, X_{t-}) \, dX_t \\ & + \frac{1}{2} \int_{t_i}^{t_{i+1}} \gamma'' (X_{t_i}, X_{t-}) (dX_t^c, dX_t^c) \\ & + \int_{t_i}^{t_{i+1}} \int (\gamma (X_{t_i}, y) - \gamma (X_{t_i}, x) \\ & \qquad - \gamma' (X_{t_i}, x) (y - x)) \, \mu (dt, dx, dy) \end{aligned}$$

donc, comme tout est intégrable (X étant de processus dominant borné),

$$\begin{aligned}
&\mathbb{E}\left[\gamma(X_{t_i}, X_{t_{i+1}}) \mid \mathcal{F}_{t_i}\right] \\
&= \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \gamma'(X_{t_i}, X_{t-})\, dF_t \right.\\
&\qquad + \frac{1}{2}\int_{t_i}^{t_{i+1}} \gamma''(X_{t_i}, X_{t-})(dX_t^c, dX_t^c)\\
&\qquad + \int_{t_i}^{t_{i+1}} \int (\gamma(X_{t_i}, y) - \gamma(X_{t_i}, x)\\
&\qquad\qquad \left. - \gamma'(X_{t_i}, x)(y-x))\, \nu(dt, dx, dy) \,\middle|\, \mathcal{F}_{t_i}\right]. \qquad (5.7)
\end{aligned}$$

Remarquons que dA_t domine uniformément

$$|dF_t|, \ |dX_t^c|^2 \qquad \text{et} \qquad \int |y-x|^2\, \nu(dt, dx, dy).$$

En utilisant les estimations

$$\begin{gathered}
|\gamma'(X_{t_i}, X_{t-}) - I| \leqq C\,|X_{t-} - X_{t_i}|,\\
|\gamma''(X_{t_i}, X_{t-}) - 2\rho(X_{t-})| \leqq C\,|X_{t-} - X_{t_i}|,\\
|(\gamma(X_{t_i}, y) - \gamma(X_{t_i}, x) - \gamma'(X_{t_i}, x)(y-x)) - (\gamma(x, y) - (y-x))|\\
\leqq C\,|y-x|^2\,|x - X_{t_i}|
\end{gathered}$$

ainsi que la convergence

$$\lim \mathbb{E} \sum_i \int_{t_i}^{t_{i+1}} |X_{t-} - X_{t_i}|\, dA_t = 0,$$

nous pouvons en déduire que

$$\begin{aligned}
&\limsup \mathbb{E} \sum_i |\mathbb{E}[\gamma(X_{t_i}, X_{t_{i+1}}) \mid \mathcal{F}_{t_i}]|\\
&\quad = \limsup \mathbb{E} \sum_{\mathrm{i}} \left| \mathbb{E}\left[\int_{t_i}^{t_{i+1}} (dF_t + \rho(X_{t-})(dX_t^c, dX_t^c) \right.\right.\\
&\qquad\qquad \left.\left. + \int (\gamma(x, y) - (y-x))\, \nu(dt, dx, dy)) \,\middle|\, \mathcal{F}_{t_i}\right] \right|.
\end{aligned}$$

Le lemme 5.8 permet alors de conclure. □

Démonstration du théorème 4.4. – Soit A_t un processus dominant pour X_t; soit τ_n une suite de temps d'arrêt croissant de façon stationnaire vers 1 et telle que chacun des A_{τ_n} est borné. Nous avons vu dans la proposition 4.3 que X_t est une γ-martingale si et seulement si chacun des $X_{t \wedge \tau_n}$ en est une; il est clair que les conditions (*b*) et (*c*) sont également localisables, donc nous pouvons nous ramener au cas où A_t est borné. L'équivalence de (*a*) et (*b*) est alors une conséquence immédiate du lemme 5.9, et il nous reste à montrer l'équivalence de (*b*) et (*c*). Supposons que (*c*) est satisfaite; en travaillant composante par composante, elle l'est aussi pour les fonctions f à valeurs vectorielles; en prenant pour f le plongement de V dans $\mathbb{R}^n$, en remarquant que $f' = I$ et Hess $f = -2\,\rho$, nous vérifions que (*c*) implique (*b*). Soit maintenant f une fonction régulière sur V quelconque; nous pouvons la prolonger en une fonction régulière sur $\mathbb{R}^n$, et écrire la formule de Ito

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})\,dX_s + \frac{1}{2}\int_0^1 f''(X_{s-})(dX_s^c,\, dX_s^c) + \int_0^1 \int (f(y) - f(x) - f'(x)(y-x))\,\mu(ds,\, dx,\, dy)$$

où f'' désigne le hessien de f calculé dans $\mathbb{R}^n$; si nous le comparons au hessien calculé pour la connexion de V, nous obtenons

$$\text{Hess}\, f(x) = f''(x) - 2\, f'(x)\,\rho(x).$$

Les processus M_t et M_t^f définis par (4.3) et (4.4) vérifient donc la relation

$$M_t^f = f(X_0) + \int_0^t f'(X_{s-})\,dM_s + \int_0^t \int (f(y) - f(x) - f'(x)(y-x))(\mu - \nu)(dt,\, dx,\, dy). \tag{5.8}$$

Comme d'autre part ν est le compensateur prévisible de μ, M_t^f est une martingale locale si M_t en est une, donc (*b*) implique (*c*), ce qui termine la démonstration du théorème 4.4. □

Démonstration du théorème 4.5. – D'après le lemme 5.2, la condition (*d*) est localisable, donc on peut à nouveau se ramener au cas où A_t est borné. De plus, si X_t est qcàg, on peut choisir A_t continu. Si S et S' sont définis par (5.1) et (5.3), le théorème se ramène donc à l'équivalence des convergences vers 0 de S et S', et ceci a été montré aux lemmes 5.6 et 5.7. □

Démonstration de la proposition 4.6. – L'équivalence de (b') et (c') se démontre comme celle de (b) et (c); pour montrer que (c') implique (b'), on prend pour f le plongement de V dans $\mathbb{R}^n$, et pour montrer la réciproque, on vérifie une relation identique à (5.8) pour $\bar{M}_t$ et $\bar{M}_t^f$. □

Nous terminons ce chapitre par un résultat technique qui nous sera utile dans la suite; dans ce résultat, nous montrons une condition plus faible que (*d*) mais qui est vérifiée par toutes les $\mathbb{B}$-marttingales.

PROPOSITION 5.10. – *Soit X_t, $0 \leq t \leq 1$, une martingale de processus dominant A_t borné. Décomposons $A_t = A_t^c + A_t^d$ en parties continue et purement discontinue. Il existe un $C > 0$ tel que pour $0 \leq s \leq t \leq 1$, si $\sigma = (t_i)$ est une subdivision de $[s, t]$, alors*

$$\mathbb{E}\left[\sum_i \delta\left(X_{t_i}, \mathbb{B}\left[X_{t_{i+1}} | \mathcal{F}_{t_i}\right]\right) \,\middle|\, \mathcal{F}_s\right] \leq C\mathbb{E}\left[A_t^d - A_s^d | \mathcal{F}_s\right] + o(1), \quad (5.9)$$

où $o(1)$ tend vers 0 dans L^1 quand $|\sigma| \to 0$.

Démonstration – Définissons ξ_u, ξ'_u, a_u pour $s \leq u \leq t$ comme en (5.2), (5.4), (5.5) et plongeons V dans $\mathbb{R}^n$. Soit $F_u$la partie à variation finie prévisible de X_u dans $\mathbb{R}^n$. Si nous appliquons la caractérisation (b′) de la notion de martingale, ρ étant défini par (4.2), nous obtenons

$$dF_u = -\rho\left(X_{u-}\right)\left(dX_u^c, dX_u^c\right) - \int \left(\gamma(x, y) - (y - x)\right) \nu\left(du, dx, dy\right)$$
$$+ \int \gamma(x, y)\, \nu^d\left(\{u\}, dx, dy\right).$$

En appliquant le lemme 5.9 sur l'intervalle de temps $[s, t]$ conditionnellement à $\mathcal{F}_s$,

$$\mathbb{E}\left[\int_s^t \xi_u dA_u \,\middle|\, \mathcal{F}_s\right]$$
$$\leq \mathbb{E}\left[\sum_{s<u\leq t} \left| \int \gamma(x, y)\, \nu^d\left(\{u\}, dx, dy\right) \right| \,\middle|\, \mathcal{F}_s\right] + o(1).$$

De plus, d'après (4.6),

$$\left| \int \gamma(x, y)\, \nu^d\left(\{\tau\}, dx, dy\right) \right| = \left| \mathbb{E}\left[\gamma\left(X_{\tau-}, X_\tau\right) | \mathcal{F}_{\tau-}\right] \right| \leq C\Delta A_\tau$$

en tout temps prévisible τ, donc

$$\mathbb{E}\left[\int_s^t \xi_u dA_u \,\middle|\, \mathcal{F}_s\right] \leq C\mathbb{E}\left[A_t^d - A_s^d | \mathcal{F}_s\right] + o\left(1\right). \tag{5.10}$$

D'autre part, nous remarquons que le membre de gauche de (5.9) est égal à l'espérance conditionnelle de $\int_s^t \xi'_u dA_u$, et nous déduisons de (5.6) que

$$\mathbb{E}\left[\int_s^t \xi'_u dA_u \,\middle|\, \mathcal{F}_s\right] \leq C\mathbb{E}\left[\int_s^t \xi_u dA_u \,\middle|\, \mathcal{F}_s\right] + \mathbb{E}\left[\int_s^t \varepsilon\left(a_u\right) dA_u \,\middle|\, \mathcal{F}_s\right].$$

Le premier terme du membre de droite est estimé par (5.10); dans le second, nous décomposons l'intégrale en parties continue et purement discontinue, cette dernière pouvant entrer dans le membre de droite de (5.9). Il est donc suffisant de montrer que

$$\lim_{|\sigma|\to 0} \mathbb{E} \int_s^t \varepsilon\left(a_u\right) dA_u^c = 0,$$

ou encore que

$$\lim_{|\sigma|\to 0} \mathbb{E} \int_s^t a_u \, dA_u^c = 0.$$

Dans la définition (5.5), la décomposition $A_u = A_u^c + A_u^d$ permet d'obtenir une décomposition $a_u = a_u^c + a_u^d$, et on a

$$\begin{aligned}
\mathbb{E} \int_s^t a_u dA_u^c &= \mathbb{E} \sum_i \mathbb{E}\left[A_{t_{i+1}} - A_{t_i} | \mathcal{F}_{t_i}\right] \left(A_{t_{i+1}}^c - A_{t_i}^c\right) \\
&= \mathbb{E} \sum_i \left(A_{t_{i+1}} - A_{t_i}\right) \mathbb{E}\left[A_{t_{i+1}}^c - A_{t_i}^c | \mathcal{F}_{t_i}\right] \\
&= \mathbb{E} \int_s^t a_u^c dA_u.
\end{aligned}$$

Or on montre comme au lemme 5.7 que la λ-intégrale de a^c tend vers 0. □

6. MARTINGALES ET FONCTIONS CONVEXES

Nous allons voir comment la convexité de la géométrie peut être utilisée pour approcher les martingales à temps continu par des martingales à temps discret; cette propriété sera fondamentale dans le paragraphe suivant pour la construction des martingales avec valeur finale fixée. Nous nous limiterons aux martingales dominées par la mesure de Lebesgue (pour le cas général il faudrait utiliser des subdivisions aléatoires, ce qui amène de nombreuses complications techniques). Nous commençons par étudier l'image d'une martingale par une fonction convexe.

PROPOSITION 6.1 – *Soit f une fonction $\mathbb{B}$-convexe lipschitzienne et soit X_t une martingale à valeurs dans V. Alors $f(X_t)$ est une sous-martingale.*

Remarque – Cette proposition peut servir de base à la définition des martingales continues; en fait, si V est suffisamment petite, un processus adapté continu X_t est une martingale si et seulement si $f(X_t)$ est une sous-martingale pour toute fonction convexe bornée f (proposition 4.41 de [11]). Dans le cas continu sur une variété sans bord, la condition de Lipschitz peut être omise, car les fonctions convexes sont localement lipschitziennes sur les cartes locales ([13]). Nous ne savons pas nous passer de cette hypothèse dans notre situation.

Démonstration – Nous pouvons nous ramener à une martingale de processus dominant A_t borné. Soit $[s, t]$ un sous-intervalle de $[0, 1]$; pour toute subdivision $\sigma = (t_i)$ de $[s, t]$, nous pouvons écrire

$$\begin{aligned} &f(X_t) - f(X_s) \\ &\quad = \sum_i \left(f(X_{t_{i+1}}) - f(\mathbb{B}[X_{t_{i+1}}|\mathcal{F}_{t_i}]) + f(\mathbb{B}[X_{t_{i+1}}|\mathcal{F}_{t_i}]) - f(X_{t_i})\right). \end{aligned}$$

D'après la convexité de f, l'espérance conditionnelle sachant $\mathcal{F}_{t_i}$ pour les deux premiers termes est positive ou nulle donc, en utilisant la propriété de Lipschitz de f pour les deux autres termes,

$$\mathbb{E}[f(X_t) - f(X_s)|\mathcal{F}_s] \geq -C\,\mathbb{E}\left[\sum_i \delta(\mathbb{B}[X_{t_{i+1}}|\mathcal{F}_{t_i}], X_{t_i}) \,\middle|\, \mathcal{F}_s\right].$$

En appliquant la proposition 5.10 et en passant à la limite quand $|\sigma|$ tend vers 0, nous obtenons l'existence d'un $C > 0$ tel que $f(X_t) + CA_t^d$ est une sous-martingale; en particulier $f(X_t)$ est une semimartingale, et si F_t^f désigne sa partie à variation finie prévisible, nous voulons montrer que F_t^f est croissante. Nous savons que

$$dF_t^f \geq -C dA_t^d$$

donc sa partie continue est croissante, et il est donc suffisant de montrer que $\Delta F_\tau^f \geq 0$ en tout temps τ prévisible; or

$$\Delta F_\tau^f = \mathbb{E}\left[f(X_\tau) - f(X_{\tau-})|\mathcal{F}_{\tau-}\right] = \mathbb{E}\left[f(X_\tau)|F_{\tau-}\right] - f\left(\mathbb{B}\left[X_\tau|\mathcal{F}_{\tau-}\right]\right)$$

est positif ou nul par convexité de f. □

Nous déduisons immédiatement de cette proposition une propriété appelée non confluence des martingales.

COROLLAIRE 6.2 – *Supposons que V est à géométrie p-convexe pour un $p \geq 1$; si X et Y sont deux martingales telles que $X_1 = Y_1$, alors $X_t = Y_t$ pour tout $t \leq 1$.*

Démonstration – Le processus (X_t, Y_t) est une martingale à valeurs dans $V \times V$ donc $\delta^p(X_t, Y_t)$ est une sous-martingale. □

Remarque – Dans le cas continu, cette propriété a été étudiée sous l'hypothèse plus faible de géométrie convexe. En fait, il est montré dans [20], [21] que la géométrie convexe de la variété est équivalente à une dépendance continue des martingales continues par rapport à leur valeur terminale.

THÉORÈME 6.3 – *Supposons que V est à géométrie p-convexe pour un $p \geq 1$, et soit X_t, $0 \leq t \leq 1$, une martingale de processus dominant $A_t = t$. Pour toute subdivision $\sigma = (t_i, 1 \leq i \leq n)$ de $[0, 1]$, soit X_t^σ le processus défini par*

$$X_1^\sigma = X_1, \qquad X_t^\sigma = \mathbb{B}\left[X_{t_{i+1}}^\sigma|\mathcal{F}_{t_i}\right] \quad si\ t_i \leq t < t_{i+1}.$$

Alors pour tout t, X_t^σ converge presque sûrement vers X_t; plus précisément,

$$\lim_{|\sigma|\to 0} \sup_i \delta\left(X_{t_i}^\sigma, X_{t_i}\right) = 0 \tag{6.1}$$

dans L^∞.

Il s'agit d'un résultat d'approximation d'une martingale à temps continu par une suite de martingales à temps discret. Nous commençons par montrer un résultat préliminaire.

LEMME 6.4. – *Soit X_t une martingale de processus dominant deterministe $a(t)$. Alors pour $s \leq t$,*

$$\left|\mathbb{E}\left[\gamma(X_s, X_t)|\mathcal{F}_s\right]\right| = o\left(a(t) - a(s)\right).$$

Démonstration du lemme 6.4. – En appliquant la formule de Ito, en prenant l'espérance conditionnelle comme en (5.7) et en appliquant la caractérisation (b') de la proposition 4.6, nous pouvons montrer que

$$\begin{aligned}&|\mathbb{E}\,[\gamma\,(X_s,\,X_t)|\mathcal{F}_s]|\\&\leq \mathbb{E}\left[C\int_s^t |X_{u-}-X_s|da\,(u)\right.\\&\left.+\sum_{s<u\leq t}\left|\int\gamma\,(x,\,y)\,\nu^d\,(\{u\},\,dx,\,dy)\right|\,\middle|\,\mathcal{F}_s\right].\end{aligned}$$

De plus d'après (1.2),

$$\left|\int\gamma\,(x,\,y)\,\nu^d\,(\{u\},\,dx,\,dy)\right| = o\,(\Delta a\,(u))$$

donc

$$|\mathbb{E}\,[\gamma\,(X_s,\,X_t)|\mathcal{F}_s]| \leq C\int_s^t \mathbb{E}\,[\,|X_{u-}-X_s||\mathcal{F}_s]\,da\,(u)+\sum_{s<u\leq t} o\,(\Delta a\,(u)).$$

Nous concluons en remarquant que

$$\mathbb{E}\,[|X_{u-}-X_s|^2|\mathcal{F}_s] \leq C\,(a\,(t)-a\,(s))$$

pour $s < u \leq t$. $\square$

Démonstration du théorème 6.3. – Pour σ fixé et pour chaque i, nous pouvons considérer la martingale X_j^i, $0\leq j\leq n$, relativement à la filtration $\mathcal{F}_{t_j}$, définie par

$$X_j^i = X_{t_i} \quad \text{pour} \quad j\geq i, \qquad X_j^i = \mathbb{B}\,[X_{j+1}^i|\mathcal{F}_{t_j}] \quad \text{pour} \quad j<i.$$

Nous voulons comparer X_j^n et X_j^j. Soit δ une distance telle que δ^p est convexe. La domination de X_t par la mesure de Lebesgue et le lemme 5.4 permettent de montrer que pour $s\leq t$ et $t-s\to 0$,

$$\delta\,(X_s,\,\mathbb{B}\,[X_t|\mathcal{F}_s]) \leq C|\mathbb{E}\,[\gamma\,(X_s,\,X_t)|\mathcal{F}_s]| + o\,(t-s).$$

D'après le lemme 6.4, le premier terme est négligeable devant $t-s$, donc

$$\delta\,(X_s,\,\mathbb{B}\,[X_t|\mathcal{F}_s]) = o\,(t-s)$$

dans L^∞, uniformément en s, t. Si on applique ce résultat pour $s = t_i$ et $t = t_{i+1}$, on a $X_{t_i} = X_i^i$ et le barycentre vaut X_i^{i+1} donc

$$\delta\,(X_i^i,\, X_i^{i+1}) = o\,(t_{i+1} - t_i)$$

dans L^∞. Comme $\delta^p\,(X_j^i,\, X_j^{i+1})$ est une sous-martingale,

$$\sup_{j \le i} \delta\,(X_j^i,\, X_j^{i+1}) = o\,(t_{i+1} - t_i).$$

En faisant la somme sur i de ces expressions, nous obtenons

$$\lim_{|\sigma| \to 0} \|\sup_j \delta\,(X_j^j,\, X_j^n)\|_\infty = 0.$$

Comme $X_j^j = X_{t_j}$ et $X_j^n = X_{t_j}^\sigma$, c'est exactement (6.1). Pour $0 \le t \le 1$, soit ζ le dernier instant de la subdivision σ inférieur à t (ce n'est pas un temps d'arrêt); le processus dominant est continu donc X est presque sûrement continu en t, ce qui implique que X_ζ converge vers X_t lorsque $|\sigma| \to 0$; de plus, (6.1) montre que $\delta\,(X_t^\sigma,\, X_\zeta)$ converge vers 0, donc X_t^σ converge vers X_t. $\square$

Le théorème 6.3 peut être appliqué à la convergence de schémas de discrétisation tels que (0.3). Plus précisément, soit $\mathcal{L}$ le générateur d'un processus de Feller Z_t. Soit $\mathcal{L}_V$ l'opérateur défini dans la proposition 4.8, et supposons que l'équation de la chaleur (0.1) a une solution h suffisamment régulière pour que $X_t = h(t, Z_t)$ soit une martingale de processus dominant $A_t = t$. Si nous considérons le schéma de discrétisation défini par

$$h^\sigma\,(1, \cdot) = u, \qquad h^\sigma\,(t_i,\, z) = \mathbb{B}\,[h^\sigma\,(t_{i+1},\, Z_{t_{i+1}}) | Z_{t_i} = z],$$

le processus $h^\sigma\,(t,\, Z_t)$ est exactement le processus X_t^σ du théorème 6.3, donc h^σ converge vers h. Comme $\mathcal{L}_V$ ne dépend que du connecteur, on peut choisir n'importe quel barycentre compatible avec le connecteur, pourvu que la variété reste à géométrie p-convexe. Si de plus la diffusion Z_t est continue, $\mathcal{L}_V$ ne dépend que de la connexion donc le choix pour $\mathbb{B}$ est encore plus large; par exemple sur une calotte sphérique, on peut choisir aussi bien le barycentre riemannien que le barycentre défini comme la projection orthogonale de la moyenne calculée dans l'espace euclidien.

7. CONSTRUCTION DES MARTINGALES

Nous avons vu que nous pouvons approcher les martingales au moyen d'un procédé de discrétisation ne faisant intervenir que la valeur finale de la martingale. Nous pouvons appliquer ce procédé à toute variable, sans savoir a priori s'il existe une martingale convergeant vers elle, et nous nous demandons si nous pouvons en déduire l'existence d'une telle martingale, l'unicité étant déjà démontrée dans le corollaire 6.2. Cette technique a déjà été utilisée dans le cas continu et pour le barycentre riemannien dans [19]. Dans les démonstrations précédentes, le processus dominant joue un rôle considérable; ici il faut savoir a priori que la martingale construite va être dominée par la mesure de Lebesgue. Cela va imposer des restrictions supplémentaires sans lesquelles nous ne pouvons démontrer l'existence. Dans le cas continu, on se limite en général à l'espace de Wiener ([19], [25]); ici, nous pourrons traiter le cas de processus à accroissements indépendants plus généraux. En nous plaçant sur une variété à géométrie p-convexe ($p \geq 1$ fixé), nous commençons par traiter le cas de variables finales régulières en un sens que nous allons préciser. Ensuite nous montrerons qu'une limite uniforme de martingales est une martingale (ce résultat est valable sans hypothèse de convexité sur la variété), ce qui nous permettra de généraliser notre résultat d'existence.

DÉFINITION 7.1. – *Soit Ω' une copie de Ω; à toute variable L sur Ω correspond la variable L' sur Ω'; nous avons également sur Ω' la filtration $\mathcal{F}'_t$ et nous munissons l'espace-produit $\Omega \times \Omega'$ de la filtration $\overline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'_t$. Nous pouvons identifier $\mathcal{F}_t$ et $\mathcal{F}'_t$ à des sous-tribus de $\overline{\mathcal{F}}_t$. Nous notons $\mathcal{V}$ l'ensemble des variables aléatoires L sur Ω à valeurs dans V qui sont $\mathcal{F}_1$ mesurables et vérifient la propriété suivante : pour tout $0 \leq s \leq t \leq 1$, il existe une probabilité $\mathbb{P}_t^s$ sur $\Omega \times \Omega'$ telle que*

(i) *la probabilité $\mathbb{P}_t^s$ est symétrique de marginales $\mathbb{P}$;*

(ii) *la restriction de $\mathbb{P}_t^s$ à $\overline{\mathcal{F}}_s$ est portée par la diagonale;*

(iii) *les tribus $\mathcal{F}_t$ et $\mathcal{F}'_t$ sont conditionnellement indépendantes sachant $\overline{\mathcal{F}}_s$;*

(iv) *pour tout u, les tribus $\mathcal{F}_1$ et $\mathcal{F}'_u$ sont conditionnellement indépendantes sachant $\mathcal{F}_u$;*

(v) *pour toute distance riemannienne δ on a*

$$\mathbb{E}_t^s\,[\mathbb{E}_t^s\,[\delta^p\,(L,\,L')|\overline{\mathcal{F}}_t]^{2/p}|\overline{\mathcal{F}}_s] \leq C\,(t-s)$$

pour un $C > 0$ ne dépendant pas de s et t.

Remarque – Les conditions (i), (ii) et (iii) déterminent $\mathbb{P}_t^s$ sur $\overline{\mathcal{F}}_t$; il faut donc choisir l'évolution après l'instant t de façon à ce que L et L' soient suffisamment proches.

Cette hypothèse sur L remplace l'hypothèse de dérivabilité sur l'espace de Wiener utilisée dans [25]. Donnons un exemple dans lequel elle peut être vérifiée.

PROPOSITON 7.2. – *Soit* $(G, \mathcal{G})$ *un espace lusinien, et soit* λ *une mesure* σ*-finie sur* G. *Supposons que* $\mathcal{F}_t$ *est la filtration engendrée par un processus à accroissements indépendants* (W, N), *où* W_t *est un processus de Wiener homogène et où* $(N_t(A);\ A \in \mathcal{G},\ \lambda(A) < \infty)$ *est un processus de Poisson ponctuel homogène d'intensité* λ. *Soit* L *une fonction lipschitzienne de la trajectoire* W *et d'un nombre fini de trajectoires* $N_t(A_i)$, $1 \leq i \leq k$, *avec* $\lambda(A_i) < \infty$; *alors* L *est dans* $\mathcal{V}$.

Démonstration – Soit Y_u le processus vectoriel à accroissements indépendants formé des coordonnées de W_u et des $N_u(A_i)$, $1 \leq i \leq k$. Nous désignons par Y'_u le processus similaire défini sur Ω', et nous prenons pour $\mathbb{P}_t^s$ la probabilité sous laquelle les accroissements de Y et Y' sont égaux entre 0 et s et entre t et 1, et sont indépendants entre s et t. Il n'est alors pas difficile de vérifier les conditions (i) à (iv) de la définition 7.1. D'autre part,

$$\begin{aligned}\delta(L, L') &\leq C \sup_{0 \leq u \leq 1} |Y_u - Y'_u| \\ &\leq C \sup_{s \leq u \leq t} (|Y_u - Y_s| + |Y'_u - Y'_s|),\end{aligned}$$

donc

$$\mathbb{E}_t^s [\mathbb{E}_t^s [\delta^p (L, L') | \overline{\mathcal{F}}_t]^{2/p} | \overline{\mathcal{F}}_s] \leq C\, \mathbb{E} \sup_{s \leq u \leq t} |Y_u - Y_s|^2.$$

Comme Y_u est de processus dominant u, nous obtenons la condition (v) de la définition 7.1. □

THÉORÈME 7.3. – *Si* V *est à géométrie* p*-convexe et si* $L \in \mathcal{V}$, *il existe une martingale sur* V *de valeur finale* L *et de processus dominant* t.

Ce théorème généralise les résultats de [19] et [25] obtenus sur l'espace de Wiener. En fait, dans [25], nous avons même obtenu un résultat d'existence sans faire d'hypothèse sur la convexité de la géométrie de V; plus précisément, si la variable L, considérée comme fonction du processus de Wiener, admet une constante de Lipschitz inférieure à une certaine valeur dépendant de la géométrie de V, alors il existe une martingale de valeur finale L; ce résultat est l'analogue probabiliste de l'existence locale de solutions à l'équation de la chaleur (0.1). Nous ne chercherons cependant pas ici à le généraliser au cas avec sauts. Nous allons tout d'abord exploiter la propriété $L \in \mathcal{V}$ dans un résultat préliminaire.

LEMME 7.4. – *Sous les hypothèses du théorème* 7.3, *soit* $\sigma = (t_i)$ *une subdivision de* $[0, 1]$ *et soit* $X_{t_i}^\sigma$ *la martingale à temps discret de valeur finale* L *définie par*

$$X_1^\sigma = L, \qquad X_{t_i}^\sigma = \mathbb{B}[X_{t_{i+1}}^\sigma | \mathcal{F}_{t_i}].$$

Alors, en plongeant V dans $\mathbb{R}^n$,

$$|\mathbb{E}\,[X^\sigma_{t_{i+k}} - X^\sigma_{t_i}|\mathcal{F}_{t_i}]| + \mathbb{E}\,[|X^\sigma_{t_{i+k}} - X^\sigma_{t_i}|^2|\mathcal{F}_{t_i}] \leq C\,(t_{i+k} - t_i).$$

Démonstration du lemme 7.4. – Appliquons l'hypothèse de la définition 7.1 pour $s = t_i$ et $t = t_{i+1}$. Sur $\Omega \times \Omega'$, considérons les variables

$$X_{t_j}\,(\omega,\,\omega') = X^\sigma_{t_j}\,(\omega), \qquad X'_{t_j}\,(\omega,\,\omega') = X^\sigma_{t_j}\,(\omega').$$

En appliquant la condition (iv), on vérifie que (X_{t_j}, X'_{t_j}) est la martingale discrète de valeur finale $(L,\,L')$; on en déduit que si δ^p est convexe,

$$\delta^p\,(X_{t_{i+1}},\,X'_{t_{i+1}}) \leq \mathbb{E}^{t_i}_{t_{i+1}}\,[\delta^p\,(L,\,L')|\overline{\mathcal{F}}_{t_{i+1}}]$$

donc par (v)

$$\mathbb{E}^{t_i}_{t_{i+1}}\,[\delta^2\,(X_{t_{i+1}},\,X'_{t_{i+1}})|\overline{\mathcal{F}}_{t_i}] \leq C\,(t_{i+1} - t_i).$$

Or d'après (iii), conditionnellement à $\overline{\mathcal{F}}_{t_i}$, $X_{t_{i+1}}$ et $X'_{t_{i+1}}$ sont deux copies indépendantes de la loi de $X^\sigma_{t_{i+1}}$ sachant $\mathcal{F}_{t_i}$, donc cette estimation signifie que la variance conditionnelle de $X^\sigma_{t_{i+1}}$ sachant $\mathcal{F}_{t_i}$ est dominée par $t_{i+1} - t_i$; d'après la proposition 2.1, la distance entre $X^\sigma_{t_i}$ et la moyenne conditionnelle de $X^\sigma_{t_{i+1}}$ est aussi dominée par $t_{i+1} - t_i$, donc l'estimation annoncée est satisfaite pour $k = 1$. Il est facile d'en déduire qu'elle est aussi satisfaite pour les autres valeurs de k et uniformément en i et k. $\square$

Démonstration du théorème 7.3. – Soit σ_n une suite de subdivisions croissant vers une partie dense D de [0, 1]. Pour $n > m$, $X^{\sigma_m}_t$ peut être vu comme une discrétisation de la martingale $X^{\sigma_n}_t$. Soit δ telle que δ^p est convexe. D'après la proposition 2.4 et le lemme 7.4, on vérifie que pour t_i et t_{i+k} dans σ_n,

$$\delta\,(X^{\sigma_n}_{t_i},\,\mathbb{B}\,[X^{\sigma_n}_{t_{i+k}}|\mathcal{F}_{t_i}]) \leq C|\mathbb{E}\,[\gamma(X^{\sigma_n}_{t_i},\,X^{\sigma_n}_{t_{i+k}})|\mathcal{F}_{t_i}]| + o\,(t_{i+k} - t_i)$$

si $t_{i+k} - t_i \to 0$. D'après le lemme 7.4, la martingale à temps discret $X^{\sigma_n}_{t_i}$ est dominée par $i \mapsto t_i$, donc en appliquant le lemme 6.4 on obtient

$$\mathbb{E}\,[\gamma\,(X^{\sigma_n}_{t_i},\,X^{\sigma_n}_{t_{i+k}})|\mathcal{F}_{t_i}] = o\,(t_{i+k} - t_i). \tag{7.1}$$

On a donc

$$\delta\,(X^{\sigma_n}_{t_i},\,\mathbb{B}\,[X^{\sigma_n}_{t_{i+k}}|\mathcal{F}_{t_i}]) = o\,(t_{i+k} - t_i)$$

dans L^∞. Si on applique cette estimation à un élément t_i de σ_m et à son successeur t_{i+k} dans σ_m, on obtient

$$\begin{aligned}\delta\,(X_{t_i}^{\sigma_n},\, X_{t_i}^{\sigma_m}) &= \delta\,(\mathbb{B}\,[X_{t_{i+k}}^{\sigma_n}|\mathcal{F}_{t_i}]),\, \mathbb{B}\,[X_{t_{i+k}}^{\sigma_m}|\mathcal{F}_{t_i}]) + o\,(t_{i+k} - t_i)\\ &\le \mathbb{E}\,[\delta^p\,(X_{t_{i+k}}^{\sigma_n},\, X_{t_{i+k}}^{\sigma_m})|\mathcal{F}_{t_i}]^{1/p} + o\,(t_{i+k} - t_i).\end{aligned}$$

En utilisant cette estimation à tous les instants t_i de σ_m, comme $X_1^{\sigma_n} = X_1^{\sigma_m}$, on obtient

$$\lim_{m\to\infty}\ \sup_{n>m}\ \sup_{t_i\in\sigma_m}\ \delta\,(X_{t_i}^{\sigma_n},\, X_{t_i}^{\sigma_m}) = 0$$

dans L^∞. Si nous considérons la suite de processus en escalier $X_t^{\sigma_n}$, nous en déduisons que pour tout t fixé dans D, la suite de variables $X_t^{\sigma_n}$ est de Cauchy dans L^∞; une application du lemme 7.4 montre que pour t hors de D la suite est encore de Cauchy dans L^2 et que si X_t désigne la limite, $t \mapsto X_t$ est continu en probabilité. On a immédiatement $X_1 = L$. D'autre part, le fait que les semimartingales à temps discret X^{σ_n} soient dominées par des processus uniformément bornés montre que la suite des lois est tendue pour la topologie des pseudo-trajectoires sur l'espace de Skorokhod (*voir* [23]); plus précisément pout tout $\alpha < \beta$ dans $\mathbb{R}$, l'espérance du nombre de montées de chacune des composantes de $X_t^{\sigma_n}$ à travers l'intervalle (α, β) est borné. A la limite, le nombre de montées de X_t, $t \in D$ est presque sûrement fini. Ceci implique que en tout instant t, X a presque sûrement des limites à droite et à gauche le long de D; comme X est continu en probabilité, en prenant le processus des limites à droite le long de D, on obtient une modification càdlàg. De plus, en passant à la limite dans le lemme 7.4, pour $s \le t$ dans D, on a

$$|\mathbb{E}\,[X_t - X_s|\mathcal{F}_s]| + \mathbb{E}\,[|X_t - X_s|^2|\mathcal{F}_s] = O\,(t-s).$$

Par continuité en probabilité, c'est aussi vrai en dehors de D, donc X_t est une quasi-martingale et par conséquent une semimartingale sur V de processus dominant t. De même en passant à la limite dans (7.1), on a

$$\mathbb{E}\,[\gamma\,(X_s,\, X_t)|\mathcal{F}_s] = o\,(t-s)$$

donc X_t est une γ-martingale (il suffit d'utiliser la définition 4.2); comme X_t est qcàg, c'est aussi une $\mathbb{B}$-martingale. □

THÉORÈME 7.5. – *Soit X_t^n une suite de martingales sur V et soit X_t un processus tel que*

$$\lim_n \sup_t \delta(X_t^n, X_t) = 0$$

en probabilité. Alors X_t est une martingale.

Ce résultat sera démontré au §8; expliquons d'abord comment il peut être appliqué à la construction des martingales sur les espaces de processus à accroissements indépendants.

COROLLAIRE 7.6. – *Supposons que V est à géométrie p-convexe. Soit $\overline{\mathcal{V}}$ l'adhérence de $\mathcal{V}$* (voir *définition 7.1*) *pour la topologie de la convergence en probabilité; alors pour tout $L \in \overline{\mathcal{V}}$, il existe une et une seule martingale de valeur finale L.*

Démonstration – Soit L^n une suite de variables de $\mathcal{V}$ convergeant en probabilité vers L. Soit X_t^n la martingale de valeur finale L^n construite au théorème 7.3; alors les processus $\delta^p(X_t^n, X_t^m)$ sont des sous-martingales, et $\delta(L^n, L^m)$ converge vers 0 et est borné, donc $\sup_t(X_t^n, X_t^m)$ converge vers 0 en probabilité. On en déduit que X_t^n converge uniformément en probabilité vers un processus X_t de valeur finale L, et qui est une martingale d'après le théorème 7.5. L'unicité a été vérifiée au corollaire 6.2. □

COROLLAIRE 7.7. – *Sous les hypothèses de la proposition 7.2 et si V est à géométrie p-convexe, toute variable $\mathcal{F}_1$ mesurable à valeurs dans V est la valeur finale d'une et d'une seule martingale.*

Démonstration – Soit L une variable $\mathcal{F}_1$ mesurable à valeurs dans V; d'après le corollaire 7.6, il suffit de montrer qu'elle peut être approchée par des variables de $\mathcal{V}$. Nous commençons par l'approcher par des variables ne dépendant que des valeurs de $(W_t, N_t(A))$ pour (t, A) dans une partie finie $\{(t_j, A_j)\}$ de $[0, 1] \times \mathcal{G}$; puis nous l'approchons par des variables ne prenant qu'un nombre fini de valeurs; une telle approximation peut être vue comme l'image des $(W_{t_j}, N_{t_j}(A_j))$ par une application mesurable à valeurs dans une partie finie F de V; en plongeant F dans l'ensemble $\mathcal{P}_F$ des probabilités sur F, nous pouvons voir cette application comme une application mesurable d'un espace vectoriel de dimension finie dans une partie convexe d'un espace affine de dimension finie; par régularisation, nous pouvons l'approcher par des application lipschitziennes à valeurs dans $\mathcal{P}_F$, et en projetant $\mathcal{P}_F$ sur V au moyen de $\mathbb{B}$, nous obtenons des approximations lipchitziennes en $(W_{t_j}, N_{t_j}(A_j))$ et à valeurs dans V, donc des éléments de $\mathcal{V}$ d'après la proposition 7.2. □

En particulier, en procédant comme pour l'espace de Wiener (voir [19] et [25]) on peut en déduire l'existence probabiliste de solutions à l'équation de la chaleur (0.1) ou au problème de Dirichlet (0.2) pour des variétés à géométrie p-convexe telles que celles des propositions 3.4 et 3.6. Plus précisément, cela est possible dès que nous pouvons réaliser le processus de Feller Z_t comme la solution forte d'une équation différentielle conduite par un processus de Wiener W_t et un processus de Poisson $(N_t(A))$ (*voir* [17]). Remarquons toutefois que nous n'obtenons pas une régularité autre que la mesurabilité pour ces solutions.

8. DÉMONSTRATION DU THÉORÈME 7.5

Dans le cas continu (théorème 4.43 de [11]), le résultat du théorème 7.5 se montre en recouvrant V par des ouverts sur lesquels la propriété de martingale peut se caractériser par les fonctions convexes; par localisation, on vérifie qu'on peut se limiter à des processus à valeurs dans un de ces ouverts et on passe à la limite sur la propriété de sous-martingale. Ici nous allons utiliser un procédé analogue, mais nous n'avons plus de caractérisation en termes de fonctions convexes et de sous-martingales; nous devons passer à la limite dans la définition 4.7; pour cela, nous devons estimer la variation quadratique des martingales à valeurs dans nos ouverts. Remarquons que pour le corollaire 7.6, il suffirait de montrer le théorème dans le cas de martingales quasi-continues à gauche, ce qui simplifie un peu la démonstration; cependant, nous préférons montrer le résultat général en raison de son intérêt propre.

LEMME 8.1. – *Plongeons V dans un espace euclidien $\mathbb{R}^n$. Pour tout point x de V, il existe un voisinage B de x et une constante $\alpha > 0$ telles que toute martingale sur V arrêtée à la sortie de B a une variation quadratique dans $\mathbb{R}^n$ d'espérance majorée par α.*

Remarque – Dire que la martingale est arrêtée à la sortie de B ne signifie pas qu'elle est à valeurs dans l'adhérence de B car il peut y avoir un saut à l'instant de sortie. Cela implique toutefois que la mesure $\nu(dt, dx, dy)$ définie au §4 ne charge que $[0, 1] \times B \times V$.

Démonstration – La lettre C désignera une constante strictement positive qui pourra varier de ligne à ligne; plusieurs constantes intervenant dans la même ligne seront différenciées par des indices. Commençons par supposer que x n'est pas sur le bord de V. Choisissons une fonction f régulière positive sur V s'annulant seulement au point x, et dont le hessien calculé pour la connexion associée à $\mathbb{B}$ soit non dégénéré en x; prenons $B = \{f < \beta\}$. En prenant β suffisamment petit, le hessien de f est uniformément elliptique sur B, et deux points de B sont reliés par une géodésique dans B. D'autre part, le hessien de la fonction $z \mapsto \gamma(x, z)$ est nul en $z = x$, donc quitte à diminuer β, le hessien de $z \mapsto \gamma(y, z)$ peut être majoré par une constante positive aussi petite que l'on veut pour y et z dans B; le hessien de

$$z \mapsto f(z) - f(y) - f'(y)\,\gamma(y, z)$$

peut donc être pris uniformément elliptique et dans ce cas, comme la dérivée est nulle en $z = y$ et comme y et z sont reliés par une géodésique dans B, on a

$$f(z) - f(y) - f'(y)\,\gamma(y, z) \geq C|z - y|^2$$

pour y et z dans B. L'existence d'un voisinage B et d'une fonction f satisfaisant cette propriété se généralise au cas où x est sur le bord de V en agrandissant un peu V. Si $y \in B$ et $z \in V$, nous pouvons écrire

$$f(z) - f(y) - f'(y)\,\gamma(y, z) \geq C_1|z - y|^2 - C_2\, 1_{\{z \notin B\}}. \tag{8.1}$$

D'autre part, si X est une variable à valeurs dans V dont le barycentre est dans B, en développant f au voisinage de $\mathbb{B}[X]$, nous obtenons

$$\begin{aligned}
&\mathbb{E} f(X) - f(\mathbb{B}[X]) \\
&\quad = f'(\mathbb{B}[X])\,\mathbb{E}\gamma(\mathbb{B}[X], X) \\
&\quad\quad + \frac{1}{2}\operatorname{Hess} f(\mathbb{B}[X])\,\mathbb{E}\,\langle \gamma(\mathbb{B}[X], X), \gamma(\mathbb{B}[X], X)\rangle \\
&\quad\quad + O(\mathbb{E}|\mathbb{B}[X] - X|^3).
\end{aligned} \tag{8.2}$$

Dans le membre de droite, nous allons majorer les valeurs absolues des permier et troisième terme et minorer le deuxième terme de façon à obtenir une minoration du membre de gauche. Le premier terme est $o(\operatorname{var} X)$; de plus

$$\operatorname{var} X \leq \varepsilon + C\,\mathbb{P}[X \notin B],$$

où ε peut être choisi aussi petit que l'on veut en diminuant β. Si on se fixe arbitrairement $\varepsilon' > 0$, alors $o(\operatorname{var} X)$ est majoré par $\varepsilon' \operatorname{var} X$ dès que $\operatorname{var} X$ est assez petit, donc, en choisissant ε suffisamment petit, dès que $\mathbb{P}[X \notin B]$ est assez petit. On en déduit qu'il existe $C > 0$ (dépendant de ε') tel que

$$|\mathbb{E}\gamma(\mathbb{B}[X], X)| \leq \varepsilon' \operatorname{var} X + C\,\mathbb{P}[X \notin B].$$

Le troisième terme de (8.2) admet une majoration du même type. Quant au deuxième, l'ellipticité de Hess f montre que

$$\begin{aligned}\mathbb{E}\,\langle\gamma\,(\mathbb{B}\,[X],\,X),\,\gamma\,(\mathbb{B}\,[X],\,X)\rangle &\geq C\,\mathbb{E}|\gamma\,(\mathbb{B}\,[X],\,X)|^2\\ &\geq C_1\,\mathrm{var}\,X - C_2\,\mathbb{P}\,[X\notin B].\end{aligned}$$

En choisissant β suffisamment petit, on peut donc s'arranger pour que

$$\mathbb{E}\,f\,(X) - f\,(\mathbb{B}\,[X]) \geq C_1\,\mathrm{var}\,X - C_2\,\mathbb{P}\,[X\notin B]. \qquad (8.3)$$

Cela signifie que f est strictement $\mathbb{B}$-convexe sur B. Soit maintenant X_t une martingale arrêtée à la sortie de B, et considérons la martingale locale $\overline{M}_t^f$ de (4.8) mise sous la forme

$$\begin{aligned}\overline{M}_t^f = f\,(X_t) &- \frac{1}{2}\int_s^t \mathrm{Hess}\,f\,(X_{s-})\,(dX_s^c,\,dX_s^c)\\ &- \int_0^t\int (f\,(z) - f\,(y) - f'\,(y)\,\gamma\,(y,\,z))\,\nu^c\,(ds,\,dy,\,dz)\\ &- \sum_{s\leq t}\int (f\,(z) - f\,(y))\,\nu^d\,(\{s\},dy,\,dz).\end{aligned}$$

L'intégrale par rapport à la variation quadratique continue s'estime en utilisant l'ellipticité du hessien de f. L'intégrale par rapport à ν^c s'estime à l'aide de (8.1). Quant à l'intégrale par rapport à ν^d, d'après (4.6) et (4.5), elle peut se mettre sous la forme

$$\sum_{\tau}(\mathbb{E}\,[f\,(X_\tau)|\mathcal{F}_{\tau-}] - f\,(\mathbb{B}\,[X_\tau|\mathcal{F}_{\tau-}]))$$

où la somme se fait sur une suite de temps énumérant les temps de saut prévisibles de X; d'après (8.3), cette expression est minorée par

$$\sum_{\tau}(C_1\,\mathrm{var}\,[X_\tau|\mathcal{F}_{\tau-}] - C_2\,\mathbb{P}\,[X_\tau\notin B|\mathcal{F}_{\tau-}]),$$

et nous minorons la variance conditionnelle de X_τ par la proposition 2.2. Toutes ces estimations permettent de déduire que

$$\begin{aligned}\overline{M}_t^f \leq f\,(X_t) - C_1\int_0^t |dX_s^c|^2 - C_1\int_0^t\int |z-y|^2\,\nu\,(ds,\,dy,\,dz)\\ + C_2\int_0^t\int 1_{\{z\notin B\}}\,\nu\,(dt,\,dy,\,dz).\end{aligned}$$

Par localisation nous pouvons supposer que $\overline{M}_t^f$ est une martingale et nous pouvons alors prendre l'espérance dans cette expression en $t = 1$; comme X_t est arrêté au temps de sortie τ de B, l'espérance du dernier terme vaut $\mathbb{P}\,[\tau \leq 1]$ et est donc majorée par 1; on en déduit que

$$\mathbb{E}\left[\int_0^1 |dX_t^c|^2 + \int_0^1 \int |z - y|^2 \,\nu\,(dt, dy,\, dz)\right] \leq C.$$

Or le terme apparaissant à l'intérieur du crochet est le compensateur prévisible de la variation quadratique de X. □

LEMME 8.2. – *Soit X_t une martingale sur V dont la variation quadratique Q_t est intégrable. Alors la variation sur* $[0, 1]$ *de X_t considéré comme quasi-martingale de $\mathbb{R}^n$ (au sens de* [4]) *est majorée par $C\,\mathbb{E}\,Q_1$ pour une constante C ne dépendant pas de X.*

Démonstration – En utilisant la condition (b') de la proposition 4.6, il faut majorer l'espérance de la variation de $F_t = X_t - \overline{M}_t$. Pour cela, il suffit de dominer la variation de chacune des intégrales de (4.7) par le compensateur prévisible de Q_t. C'est évident pour l'intégrale par rapport à la variation quadratique continue; pour l'intégrale par rapport à ν, il suffit de remarquer que

$$|\gamma\,(x,\, y) - (y - x)| \leq C|y - x|^2.$$

Quant à l'intégrale par rapport à ν^d, la condition (4.5) sur les temps de sauts prévisibles et (1.2) montrent que

$$\left|\int \gamma\,(x,\, y)\,\nu^d\,(\{\tau\},\, dx,\, dy)\right| \leq C \int |y - x|^2\,\nu^d\,(\{\tau\}, dx,\, dy),$$

ce qui est suffisant pour nos estimations. □

LEMME 8.3. – *Supposons que X_t est arrêté à la sortie d'un voisinage satisfaisant la condition du lemme* 8.1. *Alors la conclusion du théorème* 7.5 *est satisfaite.*

Démonstration – Remarquons tout d'abord que X_t est càdlàg comme limite uniforme de processus càdlàg. De plus les processus X_t^n arrêtés à la sortie de B convergent aussi uniformément vers X_t et sont des martingales, donc on supposera les X_t^n arrêtés. L'estimation du lemme 8.1 jointe au lemme 8.2 permet de dominer uniformément en n la variation de la quasi-martingale X_n. En passant à la limite, on en déduit que X_t est une quasi-martingale, donc une semimartingale. Pour montrer que c'est une martingale, nous voulons passer à la limite dans la caractérisation (b')

de la proposition 4.6, la condition (4.5) sur les sauts prévisibles étant elle facilement vérifiée. Comme ν est le compensateur prévisible de μ, il suffit en fait de vérifier que

$$\begin{aligned}\widetilde{M}_t = X_t + \int_0^t \rho(X_{s-})(dX_s, dX_s) \\ + \sum_{s \leq t} (\gamma(X_{s-}, X_s) - \Delta X_s - \rho(X_{s-})(\Delta X_s, \Delta X_s)) \\ - \sum_{\tau} \mathbb{E}[\gamma(X_{\tau-}, X_\tau)|\mathcal{F}_{\tau-}] 1_{\{\tau \leq t\}}\end{aligned} \tag{8.4}$$

est une martingale locale, où la dernière somme se fait sur une suite énumérant les temps prévisibles qui sont un temps de saut pour l'un des X^n. Par localisation, nous pouvons supposer que la variation quadratique de X et son compensateur prévisible (c'est-à-dire le crochet prévisible de X) sont bornés. La convergence de X^n vers X implique la convergence en probabilité de la variation quadratique de $X^n - X$ vers 0 (théorème VI.6.1 de [18] ou corollaire 2.2.5 de [24]); en particulier nous pouvons arrêter les processus X^n de façon à ce que leurs variations quadratiques soient uniformément bornées. Le crochet prévisible de $X^n - X$ converge également vers 0, donc nous pouvons aussi supposer que le crochet prévisible de X^n est uniformément borné. De plus quitte à prendre une sous-suite, nous pouvons supposer que la convergence vers 0 de $X^n - X$ et de sa variation quadratique est presque sûre. Sous ces hypothèses, il suffit de montrer que si $\widetilde{M}_t^n$ est la martingale bornée obtenue en ramplaçant X par X^n dans la définition (8.4) de $\widetilde{M}_t$, alors $\sup_t |\widetilde{M}_t^n - \widetilde{M}_t|$ converge presque sûrement vers 0. La convergence de l'intégrale est immédiate d'après la convergence de la variation quadratique de $X^n - X$; pour la première somme, nous remarquons que chacun des termes converge et est dominé par $|\Delta X_s^n|^3$; or

$$\sum |\Delta X_s^n|^3 1_{\{|\Delta X_s^n| < \varepsilon\}} \leq \varepsilon \sum |\Delta X_s^n|^2 \leq C\varepsilon$$

converge uniformément vers 0 quand $\varepsilon \to 0$, donc nous pouvons passer à la limite dans la somme; pour la dernière somme nous avons également la convergence de chacun des termes et

$$|\mathbb{E}[\gamma(X_{\tau-}^n, X_\tau^n)|\mathcal{F}_{\tau-}]| = o(\text{var}[X_\tau^n|\mathcal{F}_{\tau-}]) = o(\mathbb{E}[|\Delta X_\tau^n|^2 \mid \mathcal{F}_{\tau-}]),$$

la dernière expression étant le saut du crochet prévisible de X^n en τ; ces termes vérifient donc une estimation similaire à la précédente, et nous pouvons donc passer à la limite. □

Démonstration du théorème 7.5. – La variété V étant compacte, on peut la recouvrir par un nombre fini d'ouverts B_k, $1 \leq k \leq K$, satisfaisant à la condition du lemme 8.1. On prolonge cette suite par périodicité de façon à obtenir une suite B_k, $k \geq 1$. On pose $\tau_0 = 0$ et

$$\tau_k = \inf \{t \geq \tau_{k-1}; \quad X_t \notin B_k\}$$

pour $k \geq 1$. La suite τ_k est une suite de temps d'arrêt croissant de façon stationnaire vers 1. Pour tout $k \geq 0$, on peut appliquer le lemme 8.3 au processus $X_{t+\tau_k}$ considéré relativement à la filtration $\mathcal{F}_{t+\tau_k}$ et arrêté en $\tau_{k+1} - \tau_k$; on peut donc dire que X_t est une martingale sur l'intervalle de temps $[\tau_k, \tau_{k+1}]$; dans le cas des martingales réelles, une telle propriété implique que X_t est une martingale locale sur tout l'intervalle de temps [0, 1]; ici, on utilise la caractérisation (b') de la proposition 4.6 pour en déduire que $\overline{M}_t$ est une martingale locale sur [0, 1], donc X_t est une martingale (la condition (4.5) étant là aussi facilement vérifiée). □

RÉFÉRENCES

[1] K. C. Chang, W. Y. Ding et R. Ye, Finite-time blow-up of the heat flow of harmonic maps from surfaces, *J. Diff. Geometry*, Vol. **36**, 1992, pp. 507-515.

[2] S. Cohen, Géométrie différentielle avec sauts, *C. R. Acad. Sci. Paris, Série I*, Vol. **314**, 1992, pp. 767-770.

[3] R. W. R. Darling, Martingales in manifolds, definition, examples and behaviour under maps, dans : *Séminaires de Probabilités XVI, Supplément Géométrie différentielle stochastique*, *Lect. Notes Math.*, Vol. **921**, Springer, 1982.

[4] C. Dellacherie et P. A. Meyer, *Probabilités et potentiel, Chapitres V à VIII*, Hermann, 1980.

[5] C. Dellacherie et P. A. Meyer, *Probabilités et potentiel, Chapitres XII à XVI*, Hermann, 1987.

[6] S. Doss, Moyennes conditionnelles et martingales dans un espace métrique, *C. R. Acad. Sci. Paris, Série I*, Vol. **254**, 1962, pp. 3630-3632.

[7] T. E. Duncan, Stochastic integrals in Riemannian manifolds, *J. Multivariate Anal.*, Vol. **6**, 1976, pp. 397-413.

[8] J. Eells et L. Lemaire, A report on harmonic maps, *Bull. London Math. Soc.*, Vol. **10**, 1978, pp. 1-68.

[9] J. Eells et L. Lemaire, Another report on harmonic maps, *Bull. London Math. Soc.*, Vol. **20**, 1988, pp. 385-524.

[10] M. Emery, En cherchant une caractérisation variationnelle des martingales, dans : *Séminaire de Probabilités XXII*, Lect. N. Math., Vol. **1321**, Springer, 1988.

[11] M. Emery, *Stochastic calculus in manifolds*, Universitext, Springer, 1989.

[12] M. Emery et G. Mokobodzki, Sur le barycentre d'une probabilité dans une variété, dans : *Séminaire de Probabilités XXV*, *Lect. Notes Math.*, Vol. **1485**, Springer, 1991.

[13] M. Emery et W. A. Zheng, Fonctions convexes et semimartingales dans une variété, dans : *Séminaire de Probabilités XVIII*, *Lect. Notes Math.*, Vol. **1059**, Springer, 1984.

[14] A. Estrade, Exponentielle stochastique et intégrale multiplicative discontinues, *Ann. Inst. Henri Poincaré, Prob. Stat.*, Vol. **28**, 1992, 1, pp. 107-129.

[15] M. Fréchet, Les éléments aléatoires de nature quelconque dans un espace distancié, *Ann. Inst. Henri Poincaré*, Vol. **10**, 1948, pp. 215-310.

[16] W. Herer, Mathematical expectation and martingales of random subsets of a metric space, *Prob. and Math. Stat.*, Vol. **11**, 1991, 2, pp. 291-304.
[17] J. Jacod, Calcul stochastique et problèmes de martingales, *Lecture Notes Math.*, Vol. **714**, Springer, 1979.
[18] J. Jacod et A. N. Shiryaev, Limit theorems for stochastic processes, *Grundlehren der math. Wissenschaften*, Vol. **288**, Springer, 1987.
[19] W. S. Kendall, Probability, convexity and harmonic maps with small image I: uniqueness and fine existence, *Proc. London Math. Soc.*, Vol. **61**, 1990, 3, pp. 371-406.
[20] W. S. Kendall, Convex geometry and nonconfluent Γ-martingales I: tightness and strict convexity, dans : *Stochastic analysis (Durham 1990), London Math. Soc. Lecture Note Ser.*, Vol. **167**, Cambridge University Press, 1991.
[21] W. S. Kendall, Convex geometry and nonconfluent Γ-martingales II: well-posedness and Γ-martingale convergence, *Stochastics Stoch. Rep.*, Vol. **38**, 1992, pp. 135-147.
[22] P. A. Meyer, Géométrie stochastique sans larmes, dans : *Séminaire de Probabilités XV, Lecture Notes Math.*, Vol. **850**, Springer, 1981.
[23] P. A. Meyer et W. A. Zheng, Tightness criteria for laws of semimartingales, *Ann. Inst. Henri Poincaré, Prob. Stat.*, Vol. **20**, 1984, 4, pp. 353-372.
[24] J. Picard, Convergence in probability for perturbed stochastic integral equations, *Probab. Th. Rel. Fields*, Vol. **81**, 1989, pp. 383-452.
[25] J. Picard, Martingales on Riemannian manifolds with prescribed limit, *J. Functional Anal.*, Vol. **99**, 1991, pp. 223-261.
[26] J. Picard, Calcul stochastique avec sauts sur une variété, dans : *Séminaire de Probabilités XXV, Lecture Notes Math.*, Vol. **1485**, Springer, 1991.
[27] L. Schwartz, Géométrie différentielle du 2ème ordre, semimartingales et équations différentielles stochastiques sur une variété différentielle, dans : *Séminaire de Probabilités XVI, Supplément Géométrie différentielle stochastique, Lecture Notes Math.*, Vol. **921**, Springer, 1982.

(Manuscrit reçu le 21 janvier 1993;
corrigé le 23 juin 1993.)