

Astérisque

MIREILLE CHALEYAT-MAUREL
ROBERT AZENCOTT
PHILIPPE BOUGEROL
PAOLO BALDI
LAURE ELIE
JACQUES GRANARA
ANDRÉ BELLAICHE
CATHERINE BELLAICHE
MIREILLE CHALEYAT-MAUREL
Géodésiques et diffusions en temps petit

Astérisque, tome 84-85 (1981)

http://www.numdam.org/item?id=AST_1981__84-85__1_0

© Société mathématique de France, 1981, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (<http://smf4.emath.fr/Publications/Asterisque/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

géodésiques et diffusions en temps petit

**séminaire de probabilités
université de Paris VII**

Avec la participation de : R. AZENCOTT ; P. BALDI
A. BELLAICHE ; C. BELLAICHE ; P. BOUGEROL
M. CHALEYAT-MAUREL ; L. ELIE ; J. GRANARA

TABLE DES MATIÈRES

R. AZENCOTT : Introduction.	3
1 - M. CHALEYAT-MAUREL : Equations différentielles stochastiques et problème des martingales. Une introduction.	9
2 - R. AZENCOTT : Diffusions sur les variétés. Généralités.	17
3 - P. BOUGEROL : Géodésiques et connexions.	33
4 - P. BALDI : Premières majorations de la densité d'une diffusion sur $\mathbb{R}^m$. Méthode de la paramétrix.	43
5 - L. ELIE : Equivalent de la densité d'une diffusion en temps petit. Cas des points proches.	55
6 - J. GRANARA : Champs de Jacobi et dérivée de l'application exponen- tielle.	73
7 - A. BELLAICHE : Propriétés extrémales des géodésiques.	83
8 - R. AZENCOTT : Un problème posé par le passage des estimées locales aux estimées globales pour la densité d'une diffusion.	131
9 - C. BELLAICHE : Comportement asymptotique de $p(t,x,y)$ quand $t \rightarrow 0$ (points éloignés).	151
10 - M. CHALEYAT-MAUREL : La condition d'hypoellipticité d'Hörmander.	189
11 - M. CHALEYAT-MAUREL : Densité des diffusions invariantes sur certains groupes nilpotents. Calcul d'après B. Gaveau. ...	203
12 - R. AZENCOTT : Transformées de Cramer, diffusions en temps petit, et action minimale.	215
13 - R. AZENCOTT : Diffusions invariantes sur le groupe d'Heisenberg ; une étude de cas d'après B. Gaveau.	227
14 - L. ELIE : Approximation par des champs invariants.	237
15 - M. CHALEYAT-MAUREL ; L. ELIE : Diffusions gaussiennes.	255

INTRODUCTION

1. Le thème du séminaire :

Sur une variété différentielle M , de dimension m , considérons un opérateur différentiel L d'ordre 2, hypoelliptique, tel que $L1 = 0$. Soit $p(t, x, y)$, $t > 0$, $x, y \in M$ la solution fondamentale positive minimale de $(\frac{\partial}{\partial t} - L)$, associée à un volume riemannien arbitraire π sur M . Si $\mathcal{G}$ est l'unique processus de diffusion (sur M) de générateur infinitésimal L , alors p est la densité par rapport à π du semi-groupe de transition de $\mathcal{G}$.

La signification géométrique du comportement de $p(t, x, y)$ quand $t \rightarrow 0$, mise en évidence par Kac [4] et Mc Kean-Singer [5] a suscité une littérature fournie. Le but de ce séminaire était de faire le point sur l'approche probabiliste de ces questions, à partir des résultats de Varadhan [8], Molchanov [6], Gaveau [2][3]. Il s'agissait aussi de préciser l'impact sur ce problème des estimations à la Ventcell-Freidlin [9] et de leurs extensions (Azencott [1]).

Ce travail a permis de clarifier certaines obscurités non triviales de [2][6], d'améliorer plusieurs preuves, de mieux cerner les problèmes ouverts. La rédaction, volontairement extensive pour être d'accès plus aisé, présente séparément différentes techniques pertinentes (diffusions, paramétrix, opérateur hypoelliptiques, énergie, géodésiques, transformées de Cramer, etc...). Les nombreuses discussions entre les participants actifs au séminaire, un effort éditorial collectif, ont permis de donner cohérence à ce travail, bien que chaque auteur reste responsable de son texte.

Les auteurs remercient D. Meyer pour un efficace soutien moral, et Mme S. Force qui, avec l'aide de Mmes B. Chaouche et C. Pradier, a mené à bien un lourd travail dactylographique.

2. Structure de l'exposé :

Le chapitre 1, après quelques rappels sur les intégrales stochastiques browniennes et le problème des martingales, donne la formule de changement de drift de Cameron-Martin. Le chapitre 2 expose les théorèmes d'existence et unicité pour les diffusions sur une variété, les résultats de régularité sur leurs densités de transition, et une formule de localisation (à la Doob) liant la densité d'une diffusion à celle de la diffusion induite sur un ouvert. Le chapitre 3 résume les éléments de géométrie riemannienne : connexions, exponentielles, géodésiques.

Au chapitre 5, l'opérateur L est elliptique de classe 4, et on lui associe une distance riemannienne naturelle d sur M . Suivant Molchanov [6] on prouve que pour x, y fixés suffisamment proches dans M , on a

$$(1) \quad p(t, x, y) \sim t^{-\frac{m}{2}} F(x, y) \exp \left[- \frac{d^2(x, y)}{2t} \right]$$

où F se calcule explicitement à partir du jacobien de l'exponentielle et du travail du drift naturel de L le long de la géodésique joignant x à y . Une amélioration

du procédé de localisation évite le recours au résultat de Varadhan, qui pour x, y suffisamment proches prouve (sur $\mathbb{R}^m$)

$$(2) \quad \lim_{t \rightarrow 0} t \log p(t, x, y) = -\frac{1}{2} d^2(x, y).$$

L'interprétation géométrique de $F(x, y)$ dans (1) utilise une étude des champs de Jacobi et coordonnées semi-géodésiques, qui est présentée au chapitre 6. Un autre ingrédient de la preuve de (1) est fourni par le chapitre 4, où $M = \mathbb{R}^m$, L est uniformément elliptique à coefficients bornés (dépendant du temps), et où la méthode de la paramétrix donne des majorations classiques pour p et ses dérivées partielles, ainsi que, suivant [6], un équivalent de $p(s, t, x, x)$ quand $0 < s < t$, $t \rightarrow 0$.

Contrairement à une assertion trop générale de Molchanov [6], pour x, y non conjugués et joints par une unique géodésique minimisante, l'équivalent (1) et même son corollaire (2) peuvent être faux - avec L elliptique - lorsque M n'est pas complète pour la distance d . Au chapitre 8 on construit en effet des diffusions elliptiques telles que $\lim_{t \rightarrow 0} t \log p(t, x, y) > -\frac{1}{2} d^2(x, y)$, pour x, y dans un ouvert de $M \times M$. Une itération méticuleuse propage cependant l'équivalent local (1) en majorations globales de $p(t, x, y)$ du type $C(x, y) t^{-N(x, y)} \exp \left[-\frac{d^2(x, y)}{2t} \right]$, pourvu que (relation toujours vraie si M est complète)

$$(3) \quad d(x, y) < \max [d(x, \infty), d(y, \infty)]$$

Le chapitre 9 utilise (1) et ces majorations, pour montrer, toujours avec L elliptique de classe 4, que si (3) est vraie on a

$$(4) \quad p(t, x, y) \sim \int_{\Omega_{x, y}^{n, h}} \exp \left[-\frac{E(f)}{t} \right] H(f) d\pi_n(f)$$

où $\Omega_{x, y}^{n, h}$ est la variété $m+n$ dimensionnelle des géodésiques f brisées n fois joignant x à y en temps unité, et d'énergie $E(f) \leq \frac{1}{2} d^2(x, y) + h$, avec $n, h > 0$ convenables. Précisant avec soin la démarche de Molchanov [6], à l'aide des résultats du chapitre 7, on étudie l'intégrale de (4) pour $t \rightarrow 0$, suivant la nature des minima de E sur $\Omega_{x, y}^{n, h}$. Ainsi, pour x, y vérifiant (3), non conjugués et joints par une unique géodésique minimisante, l'équivalent (1) est vrai. Lorsque x, y sont conjugués (et vérifient toujours (3)), les équivalents du type

$$(5) \quad p(t, x, y) \sim G(x, y) t^{-\frac{m+\alpha}{2}} |\log t|^\beta \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

avec $\beta \in \mathbb{N}$, $\alpha \in \mathbb{Q}$, $0 \leq \alpha \leq m-1$, recouvrent tous les cas en classe analytique, ainsi que toutes les situations génériques en dimension inférieure à 10 en classe C^∞ , sans toutefois épuiser les possibilités en classe C^∞ .

INTRODUCTION

La relation (4) introduit la fonctionnelle énergie $E(f) = \frac{1}{2} \int_0^1 \|f'_t\|_{f_t}^2 dt$, où f est un chemin (absolument continu) sur la variété riemannienne M . Au chapitre 7 on considère l'espace $\Omega_{x,y}$ des chemins d'énergie finie joignant x à y comme variété modelée sur un espace de Hilbert, les points critiques de E sur $\Omega_{x,y}$ étant les géodésiques de M joignant x à y . Les notions de points conjugués, cut-locus, etc... s'interprètent par la nature des minima de E sur $\Omega_{x,y}$, et plus spécialement sur la sous-variété $\Omega_{x,y}^{n,h}$; pour les métriques C^∞ on montre que toutes les dégénérescences de E à priori possibles sont réalisées.

Lorsque L est elliptique, l'énergie $E(f)$ d'un chemin f non explosif coïncide avec $\lambda(f)$, où λ est la transformée de Cramer de f définie par

$$(6) \quad \lambda(f) = \frac{1}{2} \int_0^{1 \wedge \zeta} Q_{f(t)}^* [f'(t)] dt ;$$

ici $Q^* : TM \rightarrow [0, +\infty]$ est une forme quadratique naturellement déterminée par L , et ζ est le temps d'explosion de f . Le ch. 12 résume d'après [1] l'extension au cas hypoelliptique des estimations à la Ventcell-Freidlin

$$(7) \quad t \log P(X_{[0,t]} \in A) \sim - \inf_{f \in A} \lambda(f), \text{ pour } t \rightarrow 0,$$

où $X_{[0,t]}$ est la trajectoire sur $[0,t]$ d'une L -diffusion sur M , et A est un paquet de chemins $\gamma : [0,t] \rightarrow M \cup \{\infty\}$. Avec la définition

$$(8) \quad E(x,y) = \inf \{ \lambda(f) \mid f(0) = x, f(1) = y \}, \text{ où } x,y \in M \cup \{\infty\},$$

on passe de (7) à la conjecture

$$(9) \quad t \log p(t,x,y) \sim - E(x,y), \text{ pour } E(x,y) \text{ fini majoré par } E(x,\infty) ;$$

celle-ci est vraie dans le cas elliptique, où $E(x,y) = \frac{1}{2} d^2(x,y)$, cf. ch.9. Pour le cas hypoelliptique, des éléments de réponse sont apportés par les ch. 10 à 15.

Ayant rappelé (ch. 10) quelques aspects essentiels de l'hypoellipticité à la Hörmander, et de la théorie du potentiel associée (Bony), on considère (ch. 11) le groupe de Lie nilpotent libre G_r d'ordre 2, à r générateurs; on calcule selon Gaveau [2] la solution fondamentale de $(\frac{\partial}{\partial t} - L)$ avec $L = \frac{1}{2} \sum_{i=1}^r X_i^2$ où les champs X_i sont les générateurs de l'algèbre de Lie de G_r . L'étude du cas $r = 2$ (groupe d'Heisenberg de dimension 3) se poursuit au ch. 13, où un calcul de Gaveau [2] donne un équivalent explicite de $p(t,x,y)$ et identifie $J(x,y) = \lim_{t \rightarrow 0} [-t \log p(t,x,y)]$ avec

l'action de la bicaractéristique de L joignant x à y ; on prouve $J(x,y) = E(x,y)$, d'où (9) dans ce cas particulier, et on clarifie l'interprétation de J comme action minimale, interprétation esquissée trop succinctement dans Gaveau [2] [3].

La méthode de Molchanov [6] repose en partie sur les estimations du ch. 4 obtenues par la méthode de la paramétrix. Une éventuelle construction de paramétrix pour le cas hypoelliptique, basée sur l'approximation de L par des opérateurs invariants à gauche sur un groupe nilpotent, à la Rothschild-Stein [7], est esquissée avec pessimisme au ch. 14.

Enfin le ch. 15 étudie en détail une classe particulière de diffusions hypoelliptiques, les diffusions gaussiennes sur $\mathbb{R}^m$; les calculs explicites permettent de vérifier (9) dans ce cas, et surtout d'explorer les situations où l'hypoellipticité de L est due essentiellement à la présence d'un drift ; en effet on a alors $E(x,y) = +\infty$ pour presque tout $x,y \in \mathbb{R}^m$, et on trouve des équivalents du type

$$(10) \quad \log p(t,x,y) \sim - \frac{R(x,y)}{t^{r(x,y)}} ; \quad \log P(X_{[0,t]} \in A) \sim - \frac{C(A)}{t^{r(A)}}$$

avec $R(x,y)$, $C(A)$ finis, et $r(A), r(x,y)$ entiers impairs. Les résultats du ch. 15 suggèrent bon nombre de conjectures ; en particulier, avec les notations de (6) (7), on devrait avoir, pour $t \rightarrow 0$,

$$(11) \quad \log P(X_{[0,t]} \in A) \sim - \inf_{f \in A} \lambda^t(f)$$

$$(12) \quad \log p(t,x,y) \sim - \inf \{ \lambda^t(f) \mid f(0) = x, f(t) = y \}$$

si on pose

$$(13) \quad \lambda^t(f) = \frac{1}{2} \int_0^{t \wedge \zeta} Q_f^*(s) [f'(s) - Y_{f(s)}] ds$$

après avoir écrit $L = \frac{1}{2} \sum_i Z_i^2 + Y$.

Lorsque $(\frac{1}{2} \sum_i Z_i^2)$ est aussi hypoelliptique (à la Hörmander), il est plausible de tenter de déduire (11) de (7) et (12) de (9) ; par contre si ce n'est pas le cas, toujours avec L hypoelliptique ainsi que $(\frac{\partial}{\partial t} - L)$, une preuve de (11) et à fortiori de (12) semble pour le moment difficile à atteindre dans le cas "général" ; signalons des résultats (à paraître) de M. Chaleyat-Maurel et L. Elie qui abordent les situations "bien approchables" par des diffusions gaussiennes.

Robert Azencott

INTRODUCTION

BIBLIOGRAPHIE

- [1] R. Azencott, *Grandes déviations et applications*. Ecole d'été, Saint-Flour 1978, Lecture Notes Math. vol. 774 p.1-176. Springer Verlag 1980.
- [2] B. Gaveau, *Principe de moindre action. Propagation de la chaleur. Estimées sous elliptiques sur certains groupes nilpotents*. Acta Math. 139 (1977) p.96-153.
- [3] B. Gaveau, *Systèmes hamiltoniens associés à certains opérateurs hypoelliptiques*. Bull. Sci. Math. 102 (1978) p.203-229.
- [4] M. Kac - *Can one hear the shape of a drum*, Am. Monthly. 73 (1966), p. 1-23.
- [5] H. Mac Kean - J. Singer, *Curvature and eigenvalues of the Laplacien*, J. Diff. Geometry 1 (1967) p. 43-69.
- [6] S. Molchanov, *Diffusion processes and riemannian geometry*. Russ. Math. Surveys 30 (1975) p.1-63.
- [7] L. Rotschild - E. Stein, *Hypoelliptic differential operators and nilpotent groups*. Acta Math.137 (1976) p.248-315.
- [8] S. Varadhan, *Diffusion processes in a small time interval*. Com. Pure. Appl. Math. 20 (1967) p.659-685.
- [9] A. Ventsel - M. Freidlin, *Small random perturbations of dynamical systems*. Russ. Math. Surveys. 25 (1970) p.1-75.

CHAPITRE 1

UNE INTRODUCTION AUX EQUATIONS DIFFERENTIELLES STOCHASTIQUES ET AU PROBLÈME DES MARTINGALES

par

Mireille CHALEYAT-MAUREL

Le but de ce premier chapitre est de démontrer la formule de Cameron-Martin qui joue un rôle important dans l'existence et l'unicité des solutions du problème des martingales, surtout dans le cas non dégénéré ; et qui, dans le travail de S.A. Molchanov [5] sert à changer de coefficient de drift.

Auparavant, nous allons donner une brève présentation de l'intégrale stochastique par rapport au mouvement brownien dans un cadre élémentaire ; puis nous énoncerons un théorème d'existence et d'unicité pour les solutions d'équations différentielles stochastiques browniennes dans le cas où les coefficients dépendent du temps.

Nous rappellerons ensuite quelques propriétés du problème des martingales, pour terminer enfin par la démonstration de la formule de Cameron-Martin.

1. Définitions

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité complet muni d'une filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfaisant aux conditions habituelles de P.A. Meyer, c'est-à-dire :

- i) les sous-tribus $\mathcal{F}_t$ de $\mathcal{F}$ sont croissantes et continues à droite.
- ii) $\mathcal{F}_0$ contient les ensembles négligeables de $\mathcal{F}$.

Nous supposerons connues quelques notions élémentaires sur les martingales et pour ceci on pourra consulter le cours de M. Yor [8].

Cependant nous allons rappeler la notion de processus croissant associé à une martingale dont nous aurons besoin dans notre exposé.

Si $\mathfrak{M}$ désigne l'espace des martingales continues de carré intégrable et $\mathcal{A}$ l'espace des processus croissants continus, adaptés, intégrables, nuls en zéro, pour tout M de $\mathfrak{M}$, il existe un A unique appartenant à $\mathcal{A}$ tel que $M^2 - A$ soit une martingale ; ce résultat découle immédiatement de la décomposition de Doob-Meyer des surmartingales, et on note $A = \langle M, M \rangle$.

Par exemple, si M est un mouvement brownien réel β , alors $\langle \beta, \beta \rangle_t = t$ et le mouvement brownien réel est caractérisé par le fait que c'est la seule martingale continue de processus croissant t .

On définit également des classes plus générales de ces processus en les "localisant" : $\mathfrak{M}_{\text{loc}}$ et $\mathcal{A}_{\text{loc}}$ (localiser veut dire qu'il existe une suite croissante de temps d'arrêt $(T_n)_{n \in \mathbb{N}}$ tendant vers l'infini tels que les processus arrêtés en T_n vérifient la propriété voulue).

2. Intégrale stochastique

Nous allons donner une idée, dans le cadre L^2 , de la construction de l'intégrale stochastique par rapport au mouvement brownien.

Soit $(\beta_t)_{t \geq 0}$ un mouvement brownien réel, et soit $f : (s, \omega) \rightarrow f(s, \omega)$ une fonction continue, adaptée telle que :

$$E\left[\int_0^t f^2(s, \omega) ds\right] < \infty \text{ pour tout } t.$$

Nous allons donner un sens dans L^2 à $\int_0^t f(s, \omega) d\beta_s(\omega)$ bien que $\beta_s(\omega)$ ne soit pas à variation bornée.

2.1. PROPOSITION. Soit $\sigma^n = \{0 \leq t_0^n < t_1^n < \dots < t_n^n \leq t\}$ une subdivision de $[0, t]$, alors

$\sum_{\sigma^n} f(t_i^n) (\beta_{t_{i+1}^n} - \beta_{t_i^n})$ converge dans $L^2(\Omega, \mathcal{F}, P)$ vers une limite lorsque le pas de σ^n tend vers zéro.

Pour démontrer cette proposition, on utilise essentiellement la propriété caractéristique du mouvement brownien : β_t et $\beta_t^2 - t$ sont deux martingales ; on montre alors que $\sum(\sigma^n)$ converge dans L^2 vers une limite qu'on appelle $(f \cdot \beta)_t$, intégrale stochastique de f par rapport à β . $(f \cdot \beta)$ est une martingale qui vérifie l'égalité isométrique :

$$E(f \cdot \beta)_t^2 = E\left[\int_0^t f^2(s) ds\right]$$

et on a : $\langle f \cdot \beta, f \cdot \beta \rangle_t = \int_0^t f^2(s) ds$. (Pour un exposé complet voir [4])

3. Equations différentielles stochastiques

L'intégrale stochastique définie au paragraphe 2 permet de donner un sens à des équations différentielles stochastiques du type :

$$dX_t = \sigma(t, X_t) d\beta_t + b(t, X_t) dt,$$

dont la résolution permet d'obtenir l'existence et l'unicité de diffusions dont le générateur différentiel est donné. La proposition 3.1 explicitera exactement cette relation.

Rappelons tout d'abord que si δ désigne le point à l'infini du compactifié d'Alexandrov de $\mathbb{R}^m$ (ou d'un ouvert U de $\mathbb{R}^m$), on appelle temps de mort d'une fonction $\gamma : [0, \infty[\rightarrow \mathbb{R}^m \cup \delta$ (ou $U \cup \delta$), le temps d'atteinte de δ défini par $\zeta(\gamma) = \inf\{t \geq 0 : \gamma(t) = \delta\}$.

3.1. PROPOSITION. Soient respectivement σ et b des applications de $\mathbb{R}^+ \times U$ dans l'espace des matrices (m, k) et dans $\mathbb{R}^m$ vérifiant les conditions suivantes :

i) pour tout x de U , $\sigma(\cdot, x)$ et $b(\cdot, x)$ sont mesurables et localement bornées.

ii) pour tout entier n , pour tout t de $\mathbb{R}^+$, il existe des fonctions $c^n : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, bornées sur tout compact et telles que si $\|x\| \leq n$, $\|y\| \leq n$, on ait :

$$\|\sigma(t, x) - \sigma(t, y)\| \leq c^n(t) \|x - y\|$$

$$\text{et } \|b(t, x) - b(t, y)\| \leq c^n(t) \|x - y\|.$$

Et soit $\beta_t : \Omega \rightarrow \mathbb{R}^k$ un mouvement brownien. Considérons sur U l'équation différentielle stochastique :

$$(1) \quad dX_t = \sigma(t, X_t) d\beta_t + b(t, X_t) dt,$$

et notons ζ le temps de mort de X_t . Alors pour chaque (s, x) de $\mathbb{R}^+ \times U$, il existe une solution X_t de (1) telle que $X_s = x$, solution qui est unique sur $[s, \zeta[$. La collection des solutions de (1) issues des différents points de $\mathbb{R}^+ \times U$ définit une diffusion sur U de générateur différentiel :

$$L_t = \frac{1}{2} \sum_{1 \leq i, j \leq n} a_{ij}(t, \cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq n} b_i(t, \cdot) \frac{\partial}{\partial x_i}.$$

$$\text{où } [a_{ij}(t, x)] = \sigma(t, x) \sigma^*(t, x), \quad b(t, x) = (b_1(t, x), \dots, b_n(t, x)).$$

Preuve : Renvoyons à H. Doss et E. Lenglart [2] pour l'existence et l'unicité des solutions de (1), à D. Stroock et S. Varadhan [7] et A. Friedman [3] pour le reste.

Une fois prouvées l'existence et l'unicité des solutions de (1), la propriété de Markov forte du processus qu'elles définissent se déduit de l'unicité des solutions

de (1) et de la possibilité de les recoller à un temps d'arrêt.

Pour trouver le générateur différentiel on utilise la formule de Ito et le problème des martingales que nous allons rappeler ci-dessous.

4. Problème des martingales

Soit $(\Omega, \mathcal{F}, \mathcal{F}_t)$ l'espace canonique usuel où Ω est l'espace des fonctions continues de $\mathbb{R}^+$ dans $\mathbb{R}^m$: $\mathcal{C}(\mathbb{R}^+, \mathbb{R}^m)$, $X_t(\omega) = \omega_t$, $\mathcal{F}_t = \sigma(X_s, s \leq t)$, $\mathcal{F} = \bigvee_t \mathcal{F}_t$ avec les tribus complétées et rendues continues à droite.

Soient $a : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathcal{S}_m$ et $b : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ deux fonctions bornées, mesurables avec pour tout (t, x) de $\mathbb{R}^+ \times \mathbb{R}^m$, $a(t, x) = \{a_{ij}(t, x)\}$ et $b(t, x) = \{b_i(t, x)\}$; où $\mathcal{S}_m$ désigne l'ensemble des matrices symétriques positives, c'est-à-dire

$$\sum_{1 \leq i, j \leq m} a_{ij} \theta_i \theta_j \geq 0 \text{ pour tout } \theta \text{ de } \mathbb{R}^m.$$

On définit pour tout t positif,

$$L_t = \frac{1}{2} \sum_{1 \leq i, j \leq m} a_{ij}(t, \cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq m} b_i(t, \cdot) \frac{\partial}{\partial x_i}.$$

Etant donné (s, x) appartenant à $\mathbb{R}^+ \times \mathbb{R}^m$, une solution du problème des martingales pour L_t partant de (s, x) (en abrégé PBM (s, x, a, b)) est une probabilité P sur $(\Omega, \mathcal{F})$ telle que :

- i) $P[X_t = x, 0 \leq t \leq s] = 1$
- ii) $f(X_t) - f(X_s) - \int_s^t L_u f(X_u) du$ est une P -martingale pour toute f de $C_b^2(\mathbb{R}^m)$.

Cette formule en termes de martingale est due à D. Stroock et S. Varadhan ; pour plus de détails on pourra consulter leur récent livre [7].

Nous allons rappeler un théorème d'équivalence qui nous servira dans la formule de Cameron-Martin.

4.1. THÉOREME. Soit (s, x) appartenant à $\mathbb{R}^+ \times \mathbb{R}^m$ et P une probabilité sur $(\Omega, \mathcal{F})$ telle que $P[X_t = x, 0 \leq t \leq s] = 1$; il y a équivalence entre :

i) P résout le PBM (s, x, a, b) .

ii) pour tout θ de $\mathbb{R}^m$,

$$\exp[\langle \theta, X_t - X_s \rangle - \int_s^t b(u, X_u) du - \frac{1}{2} \int_s^t \langle \theta, a(u, X_u) \theta \rangle du]$$

est une P -martingale.

Donnons quelques indications sur la démonstration :

i) $\implies$ ii)

Il y a deux étapes : d'abord on montre que i) entraîne que :

pour tout θ de $\mathbb{R}^m$, $M_t^\theta = \langle \theta, X_t - X_s - \int_s^t b(u, X_u) du \rangle$ est une P-martingale de processus croissant : $\int_s^t \langle \theta, a(u, X_u) \theta \rangle du$.

En effet, on passe des fonctions de C_b^2 à celles de C^2 par approximation, on prend $f(x) = \langle \theta, x \rangle$ et on montre que les martingales locales qui interviennent sont de vraies martingales car les coefficients sont bornés.

La formulation ii) est alors la conséquence de la propriété suivante :

Soient A appartenant à $\mathcal{A}_{loc}$ et M un processus continu tel que $M_0 = 0$, il y a équivalence entre :

a) $M \in \mathcal{M}_{loc}$ et $\langle M, M \rangle = A$.

b) pour tout λ de $\mathbb{R}$, $\exp(\lambda M - \frac{\lambda^2}{2} A) \in \mathcal{M}_{loc}$.

De plus,

. $E \int_0^t \exp(2\lambda M_s) dA_s < \infty \implies \exp(\lambda M - \frac{\lambda^2}{2} A) \in \mathcal{M}$.

. Si, pour tout λ et t , $\exp(\lambda M - \frac{\lambda^2}{2} A) \in \mathcal{M}$ et $E(\exp \lambda M_t) < \infty$, alors M appartient à $\mathcal{M}$.

(Pour les détails de ce théorème et la démonstration du lemme, voir le cours de P. Priouret [6]).

ii) $\implies$ i)

On utilise un théorème de classe monotone pour obtenir un résultat sur les fonctions de C_b^2 à partir de la propriété sur les exponentielles.

Remarquons que si P est solution du PBM(s,x,a,b), les intégrales stochastiques par rapport à X de processus progressivement mesurables ont bien un sens car alors X est la somme d'une martingale et d'un processus à variation finie.

(On rappelle qu'une fonction f de $\mathbb{R}^+ \times \Omega$ est progressivement mesurable si pour tout t de $\mathbb{R}^+$, la restriction de f à $[0, t] \times \Omega$ est $\mathcal{B}_{[0, t]} \otimes \mathcal{F}_t$ - mesurable, où $\mathcal{B}_{[0, t]}$ désigne la tribu borélienne de $[0, t]$.)

On peut alors énoncer :

4.2. PROPOSITION. Soit P solution du PBM(s,x,a,b), alors pour tout h(u) processus à valeurs $\mathbb{R}^m$, progressivement mesurable, borné

$$\exp\left[\int_s^t \langle h(u), dX_u \rangle - \int_s^t \langle h(u), b(u, X_u) \rangle du - \frac{1}{2} \int_s^t \langle h(u), a(u, X_u) h(u) \rangle du\right]$$

est une P-martingale.

Pour démontrer ceci, on approxime h par des fonctions simples et on applique l'équivalence ii) du théorème 4.1 (voir le livre de D. Stroock et S. Varadhan [7] p. 100).

5. Formule de Cameron-Martin

Avant d'énoncer le résultat, rappelons un résultat d'existence de probabilités.

Soit $(\Omega, \mathcal{F}, \mathcal{F}_t)$ l'espace canonique considéré au paragraphe 4 et Z une $(\mathcal{F}_t, P)$ martingale telle que Z_t soit positive et $E(Z_t) = 1$; alors la formule

$$Q_t(A) = \int_A Z_t dP \text{ pour } A \text{ appartenant à } \mathcal{F}_t,$$

définit une probabilité Q_t sur $\mathcal{F}_t$ qui peut se prolonger en une probabilité Q sur $\mathcal{F}$ telle que $Q/\mathcal{F}_t = Q_t$. En effet, la propriété de martingale de Z prouve que la restriction de Q_t à $\mathcal{F}_u$ pour u inférieure à t est Q_u ; l'existence de Q sur $\mathcal{F}$ résulte alors d'un théorème de C. Dellacherie et P.A. Meyer [1] (p. 112) sur les limites projectives de mesures régulières ; la régularité provenant du fait que l'espace canonique est polonais.

Pour la démonstration de la formule de Cameron-Martin nous suivons de près le cours de P. Priouret [6] qui traite le cas homogène.

5.1. THÉOREME. Soient $a : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathcal{J}_m$, $b : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, $c : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ trois fonctions mesurables bornées et P une solution du PBM(s,x,a,b).

On définit :

$$R_t = \exp \left[\int_s^t \langle c(u, X_u), dX_u \rangle - \int_s^t \langle c(u, X_u), b(u, X_u) \rangle du \right. \\ \left. - \frac{1}{2} \int_s^t \langle ac(u, X_u), c(u, X_u) \rangle du \right]$$

Soit Q la probabilité définie par $\frac{dQ}{dP}/\mathcal{F}_t = R_t$, alors Q est solution du PBM(s,x,a,b+ac).

Preuve : La proposition 4.2 assure que R est une martingale d'espérance 1, donc l'existence de Q est garantie par la remarque ci-dessus.

Pour montrer que Q est solution du PBM(s,x,a,b+ac), il suffit de vérifier d'après le théorème 4.1 que,

pour tout θ de $\mathbb{R}^m$,

$$X_t^\theta = \exp \left[\langle \theta, X_t - X_s \rangle - \int_s^t \langle \theta, b(u, X_u) \rangle du - \frac{1}{2} \int_s^t \langle \theta, (b+ac)(u, X_u) \rangle du \right]$$

est une Q -martingale.

Or P résout le PBM(s,x,a,b) ; donc en appliquant le résultat de la proposition 4.2 avec $h(u) = \theta + c(u, X_u)$, on a :

$$M_t^\theta = \exp \left[\int_s^t \langle \theta + c(u, X_u), dX_u \rangle - \int_s^t \langle \theta + c(u, X_u), b(u, X_u) \rangle du \right. \\ \left. - \frac{1}{2} \int_s^t \langle \theta + c(u, X_u), a(u, X_u)(\theta + c(u, X_u)) \rangle du \right]$$

est une P-martingale, mais elle vaut :

$$M_t^\theta = X_t^\theta R_t^\theta ; \text{ donc, pour tout A de } \mathcal{F}_v \text{ avec } s \leq v \leq t,$$

$$\int_A X_t^\theta dQ = \int_A X_t^\theta R_t^\theta dP = \int_A M_t^\theta dP = \int_A M_v^\theta dP = \int_A X_v^\theta dQ.$$

BIBLIOGRAPHIE

- [1] C. Dellacherie et P.A. Meyer, *Probabilités et potentiel*, Hermann (1975).
- [2] H. Doss et E. Lenglart, *Sur l'existence, l'unicité et le comportement asymptotique des solutions d'équations différentielles stochastiques*, Ann. Inst. Poincaré, vol. XIV, n°2 (1978).
- [3] A. Friedman, *Stochastic differential equations and applications*, vol 1, Academic Press (1975).
- [4] H.P. Mc Kean, *Stochastic integrals*, Academic Press (1969).
- [5] S.A. Molchanov, *Diffusion processes and riemannian geometry*, Russian Math. Survey, 30, 1 (1975) p.1-63.
- [6] P. Priouret, *Cours de l'Ecole d'Eté de St Flour*, Springer (1973).
- [7] D. Stroock et S. Varadhan, *Multidimensional diffusion processes*, Springer (1979).
- [8] M. Yor, *Cours de l'école d'Eté de St Flour*, 1979, (à paraître).

CHAPITRE 2

DIFFUSIONS SUR LES VARIÉTÉS GÉNÉRALITÉS

par

Robert AZENCOTT

0. Introduction

Nous esquissons dans ce chapitre quelques définitions équivalentes des diffusions sur les variétés différentiables, les principaux théorèmes d'existence et unicité, les théorèmes d'existence de densités régulières pour ces diffusions. Nous donnons une formule de localisation, généralisant une formule de Doob, qui relie la densité d'une diffusion sur M à celle de la diffusion induite sur U . Le cadre adopté englobe celui des générateurs différentiels L elliptiques (à coefficients continus, ou localement hölderiens suivant les résultats cherchés) et celui des générateurs différentiels L hypoelliptiques à coefficients C_∞ .

Le point de vue adopté insiste sur la bijection entre générateur différentiel L et classe d'équivalence de L -diffusions, sur une variété donnée. Tout l'aspect "équations différentielles stochastiques", à base de semimartingales sur une variété, a été laissé de côté pour minimiser le vocabulaire probabiliste à utiliser. Renvoyons à L. Schwartz [16], à Bismut [Cours St Flour 1980], et à P. Malliavin [13], où d'autres points de vue sont largement développés.

Nous nous sommes bornés au cas de générateurs différentiels indépendants du temps pour simplifier l'exposé, bien que la plupart des résultats énoncés s'étendent au cas de diffusions non homogènes dans le temps.

1. Processus de Markov fort :

Soit M une variété différentiable connexe, de dimension m , de classe au moins 2. On note δ le point à l'infini du compactifié d'Alexandroff $M \cup \delta$ de M (δ est isolé si M est compacte). Toute fonction numérique f sur M sera implicitement prolongée à $M \cup \delta$ en posant $f(\delta) = 0$.

Pour toute fonction $\gamma : [0, +\infty[\longrightarrow M \cup \delta$ appelons temps de mort de γ le temps d'atteinte de δ défini par

$$\zeta(\gamma) = \inf \{ t \geq 0 \mid \gamma(t) = \delta \}$$

Appelons $\mathcal{D}(M)$ l'ensemble des fonctions $\gamma : [0, +\infty[\longrightarrow M \cup \delta$ qui sont continues sur $[0, \zeta(\gamma)[$, et qui, pour $\zeta(\gamma)$ fini, admettent une limite à gauche en $\zeta(\gamma)$ et sont constantes égales à δ sur $[\zeta(\gamma), +\infty[$. Notons $X_t = \mathcal{D}(M) \longrightarrow M \cup \delta$ les applications coordonnées, $\mathcal{F}_t$ la σ -algèbre de parties de $\mathcal{D}(M)$ engendrée par les X_s , $0 \leq s \leq t$, et $\mathcal{F}_\infty$ la σ -algèbre engendrée par tous les X_t , $t \geq 0$.

Tous les processus considérés ici seront des processus de Markov forts (cf. Meyer [14]) sur M , à trajectoires presque surement dans $\mathcal{D}(M)$. Soient $\mathcal{P}$, $\mathcal{P}'$ deux tels processus ; notons P_x , P'_x les lois des trajectoires de $\mathcal{P}$, $\mathcal{P}'$ issues de $x \in M$ au temps 0. On dit que $\mathcal{P}$ et $\mathcal{P}'$ sont équivalents si pour tout $x \in M$ les probabilités P_x et P'_x coïncident sur $(\mathcal{D}(M), \mathcal{F}_\infty)$. Un classique rituel de complétient des tribus $\mathcal{F}_t$ fait de $\{\mathcal{D}(M), (\mathcal{F}_t), (X_t)(P_x)\}$ un processus équivalent à $\mathcal{P}$.

Le semi-groupe de transition (R_t) de $\mathcal{P}$ est défini par $R_t f(x) = \int_{\mathcal{D}(M)} f(X_t) dP_x$, f fonction borélienne bornée sur M . Une liaison faible entre la structure différentiable de M et le semi-groupe (R_t) va permettre de préciser la forme du générateur infinitésimal de (R_t) .

2. Quelques définitions des diffusions

2.1. Opérateurs différentiels : Nous allons considérer, sur la variété M , des opérateurs différentiels L d'ordre 2, semi-elliptiques, à coefficients mesurables et localement bornés, vérifiant $L1 \leq 0$. Pour toute carte locale $\phi : U \longrightarrow \mathbb{R}^n$, U ouvert de M , l'image L_ϕ de L par ϕ s'écrit donc

$$L_\phi = \frac{1}{2} \sum_{1 \leq i, j \leq n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq n} b_i(x) \frac{\partial}{\partial x_i} + c(x)$$

où :

- pour chaque $x \in \phi(U)$, la matrice $a(x) = [a_{ij}(x)]$ est positive au sens large, et $c(x) \leq 0$

- les coefficients a_{ij} , b_i , c sont mesurables et localement bornés sur $\phi(U)$.

Cette famille d'opérateurs s'introduit naturellement comme suit.

2.2. THÉORÈME. Soit M une variété et soit $C_k^2(M)$ l'espace des fonctions de classe 2 à support compact sur M . Soit $\mathcal{P}$ un processus de Markov fort à trajectoires p.s. dans $C(M)$, de semi-groupe de transition (R_t) . Supposons que pour chaque $f \in C_k^2(M)$ les fonctions $g_t = \frac{1}{t}(R_t f - f)$ restent uniformément bornées et convergent simplement sur M lorsque $t \rightarrow 0$. Alors il existe un opérateur différentiel L vérifiant 2.1 et tel que

$$(1) \quad \lim_{t \rightarrow 0} \frac{1}{t} [R_t f(x) - f(x)] = Lf(x) \text{ pour tout } x \in M, f \in C_k^2(M)$$

De plus on a alors

$$R_t f(x) - f(x) = \int_0^t R_s Lf(x) ds \text{ pour tout } x \in M, f \in C_k^2(M), t \geq 0$$

Preuve : Renvoyons à Azencott [3] pour une (longue) démonstration. Un coup d'oeil à Bony-Courrège-Priouret [6] permettra d'apprécier le rôle essentiel du principe du maximum.

2.3. COROLLAIRE. Soit $\mathcal{E}(M)$ l'ensemble des trajectoires $\gamma \in \mathcal{D}(M)$ qui sont continues sur $[0, +\infty[$. Si un processus $\mathcal{P}$ vérifie l'hypothèse du théorème 2.2, et si de plus ses trajectoires sont p.s. dans $\mathcal{E}(M)$, alors l'opérateur L du théorème 2.2 vérifie $L1 \equiv 0$.

Preuve : Renvoyons à [3]. Notons en passant que sur $\mathcal{E}(M)$, le temps de mort ζ est en fait un "temps d'explosion" car $\{\gamma \in \mathcal{E}(M), \zeta(\gamma) \text{ fini}\}$ implique $\lim_{t \nearrow \zeta(\gamma)} \gamma(t) = \delta$.

2.4. DÉFINITION. Soit M une variété. Appelons *diffusion* sur M tout processus de Markov fort sur M , à trajectoires p.s. dans $\mathcal{D}(M)$, dont le semi-groupe de transition (R_t) vérifie

$$(2) \quad R_t f(x) - f(x) = \int_0^t R_s Lf(x) ds \text{ pour tout } x \in M, f \in C_k^2(M), t \geq 0 \text{ où } L \text{ est un opérateur différentiel du type 2.1.}$$

L'opérateur L est appelé *générateur différentiel* de $\mathcal{P}$. Nous dirons aussi que $\mathcal{P}$ est une *L-diffusion* sur M . Clairement tout processus équivalent à une L-diffusion est encore une L-diffusion.

Remarquons que, si L est à coefficients continus, et si $\mathcal{P}$ est un processus de Markov fort à trajectoires p.s. dans $\mathcal{D}(M)$, la condition (2) équivaut à imposer que lorsque $t \rightarrow 0$, les fonctions $\frac{1}{t}(R_t f - f)$ restent uniformément bornées et convergent simplement vers Lf sur M , pour chaque $f \in C_k^2(M)$.

2.5. PROPOSITION. Soit $\mathcal{P}$ un processus de Markov fort à trajectoires p.s. dans $\mathcal{D}(M)$, et soient $(P_x)_{x \in M}$ les lois de ses trajectoires. Soit L un opérateur différentiel sur M vérifiant 2.1. Pour que $\mathcal{P}$ soit une L-diffusion, il faut et il suffit que $\mathcal{P}$ soit solution du L-problème des martingales c'est-à-dire que pour chaque $f \in C_k^2(M)$ et chaque $x \in M$, le processus $H_t = f(X_t) - \int_0^t Lf(X_s) ds$ soit une martingale pour la pro-

babilité P_x .

Preuve : Si H_t est une martingale, on a $E_x(H_t) = E_x(H_0)$, ce qui prouve immédiatement (2). La réciproque est presque aussi simple (cf. Priouret [15] par exemple).

2.6. Equations stochastiques et diffusions : Soit U un ouvert de $\mathbb{R}^m$; soient resp. σ, b des applications localement lipschitziennes de U dans les matrices (m, k) et dans $\mathbb{R}^m$. Soit β_t un brownien dans $\mathbb{R}^k$. Les solutions de l'équation stochastique $dx_t = \sigma(x_t)d\beta_t + b(x_t)dt$ issues des différents points de U définissent une L-diffusion sur U , avec $L = \frac{1}{2} \sum_{1 \leq i, j \leq m} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq m} b_i(x) \frac{\partial}{\partial x_i}$, où $[a_{ij}(x)] = \sigma(x)\sigma^*(x)$, $b(x) = (b_1(x), \dots, b_m(x))$. Renvoyons à Priouret [15] et à M. Chaleyat-Maurel, ch. 1.

3. Existence et unicité des diffusions

Introduisons deux bonnes classes d'opérateurs L sur la variété M . Soit L un opérateur différentiel d'ordre 2 sur M .

3.1. Cas elliptique : On suppose que dans toute carte locale $\phi : U \rightarrow \mathbb{R}^n$, les coefficients de L (voir formule 2.1) vérifient

- pour chaque $x \in \phi(U)$, la matrice $a(x)$ est définie positive et à coefficients continus en x
- les coefficients $b_i(x)$, $i=1 \dots n$, $c(x)$ sont mesurables et localement bornés sur $\phi(U)$; de plus on a $c(x) \leq 0$.

3.2. Cas hypoelliptique : (condition de Hörmander locale) : on suppose M de classe C^∞ et l'existence d'un recouvrement ouvert de M tel que pour chaque ouvert U du recouvrement, L puisse s'écrire $L = \sum_{1 \leq i \leq r} Y_i^2 + Y + c$, où $c \leq 0$ est une fonction C^∞ sur M , et où $Y, Y_1 \dots Y_r$ sont des champs de vecteurs de classe C^∞ sur U , qui par crochets de Lie engendrent l'espace tangent $T_x(M)$ en tout point x de U ; bien entendu $r, Y, Y_1 \dots Y_r$ dépendent de U à priori. Renvoyons à Chaleyat-Maurel ch.10 pour une forme explicite de cette hypothèse en coordonnées locales.

3.3. THÉORÈME. Soit M une variété et soit L un opérateur différentiel sur M . Sous les hypothèses du "cas elliptique 3.1" ou bien du "cas hypoelliptique 3.2", il existe une L-diffusion sur M , unique à équivalence près.

3.4. Bibliographie : Dans la situation elliptique, le cas crucial $\{M = \mathbb{R}^n, \Delta \text{ à coef-}$

ficients bornés} étudié par Stroock-Varadhan [17] permet de traiter le cas général par une méthode de localisation (Karoui [12] Azencott [1]). Pour la situation hypo-elliptique, le cas $\{M = \mathbb{R}^n, \text{ condition de Hörmander globale}\}$ a été traité par Bonami et Altri [4]. La méthode de localisation se heurte à un problème d'unicité. La preuve du théorème 3.3 est esquissée en 3.7, après quelques rappels.

3.4. Processus induits : Soit U un ouvert d'une variété M ; soit $\mathcal{P}$ un processus de Markov fort sur M , à trajectoires p.s. dans $\mathcal{D}(M)$ (cf.1). Soit τ_U le temps de première sortie de U pour $\mathcal{P}$. Si Y_t est la "position" de $\mathcal{P}$ au temps t , on "définit" le processus $\mathcal{P}^U$ induit par $\mathcal{P}$ sur U en posant $Y_t^U = Y_t$ pour $t < \tau_U$ et $Y_t^U = \delta$ pour $t \geq \tau_U$. Le processus $\mathcal{P}^U$ est alors un processus de Markov fort sur U , à trajectoires p.s. dans $\mathcal{D}(U)$ (cf. Dynkin [10]).

Soit $(U_i)_{i \in I}$ un recouvrement ouvert de M ; pour chaque $i \in I$ soit $\mathcal{P}_i$ un processus de Markov fort sur U_i à trajectoires p.s. dans $\mathcal{D}(U_i)$. Posons $U_{ij} = U_i \cap U_j$. D'après Courrège-Priouret [7], pour qu'il existe un processus de Markov fort $\mathcal{P}$ sur M à trajectoires p.s. dans $\mathcal{D}(M)$, induisant sur chaque U_i un processus équivalent à $\mathcal{P}_i$, il faut et il suffit que $\mathcal{P}_i^{U_{ij}}$ et $\mathcal{P}_j^{U_{ij}}$ soient équivalents pour tout $i, j \in I$. Le processus $\mathcal{P}$ est alors unique à équivalence près lorsque les $\mathcal{P}_i$ sont donnés.

3.5. PROPOSITION. (Karoui [12] Azencott [1]). Soit M une variété et L un opérateur différentiel sur M du type 2.1. Soit $\mathcal{P}$ un processus de Markov fort sur M à trajectoires p.s. dans $\mathcal{D}(M)$. Soit $(U_i)_{i \in I}$ un recouvrement ouvert de M . Pour que $\mathcal{P}$ soit une L -diffusion il faut et il suffit que pour chaque $i \in I$, le processus induit $\mathcal{P}^{U_i}$ soit une L -diffusion sur U_i .

Preuve : Il s'agit de localiser le problème des martingales. Renvoyons à [12][1] ou à Priouret [15].

3.6. LEMME. Soit $\mathcal{P}$ une diffusion sur M et soit L son générateur différentiel. Alors pour toute fonction f de classe 2 sur M , tendant vers zéro à l'infini, et telle que Lf soit bornée sur M , on a $R_t f(x) - f(x) \equiv \int_0^t R_s Lf(x) ds$, où (R_t) est le semi-groupe de $\mathcal{P}$.

Preuve : On se donne un ouvert relativement compact U de M et une fonction $g \in C_k^2(M)$ coïncidant avec f sur U . La martingale $H_t = g(X_t) - \int_0^t Lg(X_s) ds$, arrêtée au temps de sortie τ_U , fournit pour $x \in U$ l'égalité

$$0 = E_x(H_{t \wedge \tau_U} - H_0) = E_x[f(X_{t \wedge \tau_U}) - \int_0^{t \wedge \tau_U} Lf(X_s) ds] - f(x)$$

Il suffit de faire croître U vers M pour conclure.

3.7. Preuve du théorème 3.3. Le cas elliptique est complètement traité dans Azencott [1] Priouret [15]. Bornons nous à étudier "le cas hypoelliptique 3.2". Soit U un ouvert de M sur lequel on a $L = \sum_{1 \leq i \leq n} Y_i^2 + Y$ contenu en 3.2. Soit ϕ une carte locale de domaine V inclus dans U . Soient $v_1(x) \dots v_r(x)$ les champs images de $Y_1 \dots Y_r$ par ϕ . La matrice (rectangulaire) $\sigma(x)$ ayant pour colonnes $v_1(x) \dots v_r(x)$ est à coefficients C^∞ et vérifie $\sigma(x)\sigma(x)^* = a(x)$, en notant $\frac{1}{2} a(x)$ la matrice des coefficients du 2nd-ordre de L dans la carte ϕ . La résolution d'une équation différentielle stochastique sur $\phi(V)$ garantit alors (cf. 2.6) l'existence d'une L -diffusion sur V . L'ensemble B des ouverts de M sur lesquels il existe une L -diffusion est donc une base (topologique) de M .

Soit B_1 l'ensemble des ouverts W de M tels que pour tout $\lambda > 0$, pour toute fonction $f \in C_k^2(W)$, il existe une fonction continue h sur $\bar{W}$, de classe 2 sur W , vérifiant

$$(3) \quad (L-\lambda)h = f \text{ sur } W, \text{ et } h = 0 \text{ sur } \partial W.$$

D'après Bony [5] (cf. Chaleyet-Maurel, ch. 10), B_1 est une base (topologique) de M . D'après la proposition 3.5, toute partie ouverte d'un élément de B est encore dans B ; par suite $B \cap B_1$ est encore une base de M .

Soient $W \in B \cap B_1$, $f \in C_k^2(W)$, $\lambda > 0$. Fixons une fonction continue h sur $\bar{W}$ vérifiant (3). Considérons une L -diffusion arbitraire sur W , et soit (S_t) son semi-groupe de transition. La relation (3) montre que $h = 0$ sur ∂W et que Lh est borné sur W . Donc (lemme 3.6), on a $S_t h(x) - h(x) = \int_0^t S_s Lh(x) dx$, $x \in W$, $t \geq 0$. Multiplions cette relation par $e^{-\lambda t}$ et intégrons de 0 à $+\infty$ pour obtenir, en posant $G = \int_0^{+\infty} e^{-\lambda t} S_t dt$

$$G_\lambda h(x) - \frac{1}{\lambda} h(x) = \int_0^{+\infty} dt e^{-\lambda t} \int_0^t ds S_s Lh(x) = \frac{1}{\lambda} G_\lambda Lh(x)$$

Comme $Lh = \lambda h + f$, on en conclut que $G_\lambda f(x) = -h(x)$ pour $x \in W$. Ainsi si (R_t) est le semi-groupe de transition d'une autre L -diffusion sur W , les fonctions continues (de t) $R_t f(x)$ et $S_t f(x)$ ont même transformées de Laplace, et donc coïncident, ce qui prouve $R_t = S_t$ pour $t \geq 0$, et l'équivalence de nos deux diffusions.

Finalement, pour tout $W \in B \cap B_1$, il existe sur W une L -diffusion unique à équivalence près, que nous noterons $\mathcal{R}_W$. Soit maintenant V un ouvert quelconque de M . Soient $\mathcal{P}$ et $\mathcal{Q}$ deux L -diffusions sur V . Puisque $B \cap B_1$ est une base, on a $V = \bigcup_{i \in I} W_i$ où les W_i sont dans $B \cap B_1$. Nécessairement, $\mathcal{P}^{W_i}$ et $\mathcal{Q}^{W_i}$ sont équivalents, pour tout $i \in I$, ce qui (cf. 3.4) entraîne l'équivalence de $\mathcal{P}$ et $\mathcal{Q}$. Sur tout ouvert de M (et sur M en particulier !) il existe donc au plus une L -diffusion à équivalence près.

En particulier si $W, W' \in B \cap B_1$, les diffusions induites $(\mathcal{R}_W)^{W \cap W'}$ et $(\mathcal{R}_{W'})^{W \cap W'}$ sont équivalentes. D'après 3.4, il existe donc un processus de Markov fort $\mathcal{R}$ sur M induisant $\mathcal{R}_W$ sur W , pour chaque $W \in B \cap B_1$. Le processus $\mathcal{R}$ est alors une L -diffusion grâce à 3.5.

4. Relation entre semi-groupe local et semi-groupe global

4.1. Oscillations entre deux fermés : Soient V, U deux ouverts de la variété M , tels que $\bar{V} \subset U$. Soit $\mathcal{P}$ une diffusion sur M , de temps de mort ζ . Pour $\mathcal{P}$ définissons la suite de temps d'arrêt suivants :

τ^1 est le temps de première sortie de U
 η^1 est le temps de première entrée dans $\bar{V}$ après τ^1
 τ^{m+1} est le temps de première sortie de U après η^m , avec la convention $\tau^{m+1} = +\infty$ si $\eta^m = +\infty$.
 η^{m+1} est le temps de première entrée dans $\bar{V}$ après τ^{m+1} , avec la convention $\eta^{m+1} = +\infty$ si $\tau^{m+1} = \zeta$ ou si $\tau^{m+1} = +\infty$. On a $\tau^1 \leq \eta^1 \leq \dots \leq \tau^m \leq \eta^m \dots$ et les temps η^m marquent les retours successifs de $\mathcal{P}$ dans $\bar{V}$ après sortie de U .

Nous allons démontrer une formule permettant le calcul du semi-groupe de transition de $\mathcal{P}$ à partir de celui de $\mathcal{P}^U$; une formule analogue a été introduite par Doob [8] dans le cas du mouvement brownien sur un espace de Green.

4.2. THEOREME. (Azencott [1]). Soit $\mathcal{P}$ une diffusion sur une variété M ; soit U un ouvert de M . Notons (R_t) le semi-groupe de transition de $\mathcal{P}$, et (R_t^U) celui de la diffusion induite $\mathcal{P}^U$. Soit V un ouvert vérifiant $\bar{V} \subset U$ et ayant une frontière ∂V compacte. Alors pour chaque $x \in M$, il existe une mesure de Radon positive ν_x sur $\mathbb{R}^+ \times \partial V$ ayant la propriété suivante

$$(4) \quad R_t f(x) = I_U(x) R_t^U f(x) + \int_{[0, t] \times \partial V} R_{t-s}^U f(y) d\nu_x(s, y)$$

pour tout $x \in M$, $t > 0$, f fonction mesurable bornée sur M , avec la convention $I_U R_t^U \equiv 0$ hors de U .

La mesure ν_x est définie par

$$(5) \quad \int_{\mathbb{R}^+ \times \partial V} g d\nu_x = E_x \left[\sum_{m=1}^{+\infty} I_{\{\eta^m_{\text{fini}}\}} g(\eta^m, X_{\eta^m}) \right]$$

où g est une fonction borélienne bornée quelconque sur $\mathbb{R}^+ \times \partial V$, et où les η^m sont les temps de retour dans $\bar{V}$ après sortie de U , définis en 4.1.

De plus, pour tout compact K de $\mathbb{R}^+ \times \partial V$, la fonction $x \longrightarrow \nu_x(K)$ est bornée sur M .

Preuve : Voir plus bas en 4.6, après quelques lemmes techniques. Notons d'emblée un utile corollaire.

4.3. COROLLAIRE. Soit $\mathcal{P}$ une diffusion sur une variété M . Soit $(U_i)_{i \in I}$ un recouvrement ouvert de M . Soit π une mesure de Radon positive sur M . Pour que le semi-groupe de transition de $\mathcal{P}$ admette une densité $p(t, x, z)$ par rapport à π , il faut et il suffit

que pour chaque $i \in I$, le semi-groupe de la diffusion induite $\mathcal{P}^U_i$ admette une densité par rapport à π .

Lorsque cette situation est réalisée, on peut alors choisir $p(t, x, z)$ mesurable en (t, x, z) . De plus si U est un ouvert de M , la diffusion induite $\mathcal{P}^U$ admet une densité mesurable $q(t, x, z)$ vérifiant

$$(6) \quad p(t, x, z) = I_U(x) q(t, x, z) + \int_{[0, t[\times V} q(t-s, y, z) dv_x(s, y)$$

pour $x \in L$, $t > 0$, π -presque tout $z \in V$, où l'ouvert V et la mesure dv_x associée sont comme en 4.2. [par convention, le terme $I_U(x) q(t, x, z)$ est nul si $x \notin U$].

Preuve : La première partie se déduit de 4.2 et de l'inégalité immédiate

$$(7) \quad R_t^U f(x) \leq R_t f(x) \quad x \in U ; t > 0 ; f \geq 0 \text{ mesurable, nulle hors de } U.$$

L'existence de versions mesurables pour p et q résulte de la possibilité de construire concrètement une dérivée de Radon-Nicodým en approchant la tribu borélienne par des tribus localement finies (cf. Doob [9]). La relation entre p et q se déduit alors directement de 4.2.

4.4. LEMME. Soit $\mathcal{P}$ une diffusion sur M , et soit d une distance riemannienne sur M . Notons $B(x, r)$ la boule de centre $x \in M$ et de rayon r , et $\tau_{x, r}$ le temps de lère sortie de cette boule. Pour tout compact K de M et tout $r > 0$ il existe alors $C > 0$ tel que

$$P_x(\tau_{x, r} < t \wedge \zeta) \leq Ct \quad t \geq 0, x \in K$$

Preuve : On se ramène trivialement au cas où K est inclus dans le domaine U d'une carte locale, et où r est assez petit pour que $[\bigcup_{x \in K} \overline{B(x, r)}]$ soit une partie compacte de U . Comme les processus $\mathcal{P}$ et $\mathcal{P}^U$ sortent alors de $B(x, r)$ au même instant, il suffit d'étudier le cas $\mathcal{P} = \mathcal{P}^U$. Supposons donc $M = U$ ouvert de $\mathbb{R}^n$, et

$[\bigcup_{x \in K} \overline{B(x, r)}]$ partie compacte de U .

Il existe alors $u > 0$ tel que $x \in K$, $y \in M$, et $\|x-y\| \leq u$, entraînent $d(x, y) < \frac{1}{2}r$. Fixons une fonction g de classe 2 sur $\mathbb{R}^+$ telle que $g(0) = 1$, $0 \leq g \leq 1$ sur $[0, u^2]$, et posons $f_x(y) = g(\|x-y\|^2)$. La fonction f_x est dans $C_k^2(M)$ et vérifie

$$(8) \quad f_x \leq 1_{B(x, \frac{r}{2})} \quad \text{pour } x \in K$$

$$(9) \quad |Lf_x(y)| \leq C \quad \text{pour } x \in K, y \in M$$

où C est une constante, et où L est le générateur différentiel de $\mathcal{P}$.

Pour $x \in K$, l'évènement $\{\tau_{x, r} < t \wedge \zeta\}$ implique $\{X_{t \wedge \tau_{x, r}} \notin B(x, \frac{r}{2})\}$. D'après (8)

on a donc

$$(10) \quad P_x(\tau_{x,r} < t \wedge \zeta) \leq 1 - E_x[f_x(X_{t \wedge \tau_{x,r}})] ; x \in K.$$

La martingale $H_t = f_x(X_t) - \int_0^t Lf_x(X_s) ds$ vérifie $E_x[H_{t \wedge \tau_{x,r}} - H_0] = 0$, et donc, pour $x \in K$,

$$(11) \quad 1 - E_x[f_x(X_{t \wedge \tau_{x,r}})] = - E_x \int_0^{t \wedge \tau_{x,r}} f_x(X_s) ds.$$

Le résultat annoncé est conséquence directe de (9)(10)(11).

4.5. LEMME. (d'après Dynkin [10]) Fixons $\mathcal{P}, M, U, V$ comme en 4.2. Soient η^n , $n \geq 1$, les successifs temps de retour de $\mathcal{P}$ dans $\bar{V}$ après sortie de U (déf. 4.1). A chaque intervalle borné $J \subset \mathbb{R}^+$ associons, pour le processus $\mathcal{P}$, le nombre aléatoire N_J d'oscillations entre U^c et $\bar{V}$ dans l'intervalle de temps J ; N_J est le cardinal de l'ensemble des $n \geq 1$ tels que $\eta^n \in J$. Alors pour J fixé, il existe des constantes positives C, D , avec $D < 1$, telles que

$$P_x(N_J \geq k) \leq C D^k \text{ pour tout } k \geq 0, x \in M$$

Preuve : Pour chaque $b \geq 0$ posons

$$\phi_b(y, t) = P_y(\eta^1 \leq b - t) \quad y \in \partial V, 0 \leq t \leq b$$

La propriété de Markov forte (dans l'espace-temps) donne pour $k \geq 1, x \in M, b \geq 0$

$$(12) \quad P[\eta^{k+1} \leq b | \mathcal{F}_{\eta^k}] = \phi_b(X_{\eta^k}, \eta^k)$$

relation vraie P_x p.s. sur $\{\eta^k \leq b\}$.

Fixons une distance riemannienne sur M . Puisque ∂V est compact, il existe $r > 0$ tel que pour $y \in \partial V$ les boules $B(y, r)$ soient contenues dans U . Pour $y \in \partial V$, le temps de sortie $\tau_{y,r}$ de $\overline{B(y, r)}$ est majoré par η^1 , d'où grâce à 4.4, l'existence de C vérifiant

$$\phi_b(y, t) \leq P_y(\tau_{y,r} < b \wedge \zeta) \leq Cb \text{ pour } y \in \partial V, 0 \leq t \leq b$$

De (12) on déduit alors, pour $k \geq 1, x \in M, b \geq 0$

$$P_x(\eta^{k+1} \leq b) \leq Cb P_x(\eta^k \leq b)$$

Fixons alors $b = \frac{1}{2C}$ pour obtenir

$$(13) \quad P_x(\eta^k \leq b) \leq \left(\frac{1}{2}\right)^{k-1} \text{ pour } x \in M, k \geq 1.$$

Etant donné un intervalle borné J de $\mathbb{R}^+$ fixons un entier p et une partition $J_1 \dots J_p$ de J en p intervalles disjoints de longueurs inférieures à b . Comme $\{\eta^k \leq b\}$

équivaut à $\{N_{[0,b]} \geq k\}$, la relation (13) entraîne

$$(14) \quad P_x(N_{J_i} \geq k) \leq \left(\frac{1}{2}\right)^{k-1} \text{ pour } x \in M, k \geq 1, i=1 \dots p. \text{ d'où, pour } x \in M, k \geq 1$$

$$P_x(N_J \geq k) = P_x(N_{J_1} + \dots + N_{J_k} \geq k) \leq \sum_{i=1}^p P_x(N_{J_i} \geq \frac{k}{p}) \leq p \left(\frac{1}{2}\right)^{\frac{k}{p} - 1} \text{ ce qui}$$

prouve le lemme.

4.6. Preuve du théorème 4.2. (Notations de 4.1, 4.2, 4.5). Fixons un intervalle borné J de $\mathbb{R}^+$ et posons $h(t,y) = 1_J(t)$ pour $(t,y) \in \mathbb{R}^+ \times \partial V$. Le lemme 4.5 fournit des constantes $C, D < 1$ telles que pour tout $x \in M$ on ait

$$(15) \quad E_x \left[\sum_{m \geq 1} \frac{1}{\{\eta^m \text{ fini}\}} h(\eta^m, X_{\eta^m}) \right] = \sum_{m \geq 1} P_x(N_J \geq m) \leq C \frac{D}{1-D}.$$

La formule (5) définit donc bien une mesure de Radon ν_x sur $\mathbb{R}^+ \times \partial V$, et (15) montre que pour tout compact K de $\mathbb{R}^+ \times \partial V$, $\nu_x(K)$ reste borné quand x décrit M .

La propriété de Markov forte entraîne, pour f fonction borélienne bornée sur M , et $m \geq 1$

$$(16) \quad E[f(X_t) 1_{\{\eta^m < t < \tau^{m+1}\}} | \mathcal{F}_{\eta^m}] = \psi(\eta^m, X_{\eta^m}), \text{ } P_x\text{-p.s. sur } \{\eta^m < t\}$$

où l'on a posé

$$\psi(s,y) = E_y[f(X_{t-s}) 1_{\{t-s < \tau\}}] \text{ pour } s \leq t, y \in \partial V$$

Imposons maintenant $\{f \text{ nulle hors de } V\}$, ce qui implique $\{f \text{ nulle hors de } U\}$ et donc

$$(17) \quad \psi(s,y) = R_{t-s}^U f(y) \text{ pour } s \leq t, y \in \partial V.$$

La définition de τ^m, η^m donne, pour $m \geq 1$

$$(18) \quad f(X_t) 1_{\{\tau^m \leq t < \tau^{m+1}\}} \equiv f(X_t) 1_{\{\eta^m < t < \tau^{m+1}\}}$$

Les relations (16)(17)(18) impliquent pour $x \in M, m \geq 1$

$$(19) \quad E_x[f(X_t) 1_{\{\tau^m \leq t < \tau^{m+1}\}}] = E_x[R_{t-\eta^m}^U f(X_{\eta^m}) 1_{[0,t[}(\eta^m)]$$

grâce à (15), on peut sommer en m , et obtenir pour $x \in M, t \geq 0$

$$(20) \quad E_x[f(X_t) 1_{\{\tau \leq t\}}] = E_x \left[\sum_{m \geq 1} g(\eta^m, X_{\eta^m}) 1_{\{\eta^m \text{ fini}\}} \right] = \int_{\mathbb{R}^+ \times \partial V} g \, d\nu_x$$

où l'on a posé $g(s,y) = R_{t-s}^U f(y) 1_{[0,t[}(s)$.

Par définition on a d'autre part

$$(21) \quad E_x[f(X_t) 1_{\{t < \tau\}}] = I_U(x) 1_t^U f(x) \text{ pour } x \in M, t \geq 0$$

avec la convention $\{I_U R_t^U \text{ nul hors de } U\}$. Il suffit d'additionner (20) et (21) pour prouver (4) et donc le théorème 4.2.

5. EXISTENCE DE DENSITÉS LISSES

Rappelons qu'une fonction f est dite localement hölderienne sur un ouvert U de $\mathbb{R}^n$, si pour tout compact K de U on peut trouver $\alpha > 0$, $C > 0$ tels que $|f(x) - f(y)| \leq C \|x-y\|^\alpha$ pour $x, y \in K$.

5.1. THEOREME. Soit M une variété et $\mathcal{P}$ une diffusion sur M , de générateur différentiel L . Soit π un volume riemannien sur M .

i) Si L vérifie les hypothèses du "cas elliptique 3.1", alors le semi-groupe de transition de $\mathcal{P}$ admet une densité $p(t, x, z)$ par rapport à π , mesurable en (t, x, z) .

ii) Si L est elliptique et à coefficients localement hölderiens, le semi-groupe de $\mathcal{P}$ admet par rapport à π une densité p continue sur $\mathbb{R}^+ \times M \times M$, et pour chaque $z \in M$, la fonction $p(t, x, z)$ est de classe 2 en $x \in M$, de classe 1 en $t > 0$, et vérifie (comme fonction de t, x)

$$(22) \quad \left(\frac{\partial}{\partial t} L\right)p(t, x, z) = 0 \quad (t, x) \in]0, +\infty[\times M$$

iii) Supposons M de classe C^∞ , et L de la forme (localement) $L = \sum_{1 \leq i \leq r} Y_i^2 + Y + c$, avec c fonction C^∞ , $c \leq 0$, $Y, Y_1, \dots, Y_r$ champs C^∞ , et où l'algèbre de Lie $\mathcal{L}$ engendrée par $Y, Y_1, \dots, Y_r$ a la propriété suivante :

$[\mathcal{L}, \mathcal{L}]$ et $\{Y_1, \dots, Y_r\}$ suffisent à engendrer l'espace tangent à M en tout point (du domaine où sont définis $Y, Y_1, \dots, Y_r$).

Alors le semi-groupe de la L -diffusion admet par rapport à π une densité $p(t, x, z)$, de classe C^∞ sur $]0, +\infty[\times M \times M$, qui pour chaque $z \in M$ vérifie (22) en t, x .

5.2. Démonstration de 5.1 (iii). Notons que l'hypothèse sur L revient à supposer que L et $\frac{\partial}{\partial t} - L$ vérifient, sur M et $\mathbb{R}^+ \times M$ resp., l'hypothèse du cas hypoelliptique 3.2. La démonstration est alors calquée sur celle de Bonami et Altri [4] qui traitent le cas $M = \mathbb{R}^n$ avec une condition de Hörmander "forte" et globale. Comme en [4] on peut d'ailleurs aussi prouver que pour chaque $x \in M$, $p(t, x, z)$ vérifie en (t, z) l'équation adjointe $(\frac{\partial}{\partial t} - L^*)p = 0$, où L^* est l'adjoint de L par rapport à π .

5.3. Démonstration de 5.1 (i) (cas elliptique 3.1). Le cas crucial où $M = \mathbb{R}^n$ et où L est à coefficient bornés, avec $L1 = 0$, a été étudié par Stroock-Varadhan [17]. Une double application de 4.3 permet de passer au cas $M =$ ouvert de $\mathbb{R}^n$ puis à celui d'une variété quelconque M , avec L vérifiant 3.1 et $L1 = 0$. Pour passer au cas de $(L+c)$ où $L1 = 0$, L vérifie 3.1, et $c \leq 0$ est une fonction mesurable localement bornée sur M , il suffit de remarquer que le semi-groupe (S_t) de la $(L+c)$ -diffusion et le semi-groupe (R_t) de la L -diffusion vérifient pour $x \in M$, $f \geq 0$ mesurable,

$$S_t f(x) = E_x\{f(X_t) \exp[-\int_0^t c(X_s) ds]\} \leq R_t f(x)$$

5.4. Démonstration de 5.1 (ii) (cas elliptique, localement hölderien). La difficulté ici vient du fait qu'on ne peut pas employer la méthode de Mac-Kean, utilisée dans le cas (iii) par Bonami et altri [4], car on ne peut pas utiliser l'adjoint L^* de L .

Plaçons nous dans la situation (ii). Soit U un ouvert relativement compact de M , à bord lisse, tel que $\bar{U}$ soit contenu dans le domaine d'une carte locale ϕ . L'image de L par ϕ est un opérateur D , uniformément elliptique, à coefficients uniformément bornés et uniformément hölderiens sur $\phi(U)$. D'après Friedmann [11] la D -diffusion sur $\phi(U)$ admet une densité $r(t, \xi, \eta)$ continue sur $]0, +\infty[\times \phi(U) \times \phi(U)$.

Soit D_1 un opérateur différentiel uniformément elliptique, à coefficients uniformément hölderiens et uniformément bornés sur $\mathbb{R}^n$, tel que D_1 coïncide avec D sur $\phi(U)$. D'après Friedmann [11] la D_1 -diffusion sur $\mathbb{R}^n$ admet une densité $k(t, \xi, \eta)$ continue en (t, ξ, η) et vérifiant

(23) $k(t, \xi, \eta) \leq A t^{-n/2} \exp[-B \frac{\|\xi - \eta\|^2}{t}]$ pour $\xi, \eta \in \mathbb{R}^n$, $0 < t \leq T$ où T est arbitraire fixé, et A, B sont des constantes positives strictes dépendant de T . L'inégalité 4.3 (2) donne

(24) $r(t, \xi, \eta) \leq k(t, \xi, \eta)$ pour $t > 0, \xi, \eta \in \phi(U)$.

Fixons une distance riemannienne d sur M ; transportés par le difféomorphisme ϕ , les résultats précédents montrent que la L -diffusion sur U admet une densité $q(t, y, z)$ continue sur $]0, +\infty[\times U \times U$, et que pour chaque $T > 0$, il existe $A, C > 0$ tels que

(25) $q(t, y, z) \leq A t^{-n/2} \exp[-\frac{C d(y, z)^2}{t}]$ pour $y, z \in U$, $0 \leq t \leq T$.

Donnons nous arbitrairement $t_0 > 0$, $z_0 \in U$; fixons $T > t_0$ et deux ouverts W, V de M tels que $z_0 \in W$, $\bar{W} \subset V$, $\bar{V} \subset U$. Soient ν_x , $x \in M$, les mesures de Radon sur $\mathbb{R}^+ \times \partial V$ introduites en 4.2. Soit $\varepsilon > 0$. La fonction $q(t-s, y, z)$ est continue en s, t, y, z , et donc uniformément continue sur le compact

$$K = \{0 \leq s \leq t - \varepsilon, t \leq T, y \in \partial V, z \in \bar{W}\}$$

Comme $\nu_x([0, T] \times \partial V)$ reste borné quand x décrit M (cf. 4.2), on voit que pour chaque $\varepsilon > 0$, l'intégrale

$$\int_{[0, t-\varepsilon] \times \partial V} q(t-s, y, z) d\nu_x(s, y)$$

est continue en $z \in \bar{W}$, uniformément lorsque (t, x) décrit $[2\varepsilon, T] \times M$.

D'autre part la majoration (25) et celle de la fonction $x \rightarrow \nu_x([0, T] \times \partial V)$ montrent que lorsque $\varepsilon \rightarrow 0$, l'intégrale

$$\int_{[t-\varepsilon, t]} q(t-s, y, z) d\nu_x(s, y)$$

tend vers 0, uniformément lorsque (t, x, z) décrit $[\varepsilon, T] \times M \times \bar{W}$. Ces deux remarques impliquent que l'intégrale

$$F(t, x, z) = \int_{[0, t] \times \partial V} q(t-s, y, z) d\nu_x(s, y)$$

est continue en $z \in \bar{W}$, uniformément lorsque (t,x) décrit $J \times M$ où J est un compact de $]0, +\infty[$. De plus le même argument montre que la fonction $F(t,x,z)$ reste bornée lorsque (t,x,z) décrit $J \times M \times \bar{W}$, avec J compact fixé de $]0, +\infty[$.

Le corollaire 4.3 montre que la densité mesurable p de la L -diffusion $\mathcal{P}$ sur M , (densité qui existe d'après l'étude de (i) par exemple) vérifie pour $t > 0$; $x \in M$, et π -p.t.t. $z \in M$

$$(26) \quad p(t,x,z) = I_U(x) q(t,x,z) + F(t,x,z)$$

Les ouverts W, V, U étant arbitraires, on en conclut aisément qu'il existe une version (nécessairement unique) $p(t,x,z)$ continue en $z \in M$, uniformément en $(t,x) \in J \times M$, J compact de $]0, +\infty[$; de plus pour chaque $z \in M$, $p(t,x,z)$ reste bornée lorsque (t,x) décrit $J \times M \times L$, J comme ci-dessus, L compact de M .

Fixons $t > 0$ et $z \in M$. La fonction $g(x) = p(t,x,z)$ est mesurable bornée sur M . D'après la proposition 5.5 ci-dessous, si (R_s) est le semi-groupe de $\mathcal{P}$, la fonction $R_s g(x)$ est de classe 2 en $x \in M$, de classe 1 en $s > 0$ et vérifie

$$(27) \quad \left(\frac{\partial}{\partial s} - L\right) R_s g(x) = 0 \text{ pour } s > 0, x \in M$$

Mais puisque $R_s R_t = R_{s+t}$, il est évident que pour chaque s, x donnés, les densités $p(s+t, x, z)$ et $\psi(z) = \int_M p(s, x, y) p(t, y, z) d\pi(y)$ doivent coïncider π -pp. en $z \in M$. Comme $p(t, y, z)$ reste bornée pour $(t, y, z) \in J \times M \times L$, $J \times L$ compact de $\mathbb{R}^+ \times M$, le théorème de convergence dominée montre que $\psi(z)$ est continue en z , donc coïncide avec $p(t+s, x, z)$ pour tout z . Finalement on voit que $R_s g(x) = p(t+s, x, z)$ pour tout $s > 0$, $t > 0, x, z \in M$. On en conclut, grâce à (27), que $p(u, x, z)$ est de classe 2 en $x \in M$, de classe 1 en $u > 0$ et vérifie en (u, x)

$$\left(\frac{\partial}{\partial u} - L\right) p(u, x, z) = 0 \text{ pour } u > 0, x \in M$$

ceci pour chaque $z \in M$. La fonction $p(u, x, z)$ est continue en (u, x) pour chaque z , et uniformément continue en z lorsque (u, x) décrit un compact de $\mathbb{R}^+ \times M$; par suite p est continue sur $\mathbb{R}^+ \times M \times M$.

5.5. PROPOSITION. *Soit M une variété ; soit $\mathcal{P}$ une L -diffusion sur M , de semi-groupe de transition (R_t) . Supposons que L soit elliptique à coefficients localement hölderiens, ou bien que L et $\frac{\partial}{\partial t} - L$ vérifient la condition de Hörmander locale (hypothèse 5.1 (iii)). Alors pour toute fonction mesurable bornée f sur M , $R_t f(x)$ est de classe 2 en $x \in M$, de classe 1 en $t > 0$ et vérifie*

$$\left(\frac{\partial}{\partial t} - L\right) R_t f(x) = 0 \text{ pour } t > 0, x \in M$$

Preuve : Pour le cas elliptique localement hölderien, nous renvoyons à Azencott [2]. Considérons le cas hypoelliptique 3.2. Si $f \in C_k^2(M)$ le résultat se démontre comme dans le cas $M = \mathbb{R}^n$, condition de Hörmander globale pour lequel nous référons à

la preuve du théorème V.2 dans Bonami et altri [4]. Pour passer au cas où f est mesurable bornée, on remarque que l'ensemble des solutions de $(\frac{\partial}{\partial t} - L)g = 0$ sur $\mathbb{R}^+ \times M$ est stable par passage au limites croissantes localement bornées, car $(\frac{\partial}{\partial t} - L)$ étant hypoelliptique, les solutions faibles de $(\frac{\partial}{\partial t} - L)g = 0$ sont aussi solutions fortes (cf. Chaleyat-Maurel ch.10).

BIBLIOGRAPHIE

- [1] R. Azencott, *Methods of localization and diffusions on manifolds*, 1972, à paraître.
- [2] R. Azencott, *Behaviour of diffusion semi groups at infinity*, Bull. Soc. Math. France 102, 1974, p. 193-240.
- [3] R. Azencott, *Diffusions et équations différentielles stochastiques*. Cours E.N.S. Ulm 1978-1979, à paraître.
- [4] A. Bonami, N. Karoui, B. Roynette et H. Reinhardt. *Processus de diffusion associé à un opérateur elliptique dégénéré*, Ann. Inst. Henri Poincaré Sect.B t.7, 1971, p. 31-80.
- [5] J.M. Bony, *Opérateurs elliptiques dégénérés*, Séminaire de théorie du potentiel Brelot Choquet Deny, 1967-1968.
- [6] J.M. Bony, P. Courrège et P. Priouret, *Semi groupes de Feller sur une variété à bord compacte*, Ann. Inst. Fourier, Grenoble t.18, 1968, p. 369-521.
- [7] P. Courrège et P. Priouret, *Recollement de processus de Markov*, Publ. Inst. Stat. Univ. Paris 14, 1965, p.275-337.
- [8] J.L. Doob, *Brownian motion on a green space*, Th. Prob. Appl. vol. 2, 1957, p.1-30.
- [9] J.L. Doob, *Stochastic processes*, Wiley.
- [10] E.B. Dynkin, *Markov processes*, Springer-Verlag-Berlin, 1965.
- [11] A.P. Friedmann, *Partial differential equations of parabolic type*, Englewood Cliffs, Prentice Hall, 1964.

- [12] N. Karoui, *Processus de diffusion associé à un opérateur elliptique dégénéré et à une condition frontière*. Thèse Université de Paris VI., 1971.
- [13] P. Malliavin, *Géométrie différentielle stochastique*, Presses de l'Université de Montréal, 1978.
- [14] P.A. Meyer, *Processus de Markov*, Lecture Notes Math. v.26 Springer Verlag, 1967.
- [15] P. Priouret, *Diffusions et équations différentielles stochastiques*, Ecole d'été Saint-Flour 1973, Lecture Notes Math, v. 390, Springer Verlag, 1974.
- [16] L. Schwartz, *Semimartingales sur des variétés, martingales conformes sur des variétés analytiques complexes*. Lecture Notes Math. v. 787, Springer Verlag, 1980.
- [17] D. Stroock et S. Varadhan, *Diffusion processes with continuous coefficients*, Com. Pure App. Math. 22 (1969) p.345-400 et 479-530.

CHAPITRE 3

GÉODESIQUES ET CONNEXIONS

par

Philippe BOUGEROL

Etant donné un opérateur différentiel elliptique du second ordre sur une variété M et $\{p_t, t \in \mathbb{R}^+\}$ la famille de densités du semi-groupe associé, nous allons voir dans les chapitres suivants que l'on peut munir M d'une structure riemannienne qui permet d'exprimer géométriquement un équivalent de p_t quand t tend vers zéro. Aussi allons nous rappeler les premières propriétés des variétés riemanniennes, en particulier celles qui concernent les géodésiques sous les hypothèses de différentiabilité que nous utiliserons ensuite. On trouvera plus de détails dans Cheeger Ebin (1) ou Milnor (2) par exemple.

1. Introduction

1.1. Soient M une variété de classe $\mathcal{C}^k$, $k \geq 1$, de dimension m , $T_x M$ son espace tangent au point x , TM son fibré tangent, c'est-à-dire $\bigcup_{x \in M} T_x M$.

Considérons une carte (U, ϕ) de M où U est un ouvert de M et ϕ une application de U dans $\mathbb{R}^m$. Pour tout x de U on note $\frac{\partial}{\partial x_i}(x)$, $1 \leq i \leq m$, l'élément de $T_x M$ qui à la fonction f de $\mathcal{C}^k(M)$ associe la $i^{\text{ème}}$ dérivée partielle de $f \circ \phi^{-1}$ calculée au point $\phi(x)$ et x_i l'application de U dans $\mathbb{R}$ égale à la $i^{\text{ème}}$ composante de ϕ . Si v est un élément de $T_x M$ il est égal à $\sum_{i=1}^m v(x_i) \frac{\partial}{\partial x_i}(x)$. On en déduit que l'application

$\bar{\phi} : \bigcup_{x \in U} T_x M \longrightarrow \phi(U) \times \mathbb{R}^m$ qui à v dans $T_x M$ associe $\bar{\phi}(v) = (\phi(x), v(x_1), \dots, v(x_m))$ est une bijection. Nous munissons TM de la structure de variété de classe $\mathcal{C}^{k-1}$ définie à partir des applications $\bar{\phi}$.

Si n est un entier positif inférieur ou égal à $k-1$, soit $\mathcal{D}^{(n)}(M)$ l'ensemble des champs de vecteurs (tangents) de classe $\mathcal{C}^n$ sur M , c'est-à-dire l'ensemble des applications X de M dans TM , de classe $\mathcal{C}^n$, telles que si $x \in M$, $X(x)$ noté aussi X_x appartient à $T_x M$. Localement $X(x)$ s'écrit alors $\sum a_i(x) \frac{\partial}{\partial x_i}(x)$, où les applications a_i sont de classe $\mathcal{C}^n$.

1.2. Une structure riemannienne g (on dit aussi métrique riemannienne) sur une variété M est la donnée en chaque point x de M d'un produit scalaire g_x sur $T_x M$. Pour v, w dans $T_x M$ on note $\langle v, w \rangle_x$ ou $\langle v, w \rangle$ l'expression $g_x(v, w)$. On dit que g est une structure riemannienne de classe $\mathcal{C}^h$, ($h \leq k-1$), si pour X et Y dans $\mathcal{D}^{(h)}(M)$ la fonction $x \longmapsto \langle X, Y \rangle_x$ ($= \langle X_x, Y_x \rangle$) est de classe $\mathcal{C}^h$. Si localement $\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle_x = g_{ij}(x)$, ce que l'on écrit parfois $ds^2 = \sum g_{ij} dx_i dx_j$, cela revient à dire que les fonctions g_{ij} sont de classe $\mathcal{C}^h$.

Pour abrégé on appellera variété riemannienne de classe $\mathcal{C}^r$, $r \geq 0$, une variété de classe $\mathcal{C}^{r+1}$ munie d'une structure riemannienne de classe $\mathcal{C}^r$. Sur une sous-variété M de $\mathbb{R}^n$ on peut par exemple définir une structure riemannienne en restreignant à $T_x M$ le produit scalaire usuel de $\mathbb{R}^n$.

1.3. Si M et N sont deux variétés et f une application de M dans N différentiable au point x soit df_x sa différentielle en ce point. On note alors df la fonction de TM dans TN qui à v dans $T_x M$ associe $df_x(v)$ si celui-ci existe.

On appelle courbe sur M une application c d'un intervalle I de $\mathbb{R}$ dans M , continue, $\mathcal{C}^1$ par morceaux. Si c est différentiable en t , son vecteur vitesse $\dot{c}(t)$ est l'élément $dc_t(\frac{d}{dt})$ de $T_{c(t)} M$. La longueur de la courbe c sur I est $L(x) = \int_I \|\dot{c}(t)\| dt$. Si x et y sont dans M on appelle distance riemannienne entre x et y le nombre $d(x, y)$ égal à la borne inférieure des longueurs des courbes d'extrémités x et y . Nous verrons que si M est connexe on définit bien ainsi une distance.

Au paragraphe 3 nous allons introduire les géodésiques à partir des connexions. Pour comprendre à quoi correspond ce nouvel objet examinons le cas d'une variété riemannienne de dimension deux de façon heuristique.

Appellons dans ce cas géodésique une courbe différentiable $c : I \longrightarrow M$ telle que, sur I , $\|\dot{c}(t)\|$ est constant et pour t et t' assez proches, c réalise la distance entre $c(t)$ et $c(t')$. Soit X un champ de vecteurs sur M et x un point de M . Supposons qu'il existe une géodésique c définie sur un voisinage de 0 vérifiant $c(0) = x$, $\dot{c}(0) = X(x)$. A un élément Z de $\mathcal{D}^{(1)}(M)$ et à t assez proche de 0 on associe le vecteur

Z^t de $T_x M$ de même norme que $Z_{c(t)}$ et tel que $\langle Z^t, X(x) \rangle$ soit égal à $\langle Z_{c(t)}, \dot{c}(t) \rangle$. (On peut interpréter Z^t comme le transporté parallèle de $Z_{c(t)}$ le long de c). On considère alors le vecteur de $T_x M$ défini par $\nabla_X Z(x) = \lim_{t \rightarrow 0} \frac{Z^t - Z(x)}{t}$. Admettant que cette limite existe et que $x \mapsto \nabla_X Z(x)$ est un champ de vecteurs, on montre facilement que l'application qui à X, Z dans l'espace vectoriel réel $\mathcal{D}^{(1)}(M)$ associe $\nabla_X Z$ est bilinéaire et que si f est différentiable, $\nabla_X(fZ)(x) = f(x)\nabla_X Z(x) + Xf(x)Z(x)$.

D'après notre définition $\nabla_X Z(x)$ ne dépend que de X_x , on peut le noter $\nabla_{X_x} Z$. Remarquons que si Z est parallèle le long de c (i.e. $Z^t = Z_x$), $\nabla_{\dot{c}(t)} Z$ est nul et qu'en particulier ceci est réalisé si $Z_{c(t)} = \dot{c}(t)$.

En général nous allons commencer par définir un objet ∇ vérifiant les conditions au dessus et une géodésique sera, approximativement, une courbe c telle que si $Z_{c(t)} = \dot{c}(t)$, $\nabla_{\dot{c}(t)} Z$ est nul. Nous en retrouverons plus tard la définition "naïve" que nous avons donnée.

2. Connexions

2.1. DÉFINITION. Soient n et k deux entiers, $k \geq 3$, $1 \leq n \leq k-2$ et M une variété de classe $\mathcal{C}^k$. Une *connexion* de classe $\mathcal{C}^n$ sur M est une application bilinéaire ∇ de $\mathcal{D}^{(1)}(M) \times \mathcal{D}^{(1)}(M)$ dans $\mathcal{D}^{(0)}(M)$ vérifiant

$$a) \quad \nabla_{fX} Y(x) = f(x)\nabla_X Y(x) \text{ et } \nabla_X(fY)(x) = f(x)\nabla_X Y(x) + X(f)(x)Y(x)$$

si X, Y sont dans $\mathcal{D}^{(1)}(M)$, f dans $\mathcal{C}^1(M)$ et x dans M .

$$b) \quad \nabla_X Y \text{ est dans } \mathcal{D}^{(n)}(M) \text{ si } X \in \mathcal{D}^{(n)}(M) \text{ et } Y \in \mathcal{D}^{(n+1)}(M)$$

$\nabla_X Y$ s'appelle la *dérivée covariante* de Y dans la direction X .

Soient U un ouvert de M , X un champ nul sur U , f une fonction différentiable égale à 1 en un point x de U , à 0 hors de U . Alors, si $Y \in \mathcal{D}^{(1)}(M)$,

$$\nabla_X Y(x) = f(x)\nabla_X Y(x) = \nabla_{fX} Y(x) = 0$$

De même si Y est nul sur U , pour tout X de $\mathcal{D}^{(1)}(M)$,

$$\nabla_X Y(x) = X(f)(x)Y(x) + f(x)\nabla_X Y(x) = \nabla_X(fY)(x) = 0$$

donc en général $\nabla_X Y$ ne dépend, sur U , que de X et Y sur U .

Soit U , domaine d'une carte (U, ϕ) , on peut alors définir sans ambiguïté

$$\nabla \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j}. \text{ Si, sur } U$$

$$\nabla \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} = \sum_k \Gamma_{ij}^k \frac{\partial}{\partial x_k},$$

les Γ_{ij}^k s'appellent les symboles de Riemann Christoffel, le b) de la définition montre que ce sont des applications de classe $\mathcal{C}^{n-1}$. Réciproquement les Γ_{ij}^k caractérisent la connexion et s'ils sont de classe $\mathcal{C}^n$ la connexion est de classe $\mathcal{C}^n$.

Soient X et Y deux éléments de $\mathfrak{D}^{(1)}(M)$ s'écrivant, sur U

$$X = \sum_i X^i \frac{\partial}{\partial x_i}, \quad Y = \sum_i Y^i \frac{\partial}{\partial x_i}. \quad \text{On obtient à l'aide de a) : Si } x \in U,$$

$$\begin{aligned} (*) \quad \nabla_X Y(x) &= \sum_k \left\{ \sum_i X^i(x) \frac{\partial Y^k}{\partial x_i}(x) + \sum_{i,j} \Gamma_{ij}^k(x) X^i(x) Y^j(x) \right\} \frac{\partial}{\partial x_k}(x) \\ &= \sum_k \left\{ X_x^k(Y^k) + \sum_{i,j} \Gamma_{ij}^k(x) X^i(x) Y^j(x) \right\} \frac{\partial}{\partial x_k}(x). \end{aligned}$$

Remarquons que cette expression ne dépend que de X au point x ce qui nous permet d'introduire la notation $\nabla_{X_x} Y$ pour $\nabla_X Y(x)$. Cette relation montre aussi que $\nabla_X Y(x)$ ne dépend que de Y sur une courbe c vérifiant $\dot{c}(0) = X_x$, ce qui nous amène à généraliser cette notion de connexion.

2.2. Dérivée covariante le long d'une application.

Si M et N sont deux variétés, N éventuellement à bord, et $f : N \longrightarrow M$ une application de classe $\mathcal{C}^h$, on appelle champ de vecteurs le long de f de classe $\mathcal{C}^n$, où n est un entier inférieur ou égal à h, une application $W : N \longrightarrow TM$ de classe $\mathcal{C}^n$ vérifiant, pour tout x de N, $W(x)$ appartient à $T_{f(x)}M$. Si, localement, $W(x) = \sum_i W^i(x) \frac{\partial}{\partial x_i}[f(x)]$, les applications W^i sont alors de classe $\mathcal{C}^n$.

Un cas particulier important est le suivant : Etant donné un champ de vecteurs X de classe $\mathcal{C}^r$ sur N on définit un champ le long de f, noté $df(X)$, par $df(X)(x) = df_x(X(x))$. Il est de classe $r \wedge (h-1)$.

Si v est un élément de $T_x N$ et Y dans $\mathfrak{D}^{(1)}(M)$, on a localement, d'après la relation (*) du 2.1.

$$\nabla_{df_x(v)} Y = \sum_k \left\{ v(Y^k \circ f) + \sum_{i,j} \Gamma_{ij}^k \{f(x)\} v(x_i \circ f) Y^j(f(x)) \right\} \frac{\partial}{\partial x_k} [f(x)]$$

Puisque cette expression ne dépend que des valeurs de Y "le long de f", il est naturel de la généraliser à un champ W de classe $\mathcal{C}^1$ le long de f en posant, si localement $W(x) = \sum_i W^i(x) \frac{\partial}{\partial x_i}[f(x)]$, si $v \in T_x N$:

$$\nabla_v^f W = \sum_k \left\{ v(W^k) + \sum_{i,j} \Gamma_{ij}^k \{f(x)\} v(x_i \circ f) W^j(x) \right\} \frac{\partial}{\partial x_k} [f(x)].$$

$\nabla_v^f W$ s'appelle la dérivée covariante de W le long de f dans la direction v .

Si X est un champ de vecteurs sur N on note alors $\nabla_X^f W$ le champ le long de f qui au point x de N est égal à $\nabla_{X_x}^f W$. Remarquons que s'il existe un élément $\tilde{W}$ de $\mathcal{D}^{(1)}(M)$ tel que, pour tout x de N , $W_x = \tilde{W}_{f(x)}$, (c'est toujours le cas si f est un plongement) on a, si $v \in T_x N$:

$$\nabla_v^f W = \nabla_{df_x(v)} \tilde{W}$$

Parfois, quand il n'y a pas d'ambiguïté, on écrira ∇_v au lieu de ∇_v^f .

2.3. Parallélisme.

DÉFINITION. Si I est un intervalle de $\mathbb{R}$ et $c : I \longrightarrow M$ une courbe différentiable, un champ W le long de c de classe $\mathcal{C}^1$ est dit *parallèle* le long de c si, pour tout t de I , $\nabla_{\frac{d}{dt}(t)} W$ est nul.

Localement, si $W(x) = \sum_i W^i(x) \frac{\partial}{\partial x_i} [c(x)]$, cette condition s'écrit :

$$(*) \quad \text{Pour tout } k, \quad \frac{dW^k}{dt} + \sum_{i,j} \Gamma_{ij}^k \circ c \cdot \frac{d(x_i \circ c)}{dt} \cdot W^j \text{ est nul sur } I.$$

2.3.1. PROPOSITION. Soit $c : [a, b] \longrightarrow M$ une courbe différentiable. Pour tout v de $T_{c(a)} M$ il existe un champ W^v le long de c de classe $\mathcal{C}^1$, et un seul, parallèle le long de c tel que $W_a^v = v$. L'application de $T_{c(a)} M$ sur $T_{c(b)} M$ qui à v associe W_b^v est un isomorphisme linéaire appelé *transport parallèle* de $c(a)$ à $c(b)$ le long de c .

Pour montrer cette proposition il suffit de raisonner localement. C'est alors une conséquence du fait que les équations différentielles (*) que doit satisfaire W^v sont linéaires du premier ordre.

Remarque : Si on note $\mathcal{C}_{a,b}$ l'isomorphisme introduit dans la proposition on peut démontrer que si Y est dans $\mathcal{D}^{(1)}(M)$,

$$\nabla_{\dot{c}(a)} Y = \lim_{t \rightarrow a} \frac{1}{t-a} \{ \mathcal{C}_{a,t}^{-1} (Y_{c(t)}) - Y_{c(a)} \}$$

retrouvant la définition intuitive que nous avons donnée dans l'introduction.

2.4. Connexions riemanniennes

DÉFINITION. Soient k et n deux entiers, $k \geq 3$, $1 \leq n \leq k-1$, M une variété de classe $\mathcal{C}^k$ portant une structure riemannienne g de classe $\mathcal{C}^n$. Une connexion ∇ sur M est dite *compatible* avec g si

- a) ∇ est de classe $\mathcal{C}^{n-1}$
- b) Si X, Y sont dans $\mathcal{D}^1(M)$, $\nabla_X Y - \nabla_Y X = [X, Y]$
- c) Le produit scalaire est conservé par transport parallèle.

La relation b) se traduit localement par $\Gamma_{ij}^k = \Gamma_{ji}^k$ pour tout i, j, k . La relation c) est équivalente à celle apparaissant dans le lemme suivant, et c'est sous cette forme qu'on l'utilise.

LEMME. (Identité de Ricci)

Si ∇ est compatible avec g , pour tout X, Y, Z dans $\mathcal{D}^{(1)}(M)$,

$$X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$$

Démonstration. Montrons cette relation au point x de M . Si X_x est nul elle est claire, sinon soit c une courbe $\mathcal{C}^1$ partant de x avec la vitesse X_x et $P_1, \dots, P_m$ des champs le long de c obtenus en transportant parallèlement une base orthonormale de $T_x M$ grâce à la proposition 2.3.1.

Si $Y_{c(t)} = \sum \alpha_i(t) P_i(t)$ et $Z_{c(t)} = \sum \beta_i(t) P_i(t)$ on a

$$\langle Y, Z \rangle_{c(t)} = \sum \alpha_i(t) \beta_i(t) \text{ d'où}$$

$$X_x \langle Y, Z \rangle = \dot{\langle Y, Z \rangle}(0) = \sum_i \{ \alpha_i'(0) \beta_i(0) + \alpha_i(0) \beta_i'(0) \}.$$

$$\begin{aligned} \text{D'autre part } \nabla_{X_x} Y &= \frac{\nabla^c}{dt} (\sum \alpha_i P_i)(0) = \sum \alpha_i'(0) P_i(0) + \sum \alpha_i(0) \frac{\nabla^c}{dt} P_i(0) \\ &= \sum \alpha_i'(0) P_i(0) \end{aligned}$$

$$\text{d'où } \langle \nabla_{X_x} Y, Z \rangle = \sum \alpha_i'(0) \beta_i(0) \text{ et de même } \langle Y, \nabla_{X_x} Z \rangle = \sum \alpha_i(0) \beta_i'(0).$$

Si $f : N \rightarrow M$ est une application suffisamment différentiable on peut vérifier que

b) Si X et Y sont dans $\mathcal{D}^{(1)}(N)$,

$$\nabla_X^f df(Y) - \nabla_Y^f df(X) = df[X, Y]$$

c') Si v est un élément de TN et W_1, W_2 deux champs le long de f de classe $\mathcal{C}$

$$\langle \nabla_v^f W_1, W_2 \rangle + \langle W_1, \nabla_v^f W_2 \rangle = v \langle W_1, W_2 \rangle.$$

THÉOREME. Pour toute structure riemannienne g sur M il existe une unique connexion ∇ compatible avec g appelée connexion riemannienne associée.

La démonstration se fait à partir du lemme précédant. On obtient en particulier que localement, si (g^{ij}) est la matrice inverse de (g_{ij}) :

$$\Gamma_{ij}^k = \frac{1}{2} \sum_l \left\{ \frac{\partial g_{il}}{\partial x_j} + \frac{\partial g_{jl}}{\partial x_i} - \frac{\partial g_{ij}}{\partial x_l} \right\} g^{lk}$$

Dorénavant, si sur une variété riemannienne on emploie une connexion, c'est de celle ci qu'il s'agit.

Exemple. Sur $\mathbb{R}^m$, $ds^2 = \sum (dx_i)^2$ est la structure riemannienne canonique et la connexion ∇ associée vérifie : Γ_{ij}^k est identiquement nul. Si M est une sous-variété de $\mathbb{R}^m$, puisque $T_x M$ est plongé dans $T_x \mathbb{R}^m$ on peut lui restreindre la structure canonique. La connexion $\tilde{\nabla}$ associée est alors donnée par : Si X, Y sont dans $\mathfrak{D}^{(1)}(M)$, $\tilde{\nabla}_X Y$ est la projection orthogonale de $\nabla_X Y$ sur $T_x M$.

3. Géodésiques et exponentielle

Fixons une variété riemannienne M de classe $\mathcal{C}^n$, $n \geq 2$.

DÉFINITION. Une courbe $c : I \rightarrow M$, de classe $\mathcal{C}^2$ est une géodésique si son vecteur vitesse est parallèle le long de celle ci, c'est-à-dire si $\frac{\nabla_c^c}{dt} \dot{c}$ est nul sur I .

Une géodésique est donc une courbe qui vérifie localement l'équation :

$$\frac{d^2 x_k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{dx_i}{dt} \frac{dx_j}{dt} = 0 \quad (1 \leq k \leq m).$$

(Il suffit d'appliquer la relation $(*)$ du 2.3 et de remarquer que

$$\dot{c}(t) = \sum_i \dot{c}(t)(x_i) \frac{\partial}{\partial x_i} [c(t)] = \sum_i \frac{d(x_i \circ c)}{dt} \frac{\partial}{\partial x_i} [c(t)]$$

3.1. PROPOSITION. Pour tout v de TM , il existe un voisinage $\mathcal{V}$ de v dans TM et $\varepsilon > 0$ tels que :

Si w appartient à $\mathcal{V}$, il existe une unique géodésique $c(t; w)$ définie sur un intervalle contenant $]-\varepsilon, \varepsilon[$, vérifiant $c(0; w) = x$ si $w \in T_x M$ et $\dot{c}(0; w) = w$. De plus l'application $c :]-\varepsilon, \varepsilon[\times TM \rightarrow M$ est de classe $\mathcal{C}^{n-1}$.

Démonstration. Cela résulte immédiatement du fait que, pour que c soit une géodésique il faut et il suffit que la courbe $t \longrightarrow (c(t), \dot{c}(t))$ soit, localement, solution du système :

$$\frac{\partial x_i}{\partial t} = y_i, \quad \frac{\partial y_k}{\partial t} = - \sum_{i,j} \Gamma_{ij}^k y_i y_j \quad \text{où } \Gamma_{ij}^k \text{ est de classe } \mathcal{C}^{n-1}.$$

3.2. COROLLAIRE. Si x est dans M , il existe un voisinage U de x dans M et $\varepsilon > 0$ tels que : Si y appartient à U , $v \in T_y M$ et $\|v\|_y < \varepsilon$, $c(t;v)$ est défini sur $] -2, 2[$.

Démonstration. Appliquons la proposition au vecteur v de $T_x M$ égal au vecteur nul de $T_x M$. Il existe U , ε_1 , ε_2 tels que, si $y \in U$, $v \in T_y M$, $\|v\|_y < \varepsilon_1$, $c(t;v)$ est défini sur $]-\varepsilon_2, \varepsilon_2[$. Puisque $c(\varepsilon_2 t; \frac{v}{2})$ est une géodésique γ vérifiant $\gamma(0) = y$, $\dot{\gamma}(0) = v$ et définie sur $]-2, 2[$ le résultat est clair.

DÉFINITION. Soit v un élément de $T_x M$, si la géodésique $t \longmapsto c(t,v)$ introduite dans la proposition 3.1 est définie sur un intervalle ouvert contenant $[0,1]$, on note $\exp_x v$ l'élément $c(1;v)$ de M .

Considérons la fonction $\exp : TM \longrightarrow M \times M$ définie par, si v est dans $T_x M$ et si $\exp_x v$ a un sens, $\exp v = (x, \exp_x v)$. D'après le corollaire la fonction $\exp$ est définie sur un voisinage de $\{X \in T_x M, x \in M, \|X\|_x = 0\}$ et y est de classe $\mathcal{C}^{n-1}$. En fait le domaine de $\exp$ est ouvert et l'application exponentielle y est $\mathcal{C}^{n-1}$.

Remarque. Si c est une géodésique définie sur un intervalle I contenant 0,

$$c(t) = \exp_{c(0)} t \dot{c}(0) \quad \text{pour tout } t \in I.$$

La norme du vecteur vitesse est constante, donc la longueur de c sur $[a,b]$ est égale à $\|\dot{c}(a)\|(b-a)$.

4. Géodésiques et distance riemannienne.

Nous voulons montrer l'existence de voisinages où les géodésiques réalisent la distance. Montrons d'abord le théorème suivant.

4.1. THÉORÈME. Pour tout x de M il existe un voisinage W de x et $\varepsilon > 0$ tels que :

- Si y appartient à W , $\exp_y$ est un difféomorphisme de classe $\mathcal{C}^{n-1}$ de $\{X \in T_y M, \|X\|_y < \varepsilon\}$ sur un ouvert contenant W .
- Deux points y_1, y_2 de W peuvent être joints par une unique géodésique c de longueur inférieure à ε .

Remarque. Chaque fois que l'on parle de géodésique unique, on sous entend à un changement affine de paramètre près.

Démonstration. Il est facile de voir que l'application exponentielle $\exp : TM \longrightarrow M^2$ est régulière au point de TM égal au vecteur nul de $T_x M$. Il existe donc un ouvert U de M contenant x et $\varepsilon > 0$ tels que $\exp$ soit un difféomorphisme de $\bigcup_{y \in U} \{X \in T_y M, \|X\|_y < \varepsilon\}$ sur un ouvert $\mathfrak{D}$.

Si W est un voisinage de x tel que $W \times W$ soit contenu dans $\mathfrak{D}$, il est clair que a. est vérifié. Soient y_1 et y_2 deux éléments de W . Puisque W est contenu dans l'image de $\{X \in T_y M, \|X\|_y < \varepsilon\}$ par $\exp_{y_1}$, il existe un unique élément v de cette boule tel que $\exp_{y_1} v = y_2$. L'unique géodésique cherchée est alors $c(t) = \exp_{y_1} tv$, $t \in [0, 1]$.

On peut préciser ce théorème en montrant que tout point de M admet un voisinage V tel que deux points de V peuvent être joints par une unique géodésique contenue dans V .

(H) Fixons un point x de M et $\varepsilon > 0$ tel que $\exp_x$ soit un difféomorphisme de $\{v \in T_x M, \|v\|_x < \varepsilon\}$ sur son image.

L'application exponentielle n'est pas en général une isométrie mais elle conserve l'orthogonalité radiale comme le montre le lemme suivant :

4.2. LEMME. Sous (H), pour δ inférieur à ε , les géodésiques issues de x sont orthogonales aux surfaces $\exp_x \{v \in T_x M, \|v\|_x = \delta\}$.

Démonstration. Soit $t \longmapsto X(t)$ une courbe dans $T_x M$ telle que, pour tout t , $\|X(t)\|$ Si $f(r, t) = \exp_{x,r} X(t)$, nous devons montrer que le produit scalaire des deux champs le long de f , $df(\frac{\partial}{\partial r})$ et $df(\frac{\partial}{\partial t})$ est nul. En utilisant les relations b' et c' du 2.3, on peut écrire que

$$\begin{aligned} \nabla_{\frac{\partial}{\partial r}}^f df(\frac{\partial}{\partial t}) - \nabla_{\frac{\partial}{\partial t}}^f df(\frac{\partial}{\partial r}) &= 0 \text{ et} \\ \frac{\partial}{\partial r} \langle df(\frac{\partial}{\partial r}), df(\frac{\partial}{\partial t}) \rangle &= \langle \nabla_{\frac{\partial}{\partial r}}^f df(\frac{\partial}{\partial r}), df(\frac{\partial}{\partial t}) \rangle + \langle df(\frac{\partial}{\partial r}), \nabla_{\frac{\partial}{\partial r}}^f df(\frac{\partial}{\partial t}) \rangle \end{aligned}$$

Puisque $r \longmapsto f(r, t)$ est une géodésique le premier terme de cette somme est nul alors que le second est égal à

$$\langle df(\frac{\partial}{\partial r}), \nabla_{\frac{\partial}{\partial t}}^f df(\frac{\partial}{\partial r}) \rangle = \frac{1}{2} \frac{\partial}{\partial t} \langle df(\frac{\partial}{\partial r}), df(\frac{\partial}{\partial r}) \rangle = \frac{1}{2} \frac{\partial}{\partial t} \|X(t)\|^2 = 0$$

Donc $\langle df(\frac{\partial}{\partial r}), df(\frac{\partial}{\partial t}) \rangle$ est indépendant de r . Mais $df(\frac{\partial}{\partial t})(0, t)$ est égal à zéro donc ce produit scalaire est nul.

Considérons une courbe $\omega : [a, b] \longrightarrow M$, dont l'image est contenue dans $\exp_x\{v \in T_x M, 0 < \|v\|_x < \varepsilon\}$. On peut l'écrire sous la forme $\omega(t) = \exp_x r(t)X(t)$ où $\|X(t)\|_x$ est égal à un, $r(t) \geq 0$, de façon unique.

4.3. LEMME. La longueur de ω , $\int_b^a \|\dot{\omega}(t)\| dt$ est supérieure ou égale à $|r(b) - r(a)|$. Il n'y a égalité que si X est constant et r monotone.

Démonstration. Remarquons que si $f(r, t) = \exp_x r X(t)$, $\dot{\omega}(t) = r'(t) df(\frac{\partial}{\partial r}) + df(\frac{\partial}{\partial t})$ et d'après le lemme $\|\dot{\omega}(t)\|^2 = |r'(t)|^2 + \|df(\frac{\partial}{\partial t})\|^2$. Ce qui donne le résultat.

On dit qu'une courbe $c : [0, 1] \longrightarrow M$ peut être paramétrée par sa longueur si il existe une application $\phi : [0, 1] \rightarrow a, b$ continue croissante et une courbe $\omega : [a, b] \rightarrow M$ telles que :

- 1) $\|\dot{\omega}(t)\| = 1$, pour tout $t \in [a, b]$
- 2) $c(t) = \omega(\phi(t))$, pour tout $t \in [0, 1]$.

On dit alors que ω est "la courbe c paramétrée par sa longueur".

On déduit immédiatement des lemmes précédents :

4.4. THÉORÈME. Pour chaque point de M il existe $\varepsilon > 0$ et un voisinage V de ce point tels que chaque couple de points de V peut être joint par une unique géodésique de longueur L inférieure à ε . La longueur L' de toute courbe joignant ces deux points vérifie $L \geq L'$. De plus $L = L'$ si et seulement si cette courbe peut être paramétrée par sa longueur et est alors une géodésique.

4.5. COROLLAIRE. Si une courbe réalise la distance entre ses extrémités elle peut être paramétrée par sa longueur et est alors une géodésique.

En effet toutes ces propriétés sont locales et il suffit alors d'appliquer le théorème 4.4.

Notons aussi que, d'après le théorème, la distance riemannienne introduite est bien une distance si M est connexe et que la topologie associée est la topologie de M .

BIBLIOGRAPHIE

- [1] J. Cheeger, D.G. Ebin, *Comparison theorems in riemannian geometry*, North Holland Publishing Company, 1975
- [2] J. Milnor, *Morse theory*, Annals of Math. Studies. Princeton University Press. 1963.

CHAPITRE 4

PREMIÈRES MAJORATIONS DE LA DENSITÉ D'UNE DIFFUSION SUR $\mathbb{R}^m$ MÉTHODE DE LA PARAMETRIX

par

Paolo BALDI

Nous verrons dans le chapitre 5 (L. Elie) que l'étude du comportement en temps petit et pour des points proches de la densité de transition d'une diffusion sur une variété est un problème local et qu'à l'aide d'une carte, nous pouvons transporter cette étude sur $\mathbb{R}^m$.

Nous allons donc dans cet exposé rechercher des majorations de la densité de transition d'une diffusion sur $\mathbb{R}^m$.

1. La solution fondamentale

Soit $D = L_t - \frac{\partial}{\partial t} = \frac{1}{2} \sum_{i,j=1}^m a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(t,x) \frac{\partial}{\partial x_i} - \frac{\partial}{\partial t}$ un opérateur semi-elliptique défini sur $U = [0,T] \times \mathbb{R}^m$.

On appelle solution fondamentale de D (au sens de Cauchy), une fonction $p(\tau, t, x, \xi)$ définie sur $[0,T]^2 \times \mathbb{R}^m \times \mathbb{R}^m$ avec $0 \leq \tau < t \leq T$ telle que :

i) pour tout (τ, ξ) appartenant à U , p vérifie $D_p = 0$ en tant que fonction de (t, x) avec x appartenant à $\mathbb{R}^m$, $\tau < t \leq T$.

ii) pour toute fonction f continue bornée sur $\mathbb{R}^m$ on ait :

$$\lim_{t \searrow \tau} \int_{\mathbb{R}^m} p(\tau, t, x, \xi) f(\xi) d\xi = f(x)$$

On va d'abord construire une solution fondamentale de D , sous des hypothèses sur les coefficients que l'on précisera, par un procédé analytique (méthode de la parametrix) qui nous fournira en même temps une majoration explicite de cette solution. On identifiera ensuite la solution fondamentale ainsi obtenue avec la densité de transition de la diffusion associée à L_t , dont on démontrera en même temps l'existence.

L'idée de cette méthode de construction est très ancienne et remonte à E.E. Levi (1907). Ici on suivra de près l'exposition de A. Friedman [2] (chap. I et IX) simplifiée là où il a été possible et adaptée à nos besoins spécifiques. Pour le dernier paragraphe, les sources sont l'exposé d'Azencott sur les diffusion (chap 2) et l'article de R. Azencott [1].

2. La méthode de la parametrix.

Soit D_z l'opérateur obtenu à partir de D en "gélant" en z la variable spatiale et dont les coefficients, donc, ne dépendent que du temps.

$$D_z f(t, x) = \frac{1}{2} \sum_{i,j=1}^m a_{ij}(t, z) \frac{\partial^2 f}{\partial x_i \partial x_j}(t, x) + \sum_{i=1}^m b_i(t, z) \frac{\partial f}{\partial x_i}(t, x) - \frac{\partial f}{\partial t}(t, x).$$

Soit $W^z(\tau, t, x, \xi)$ une solution fondamentale pour D_z et posons $q(\tau, t, x, \xi) = W(\tau, t, x, \xi)$; q s'appelle la parametrix et l'idée est de chercher une solution fondamentale de la forme :

$$(1) \quad p(\tau, t, x, \xi) = q(\tau, t, x, \xi) + \int_{\tau}^t d\sigma \int_{\mathbb{R}^m} q(\tau, \sigma, x, y) \phi(\sigma, t, y, \xi) dy$$

où ϕ est une fonction à déterminer. Posons :

$$q * \phi(\tau, t, x, \xi) = \int_{\tau}^t d\sigma \int_{\mathbb{R}^m} q(\tau, \sigma, x, y) \phi(\sigma, t, y, \xi) dy.$$

Si on pouvait effectuer toutes les dérivations sous le signe somme, on aurait :

$$\begin{aligned} 0 &= Dp(\tau, t, x, \xi) = Dq(\tau, t, x, \xi) + D(q * \phi)(\tau, t, x, \xi) \\ &= Dq(\tau, t, x, \xi) + (Dq) * \phi(\tau, t, x, \xi) - \phi(\tau, t, x, \xi) \end{aligned}$$

où le dernier terme vient de la dérivation de l'intégrale par rapport à la borne d'intégration. ϕ doit donc satisfaire à l'équation :

$$(2) \quad \phi = Dq + (Dq) * \phi,$$

et on est donc amené à chercher une fonction de la forme :

$$(3) \quad \phi = \sum_{n=1}^{\infty} (Dq)^{*n}$$

Il faut donc :

- a) calculer q ;
- b) prouver que la série de (3) converge ;
- c) prouver que l'on peut effectuer les dérivations sous le signe somme et donc que p définie par (1) est une solution fondamentale.
- d) montrer que la solution fondamentale ainsi obtenue est la densité de transition de la diffusion associée à L_t .

Dans cette exposition on se contentera de prouver a), b) et d). La démonstration du point c) est en effet assez laborieuse et technique, la difficulté résidant surtout dans le fait que les q et ϕ ont des singularités qui compliquent les opérations de dérivation sous le signe somme. Pour une preuve complète, on pourra se reporter à A. Friedman [2] (chap. I et IX).

On fera dorénavant les hypothèses suivantes :

A_1 : les a_{ij} et b_i sont continues sur U . Les a_{ij} sont uniformément continues en t pour $(t, x) \in U$.

A_2 : les coefficients sont bornés sur U et höldériens d'exposant $\alpha > 0$ en x , uniformément pour $(t, x) \in U$; c'est-à-dire qu'il existe des constantes positives α, M et k telles que :

$$\begin{aligned} |b_i(t, x)| &\leq M, \quad |a_{ij}(t, x)| \leq M, \quad \text{pour tout } (t, x) \text{ de } U, \quad i, j=1, \dots, m. \\ |a_{ij}(t, x) - a_{ij}(t, y)| &\leq k \|x-y\|^\alpha, \\ |b_i(t, x) - b_i(t, y)| &\leq k \|x-y\|^\alpha, \quad \text{pour tout } x, y \text{ de } \mathbb{R}^m, \quad t \text{ appartenant à } [0, T], \\ &\quad i, j=1, \dots, m. \end{aligned}$$

A_3 : les matrices $(a_{ij}(t, x))$ sont symétriques et définies positives uniformément en (t, x) appartenant à U . On supposera donc que toutes les valeurs propres sont comprises entre deux constantes μ_1 et μ_2 avec

$$0 < \mu_1 \leq \mu_2 \leq M.$$

3. Calcul de la paramétrix.

Il s'agit de trouver une fonction $q(\tau, t, x, \xi)$, (x, ξ) appartenant à $\mathbb{R}^m \times \mathbb{R}^m$, $0 \leq \tau < t \leq T$, telle que :

$$(4) \quad \frac{\partial q}{\partial t} = \frac{1}{2} \sum_{i,j=1}^m a_{ij}(t, \xi) \frac{\partial^2 q}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(t, \xi) \frac{\partial q}{\partial x_i}$$

$$(5) \quad \lim_{t \searrow \tau} q(\tau, t, x, \xi) = \delta_x(\xi).$$

Si on calcule formellement la transformée de Fourier des deux côtés par rapport à x , on trouve :

$$\frac{\partial}{\partial t} \hat{q}(\tau, t, \theta, \xi) = \hat{q}(\tau, t, \theta, \xi) \left\{ -\frac{1}{2} \sum_{h,k=1}^m a_{hk}(t, \xi) \theta_h \theta_k - i \sum_{h=1}^m b_h(t, \xi) \theta_h \right\}$$

Il s'agit d'une équation ordinaire du premier ordre en t dont l'intégrale est donnée par :

$$\hat{q}(\tau, t, \theta, \xi) = c \exp \int_{\tau}^t \left[-\frac{1}{2} \sum_{h,k=1}^m a_{hk}(s, \xi) \theta_h \theta_k - i \sum_{h=1}^m b_h(s, \xi) \theta_h \right] ds$$

et la condition (5) entraîne que $c = \exp i\theta^* \xi$.

$$\text{Posons } A_{hk}(\tau, t, y) = \int_{\tau}^t a_{hk}(s, y) ds$$

$$B_h(\tau, t, y) = \int_{\tau}^t b_h(s, y) ds$$

$$A(\tau, t, y) = \{A_{hk}(\tau, t, y)\}_{h,k}, \quad B(\tau, t, y) = \{B_h(\tau, t, y)\}_h$$

On écrira A et B pour $A(\tau, t, y)$ et $B(\tau, t, y)$ respectivement pour alléger la notation dès qu'il n'y aura pas de confusion possible.

Par la transformation de Fourier inverse, on a alors :

$$q(\tau, t, x, \xi) = (2\pi)^{-m} \int_{\mathbb{R}^m} \exp(-i\theta^* x) \exp(i\theta^* \xi) \exp\left[-\frac{1}{2}\theta^* A \theta - i\theta^* B\right] d\theta$$

et un calcul standard donne :

$$q(\tau, t, x, \xi) = \frac{1}{(2\pi)^{\frac{m}{2}} (\det A)^{1/2}} \exp\left[-\frac{1}{2}(\xi - x - B)^* A^{-1} (\xi - x - B)\right]$$

Et comme en tant que fonction de x , q est dans L^1 ainsi que ses dérivées jusqu'à l'ordre 2, la transformation de Fourier était justifiée et il s'agit effectivement d'une solution de (4) et (5).

4. Calcul de la solution fondamentale.

4.1. Convergence de la série. On va d'abord montrer que la série $\sum_{n=1}^{\infty} (Dq)^* n$ converge et pour cela on commence par majorer chaque terme et on conclura grâce à un lemme technique.

On a pour $t > \tau$,

$$\begin{aligned} Dq(\tau, t, x, \xi) &= (D - D_{\xi})(q(\tau, t, x, \xi)) = \\ &= \frac{1}{2} \sum_{i,j=1}^m [a_{ij}(t, x) - a_{ij}(t, \xi)] \frac{\partial^2 q}{\partial x_i \partial x_j}(\tau, t, x, \xi) + \sum_{i=1}^m [b_i(t, x) - b_i(t, \xi)] \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \end{aligned}$$

On commence par majorer les dérivées de q :

$$\frac{\partial q}{\partial x_i}(\tau, t, x, \xi) = \frac{1}{(2\pi)^{m/2}} \frac{1}{(\det A(\tau, t, \xi))^{1/2}} [A^{-1}(\tau, t, \xi) [\xi - x - B(\tau, t, \xi)]]$$

$$\times \exp \left[-\frac{1}{2} (\xi - x - B(\tau, t, \xi))^* A^{-1}(\tau, t, \xi) (\xi - x - B(\tau, t, \xi)) \right].$$

Comme $\frac{1}{\mu_2(t-\tau)} = \frac{\lambda}{t-\tau}$ et $\frac{1}{\mu_1(t-\tau)}$ sont respectivement un minorant et un majorant de l'ensemble des valeurs propres de $A^{-1}(\tau, t, \xi)$, on obtient :

$$\begin{aligned} \left| \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \right| &\leq \frac{1}{(2\pi)^{m/2}} \frac{1}{[\mu_1(t-\tau)]^{m/2}} \frac{1}{\mu_1(t-\tau)} \\ &\times \|\xi - x - B(\tau, t, \xi)\| \exp \left\{ -\frac{\lambda}{2(t-\tau)} \|\xi - x - B(\tau, t, \xi)\|^2 \right\} \\ \|x\| &\text{ désignant la norme } \left(\sum_{k=1}^m |x_k|^2 \right)^{1/2} \end{aligned}$$

La fonction $y \longrightarrow y^\alpha \exp -\varepsilon y^2$ étant bornée sur $\mathbb{R}^+$ si (ε, α) appartient à $\mathbb{R}^+ \times \mathbb{R}^+$, on en déduit, en posant $\lambda_\varepsilon = \lambda - 2\varepsilon$, qu'il existe une constante c' de $\mathbb{R}^+$ telle que :

$$\left| \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \right| \leq \frac{c'}{(t-\tau)^{\frac{m}{2} + \frac{1}{2}}} \exp \left\{ -\frac{\lambda_\varepsilon}{2(t-\tau)} \|\xi - x - B(\tau, t, \xi)\|^2 \right\}$$

Mais on a :

$$\begin{aligned} \|\xi - x - B(\tau, t, \xi)\|^2 &\geq \|\xi - x\|^2 - 2 \sum_{k=1}^m B_k(\tau, t, \xi) (\xi - x)_k \\ &\geq \|\xi - x\|^2 - 2M(t-\tau) \sum_{k=1}^m |\xi - x|_k \\ &\geq \|\xi - x\|^2 - M'(t-\tau) \|\xi - x\| \end{aligned}$$

avec M' constante positive.

D'où :

$$\left| \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \right| \leq \frac{c'}{(t-\tau)^{\frac{m}{2} + \frac{1}{2}}} \exp \left\{ -\frac{\lambda_\varepsilon}{2(t-\tau)} \|\xi - x\|^2 + \frac{M'\lambda_\varepsilon}{2} \|\xi - x\| \right\}$$

$$\begin{aligned} &\frac{c''}{(t-\tau)^{\frac{m}{2} + \frac{1}{2}}} \exp \left\{ \frac{-\lambda_\varepsilon}{2(t-\tau)} \|\xi - x\|^2 \right\} \mathbb{1}_{\{\|\xi - x\| \leq \eta\}} \\ &+ \frac{c'}{(t-\tau)^{\frac{m}{2} + \frac{1}{2}}} \exp \left\{ -\frac{\lambda_\varepsilon \|\xi - x\|^2}{2(t-\tau)} \left(1 - \frac{M'(t-\tau)}{\eta} \right) \right\} \mathbb{1}_{\{\|\xi - x\| > \eta\}} \end{aligned}$$

où $\eta \in \mathbb{R}^+$. En choisissant η suffisamment grand, on obtient si $\lambda^* = \lambda - 4\varepsilon$,

$$\left| \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \right| \leq \frac{c}{(t-\tau)^{\frac{m}{2}+1}} \exp \left[- \frac{\lambda \|\xi-x\|^2}{2(t-\tau)} \right],$$

où $c = \sup(c', c'')$.

Par le même type d'arguments on a :

$$\begin{aligned} |b_i(t, x) - b_i(t, \xi)| \left| \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \right| &\leq k \|\xi-x\|^\alpha \left| \frac{\partial q}{\partial x_i}(\tau, t, x, \xi) \right| \\ &\leq \frac{c}{(t-\tau)^{\frac{m}{2}+1-\frac{\alpha}{2}}} \exp \left[- \frac{\lambda^* \|\xi-x\|^2}{2(t-\tau)} \right]. \end{aligned}$$

Et pour les dérivées secondes :

$$\begin{aligned} \frac{\partial^2 q}{\partial x_i \partial x_j}(\tau, t, x, \xi) &= \frac{1}{(2\pi)^{\frac{m}{2}} (\det A)^{1/2}} \{ (A^{-1})_{ij} + \\ &+ [A^{-1}(\xi-x-B)]_i [A^{-1}(\xi-x-B)]_j \} \exp \left\{ - \frac{1}{2} (\xi-x-B)^* A^{-1} (\xi-x-B) \right\} \end{aligned}$$

d'où

$$|a_{ij}(t, x) - a_{ij}(t, \xi)| \left| \frac{\partial^2 q}{\partial x_i \partial x_j}(\tau, t, x, \xi) \right| \leq \frac{c}{(t-\tau)^{\frac{m}{2}+1-\frac{\alpha}{2}}} \exp \left[- \frac{\lambda^* \|\xi-x\|^2}{2(t-\tau)} \right]$$

D'où finalement :

$$(6) \left| (D-D_\xi)q(\tau, t, x, \xi) \right| \leq \frac{c}{(t-\tau)^{\frac{m}{2}+1-\frac{\alpha}{2}}} \exp \left[- \frac{\lambda^* \|\xi-x\|^2}{2(t-\tau)} \right]$$

Remarque : un regard plus attentif aux majorations qu'on vient d'établir montre que dans l'inégalité ci-dessus, c ne dépend que de $k, M, \alpha, \mu_1, \mu_2$ et T ; λ^* peut être choisi aussi proche que l'on veut de $\lambda = \frac{1}{\mu_2}$ au prix de faire croître c . Cette remarque étant utile dans le théorème sur les suites d'opérateurs que l'on verra à la fin.

LEMME. Si on pose $N_\gamma(\tau, t, x, \xi) = (t-\tau)^\gamma \exp \left[- \frac{h \|\xi-x\|^2}{2(t-\tau)} \right]$ avec $0 < \gamma < \frac{m}{2} + 1$, $0 < \beta < \frac{m}{2} + 1$ et $h > 0$, alors :

$$N_\gamma * N_\beta(\tau, t, x, \xi) = \left(\frac{2\pi}{h} \right)^{\frac{m}{2}} \frac{\Gamma(\frac{m}{2}-\gamma+1) \Gamma(\frac{m}{2}-\beta+1)}{\Gamma(m-\gamma-\beta+2)}$$

$$(t-\tau)^{\frac{m}{2}-\alpha-\beta+1} \times \exp \left[-\frac{h\|\xi-x\|^2}{2(t-\tau)} \right].$$

Preuve : par la propriété de semi-groupe des densités gaussiennes, on a :

$$\int_{\mathbb{R}^m} (t-\sigma)^{-\frac{m}{2}} (\sigma-\tau)^{-\frac{m}{2}} \exp \left[-\frac{h\|y-x\|^2}{2(t-\sigma)} \right] \exp \left[-\frac{h\|\xi-y\|^2}{2(\sigma-\tau)} \right] dy = \left(\frac{2\pi}{h} \right)^{\frac{m}{2}} (t-\tau)^{-\frac{m}{2}} \exp \left[-\frac{h\|\xi-x\|^2}{2(t-\tau)} \right]$$

D'où :

$$N_Y * N_\beta = \left(\frac{2\pi}{h} \right)^{\frac{m}{2}} (t-\tau)^{-\frac{m}{2}} \exp \left[-\frac{h\|\xi-x\|^2}{2(t-\tau)} \right] \int_\tau^t (t-\sigma)^{\frac{m}{2}-\gamma} (\sigma-\tau)^{\frac{m}{2}-\beta} d\sigma.$$

Mais cette dernière intégrale est égale à :

$$(t-\tau)^{m+1-\gamma-\beta} \int_0^1 (1-\rho)^{\frac{m}{2}-\gamma} \rho^{\frac{m}{2}-\beta} d\rho = (t-\tau)^{m+1-\gamma-\beta} \frac{\Gamma(\frac{m}{2}-\gamma+1) \Gamma(\frac{m}{2}-\beta+1)}{\Gamma(m+2-\gamma-\beta)}$$

ce qui prouve le lemme.

On peut alors montrer par récurrence grâce à cette formule que :

$$|(Dq)^{*n}(\tau, t, x, \xi)| \leq \frac{c n \Gamma(\frac{\alpha}{2}) (t-\tau)^{\frac{n\alpha}{2}}}{\Gamma(\frac{n\alpha}{2})} \frac{1}{(t-\tau)^{\frac{m}{2}+1}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right].$$

La série (3) est alors absolument convergente et sa somme ϕ vérifie :

$$|\phi(\tau, t, x, \xi)| \leq \frac{c}{(t-\tau)^{\frac{m}{2}+1-\frac{\alpha}{2}}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right].$$

De plus ϕ satisfait à l'équation 2.

4.2. THÉORÈME. Sous les hypothèses A_1, A_2, A_3 , il existe une solution fondamentale p qui vérifie pour tous (τ, t, x, ξ) de $[0, T]^2 \times \mathbb{R}^m \times \mathbb{R}^m$, $\tau \leq t$

$$(7) \quad p(\tau, t, x, \xi) \leq \frac{c}{(t-\tau)^{\frac{m}{2}}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right]$$

$$(8) \quad \left| \frac{\partial p}{\partial x_i}(\tau, t, x, \xi) \right| \leq \frac{c}{(t-\tau)^{\frac{m}{2}+\frac{1}{2}}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right]$$

$$(9) \quad \left| \frac{\partial^2 p}{\partial x_i \partial x_j}(\tau, t, x, \xi) \right| \leq \frac{c}{(t-\tau)^{\frac{m}{2}+1}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right]$$

De plus, lorsque t_0 tend vers zéro uniformément en $\tau \leq t \leq t_0$, x appartenant à $\mathbb{R}^m$, on a : (cf. S. Molchanov, Russian Math. Surveys, 30, 1975)

$$(10) \quad p(\tau, t, x, x) \sim \frac{1}{(2\pi)^{\frac{m}{2}} [\det \int_{\tau}^t a(s, x) ds]^{1/2}}.$$

Donnons quelques indications sur la démonstration.

Nous posons $p = q + q * \phi$ et renvoyons à A. Friedman [2] pour montrer que les dérivations sous le signe somme sont possibles ; c'est-à-dire que :

$$\frac{\partial p}{\partial x_i} = \frac{\partial q}{\partial x_i} + \frac{\partial q}{\partial x_i} * \phi, \quad \frac{\partial p}{\partial t} = \frac{\partial q}{\partial t} + \frac{\partial q}{\partial t} * \phi - \phi.$$

$$\frac{\partial^2 p}{\partial x_i \partial x_j} = \frac{\partial^2 q}{\partial x_i \partial x_j} + \frac{\partial^2 q}{\partial x_i \partial x_j} * \phi.$$

La fonction p est alors une solution fondamentale ; en effet, $Dp = Dq + Dq * \phi - \phi$, cette dernière expression étant nulle grâce à (2).

En appliquant à nouveau le lemme on obtient :

$$|q * \phi(\tau, t, x, \xi)| \leq \frac{c(t-\tau)^{\frac{\alpha}{2}}}{(t-\tau)^{\frac{m}{2}}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right]$$

Donc on a :

$$\begin{aligned} |p(\tau, t, x, \xi)| &\leq \frac{1}{(2\pi)^{\frac{m}{2}} (\det A)^{1/2}} \exp \left[-\frac{1}{2}(\xi-x-B)^* A^{-1}(\xi-x-B) \right] \\ &\quad + \frac{c(t-\tau)^{\frac{\alpha}{2}}}{(t-\tau)^{\frac{m}{2}}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right] \end{aligned}$$

ce qui donne :

$$|p(\tau, t, x, \xi)| \leq \frac{c}{(t-\tau)^{\frac{m}{2}}} \exp \left[-\lambda^* \frac{\|\xi-x\|^2}{2(t-\tau)} \right]$$

Par contre, si $x = \xi$,

$$p(\tau, t, x, x) = \frac{1}{(2\pi)^{\frac{m}{2}} [\det A(\tau, t, x)]^{1/2}} \exp\left[-\frac{1}{2} B(\tau, t, x)^* A^{-1} B(\tau, t, x)\right]$$

$$+ q * \phi(\tau, t, x, x)$$

$$\text{Et comme } q * \phi(\tau, t, x, x) \leq \frac{c(t-\tau)^{\frac{\alpha}{2}}}{(t-\tau)^{\frac{m}{2}}}, \text{ alors que}$$

$$[\det A(\tau, t, x)]^{1/2} \sim [\det a_{ij}(\tau, x)]^{1/2} (t-\tau)^{\frac{m}{2}}, \text{ on a :}$$

$$(10) \quad p(\tau, t, x, x) \sim \frac{1}{(2\pi)^{\frac{m}{2}} [\det \int_a^t a(s, x) ds]^{1/2}}.$$

Le même type d'arguments permet de montrer (8) et (9).

5. Passage à la densité de transition.

Dans ce paragraphe, on prouve que la solution fondamentale p que l'on vient de construire est la densité de transition de la diffusion associée à L_t . C'est-à-dire (voir R. Azencott [1]) que si l'on pose :

$$(11) \quad R_s^t g(x) = \int_{\mathbb{R}^m} p(s, t, x, \xi) g(\xi) d\xi,$$

pour g appartenant à $C_k(\mathbb{R}^m)$, alors on a pour toute f de $C_k^2(\mathbb{R}^m)$,

$$R_s^t f(x) = f(x) + \int_s^t R_s^u L_u f(x) du.$$

On utilisera le théorème suivant, valable même sous des hypothèses plus faibles pour L_t .

5.1. THEOREME. Soit R_s^t un semi-groupe sous-markovien qui satisfait à :

$$(12) \quad \begin{aligned} \frac{\partial}{\partial t} R_s^t g(x) &= L_t R_s^t g(x) \\ \lim_{t \searrow s} R_s^t g(x) &= g(x) \end{aligned} \quad \begin{aligned} &\text{pour toute } g \text{ de } C_k(\mathbb{R}^m) \\ &T > t > s, \end{aligned}$$

et qui est minimal dans cette classe, dans le sens que pour tout autre semi-groupe Q_s^t semi-markovien vérifiant (12), on ait :

$$Q_s^t g(x) \geq R_s^t g(x) \text{ pour toute } g \text{ de } C_k^+(\mathbb{R}^m)$$

Alors il vérifie également :

$$R_s^t f(x) = f(x) + \int_s^t R_s^u L_u f(x) du.$$

Ce théorème se trouve démontré dans le cas homogène dans R. Azencott ([1]).

Il faut donc démontrer que :

- a) R_s^t défini par (11) est solution de (12)
- b) qu'il est un semi-groupe sous-markovien
- c) qu'il est minimal dans cette classe.

Le point a) vient du fait que p est une solution fondamentale et que, grâce aux majorations (9) et (10), on peut dériver sous le signe somme ; b) et c) sont des conséquences classiques du principe faible du maximum que nous emploierons sous la forme suivante :

5.2. THÉORÈME. Soit V un ouvert de $\mathbb{R}^m$, $Q = V \times [T_1, T_2]$, $Q_0 = \partial V \times [T_1, T_2] \cup V \times \{T_1\}$,

$$M = \frac{1}{2} \sum_{i,j=1}^m a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(t,x) \frac{\partial}{\partial x_i} - \frac{\partial}{\partial t}$$

un opérateur différentiel sur Q tel que les matrices $\{a_{ij}(t,x)\}_{i,j}$ soient définies positives. Soit u une fonction sur Q deux fois dérivable en x, une fois en t et telle que $Mu \geq 0$ dans Q. Alors :

$$\sup_Q u \leq \sup_{Q_0} u.$$

Pour une démonstration, on peut voir par exemple A. Friedman [2].

5.3. LEMME. Avec les notations du théorème 5.2, soit v une fonction sur $[T_1 \times T_2] \times \mathbb{R}^m$, deux fois dérivable en x, une fois dérivable en t et telle que :

$$\begin{aligned} v(T_1, x) &\geq 0 && \text{pour tout } x \text{ de } \mathbb{R}^m \\ Mv &\leq 0 && \text{sur } [T_1, T_2] \times \mathbb{R}^m \\ \lim_{|x| \rightarrow +\infty} v(t, x) &\geq 0 && \text{uniformément en } t \in [T_1, T_2]. \end{aligned}$$

alors $v \geq 0$ sur $[T_1, T_2] \times \mathbb{R}^m$.

Démonstration. D'après les hypothèses, pour tout $\varepsilon > 0$, on a : $v(t, x) + \varepsilon > 0$ pour tout $t \in [T_1, T_2]$ et x hors d'un compact K. Pour tout ouvert $V \supset K$, on a donc :

$v(t, x) + \varepsilon > 0$ pour tout $(t, x) \in [T_1, T_2] \times \partial V$. Le théorème 5.2 appliqué à la fonction $u = -(v + \varepsilon)$ donne alors pour $Q = V \times [T_1, T_2]$,

$$\inf_Q (v + \varepsilon) = \sup_Q -(v + \varepsilon) = \sup_{Q_0} -(v + \varepsilon) \leq 0$$

donc $\inf_Q (v + \varepsilon) \geq 0$ et comme ε et V sont arbitraires, on a le résultat.

5.4. Démonstration de b) et c). Il est clair d'abord que R_s^t est un opérateur positif car si $g \in C_k(\mathbb{R}^m)$, et $g \geq 0$, alors grâce au point a) et à (7), $R_s^t g(x)$ en tant que fonction de (t, x) vérifie les hypothèses du lemme 5.3 avec $T_1 = s$, $T_2 = T$, $M = D$.

Deuxièmement, $(R_s^t)_{s, t}$ est un semi-groupe sous-markovien. L'équation de Chapman-Kolmogorov vient du fait que si on pose :

$$v(t, x) = R_\tau^t R_s^\tau g(x) - R_s^t g(x),$$

alors v et $-v$ vérifient à la fois les hypothèses du lemme 5.3 pour $M=D$, $T_1 = \tau$, $T_2 = T$, d'où $v \equiv 0$ sur $[\tau, T] \times \mathbb{R}^m$; donc on a :

$$R_\tau^t R_s^\tau g(x) = R_s^t g(x).$$

Analogue sera la démonstration du fait que R_s^t est minimal. Si Q_s^t est un autre semi-groupe sous-markovien qui est solution de (12), et $g \in C_k^+(\mathbb{R}^m)$, le lemme 5.3. appliqué à $M=D$, $v(t, x) = Q_s^t f(x) - R_s^t f(x)$, $T_1 = s$, $T_2 = T$ donne le résultat, ce qui prouve en même temps que R_s^t est une contraction.

Comme conséquence du théorème 4.2, de la remarque du paragraphe 4.1 et des résultats de ce dernier paragraphe, on peut énoncer :

5.5. THÉORÈME.

$$D^\gamma = L_t^\gamma - \frac{\partial}{\partial t} = \frac{1}{2} \sum_{i, j=1}^m a_{ij}^\gamma(t, x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i^\gamma(t, x) \frac{\partial}{\partial x_i} - \frac{\partial}{\partial t}$$

$\gamma \in \Gamma$, une famille d'opérateurs qui satisfont aux hypothèses A_1, A_2, A_3 de façon telle que les constantes α, k, M, μ_1 et μ_2 puissent être choisies indépendamment de $\gamma \in \Gamma$. Alors les densités de transition $p^{(\gamma)}(\tau, t, x, \xi)$ associées aux D^γ vérifient les relations (7) et (8) uniformément en $\gamma \in \Gamma$.

BIBLIOGRAPHIE

- [1] R. Azencott, *Behaviour of diffusion semi-groups at infinity*. Bull. Soc. Math. France 102 (1974) p. 193-240.
- [2] A. Friedman, *Partial differential equation of parabolic type*. Prentice Hall 1964.

CHAPITRE 5

ÉQUIVALENT DE LA DENSITÉ D'UNE DIFFUSION EN TEMPS PETIT CAS DES POINTS PROCHES

par

Laure ELIE

1. Introduction.

Soit M une variété connexe de classe C^{k+1} ($k \geq 4$), de dimension m . On considère sur M un opérateur différentiel du second ordre elliptique L vérifiant $L1 = 0$. Alors, sur tout domaine U d'une carte locale (U, ϕ) , L s'écrit selon les notations du chapitre 3, sous la forme

$$L = \frac{1}{2} \sum_{1 \leq i, j \leq m} a_{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq m} b_i(x) \frac{\partial}{\partial x_i},$$

où, pour tout $x \in U$, la matrice $a(x) = [a_{i,j}(x)]_{1 \leq i, j \leq m}$ est définie positive ; nous noterons $g(x) = [g_{i,j}(x)]_{1 \leq i, j \leq m}$ la matrice inverse de $a(x)$.

Nous supposons de plus que les coefficients $a_{i,j}$ sont de classe C^k et les coefficients b_i de classe C^{k-1} . Ces conditions s'imposent de façon naturelle car ce sont les plus fortes hypothèses que l'on puisse faire sur les coefficients de L , qui soient stables par changement de carte de classe C^{k+1} .

Nous munirons alors M d'une structure riemannienne de classe C^k à l'aide de la métrique définie localement par

$$ds^2 = \sum_{1 \leq i, j \leq m} g_{i,j}(x) dx_i dx_j,$$

et nous désignerons par d la distance riemannienne et par π le volume riemannien.

Si Δ est l'opérateur de Laplace-Beltrami associé à la métrique ds^2 , alors L s'écrit $\frac{1}{2}\Delta + Y$ où Y est un opérateur différentiel du premier ordre. En effet localement Δ s'écrit [1]

$$\Delta = \frac{1}{\sqrt{\det g}} \sum_{1 \leq j \leq m} \frac{\partial}{\partial x_j} \left(\sum_{1 \leq i \leq m} g^{i,j} \sqrt{\det g} \frac{\partial}{\partial x_i} \right)$$

où la matrice $[g^{i,j}(x)]_{1 \leq i,j \leq m}$ désigne la matrice inverse de $g(x)$ et est donc $a(x)$.

Il en résulte qu'étudier L revient à étudier, sur une variété M portant une structure riemannienne de classe C^k , un opérateur du type $\frac{1}{2}\Delta + Y$, où Δ est l'opérateur de Laplace-Beltrami et Y un champ de vecteurs de classe C^{k-1} .

Nous savons (Azencott, chapitre 2) qu'il existe sur M une diffusion, unique à une équivalence près, admettant L comme générateur différentiel et que le semi-groupe de transition de cette L -diffusion admet une densité $p(t,x,y)$ par rapport à π continue sur $\mathbb{R}^+ \times M \times M$. De plus cette fonction p vérifie en tant que fonction de (t,x) l'équation parabolique

$$\left(\frac{\partial}{\partial t} - L\right) p(t,x,y) = 0 \quad (t,x,y) \in \mathbb{R}^+ \times M \times M.$$

Notre but est d'étudier le comportement de $p(t,x,y)$ lorsque $t \rightarrow 0$ pour des points x et y proches. Suivant Molchanov [3], nous prouverons dans cet exposé le théorème suivant :

1.1. THÉORÈME. *Soit M une variété portant une structure Riemannienne de classe C^k ($k \geq 4$) et $\mathcal{P}$ une diffusion sur M de générateur différentiel $L = \frac{\Delta}{2} + Y$ où Δ désigne l'opérateur de Laplace-Beltrami et Y un champ de vecteurs de classe C^{k-1} . Soit $p(t,x,y)$ la version continue sur $\mathbb{R}^+ \times M \times M$ de la densité par rapport au volume riemannien π du semi-groupe de transition de $\mathcal{P}$. On désigne par d la distance riemannienne et par $\langle \cdot \rangle$ le produit scalaire riemannien.*

Alors, si D est un compact quelconque de M , il existe une constante ε positive telle que, si on pose $D_\varepsilon = \{(x,y) \in D \times D, d(x,y) < \varepsilon\}$,

. il existe, pour tout $(x,y) \in D_\varepsilon$, une unique géodésique de longueur minimale $\gamma_{x,y} : [0,1] \rightarrow M$ joignant x et y ,

. on ait, uniformément pour $(x,y) \in D_\varepsilon$, lorsque $t \rightarrow 0$,

$$p(t,x,y) \sim (2\pi t)^{-\frac{m}{2}} H(x,y) \exp \left[\frac{-d^2(x,y)}{2t} + A(x,y) \right]$$

où $A(x,y) = \int_0^1 \langle Y(\gamma_{x,y}(s)), \dot{\gamma}_{x,y}(s) \rangle ds$

et $H(x,y) = [\det(\exp'_x(\dot{\gamma}_{x,y}(0)))]^{-1/2}$, c'est-à-dire représente l'inverse de la racine

carrée du Jacobien de $\exp_x$ au point $\dot{\gamma}_{x,y}(0)$ de $T_x(M)$.

1.2. Remarques.

1) Le terme $A(x,y)$ qui représente le "travail" du champ Y le long de la géodésique $\gamma_{x,y}$ est la seule contribution de l'opérateur Y dans l'équivalent de $p(t,x,y)$.

2) Au cours de la démonstration de ce théorème, Molchanov se sert de l'estimation de Varadhan suivante [4] :

$$\lim_{t \rightarrow 0} -2t \ln p(t,x,y) = d^2(x,y)$$

uniformément pour x et y appartenant à un compact arbitraire de M . La méthode que nous utiliserons ici permettra de court-circuiter cette estimation, qui n'est en fait démontrée par Varadhan que sur des espaces $\mathbb{R}^m$ et qui ne semble pas aisée à obtenir à priori sur une variété quelconque. Cette estimation deviendra en fait conséquence du théorème 1.1 pour des points suffisamment proches d'un compact et il sera prouvé dans le chapitre 9 (C. Bellaïche) qu'elle s'étend aux points d'un compact arbitraire d'une variété complète.

3) Signalons que Y. Kifer a obtenu dans ([2]), moyennant certaines hypothèses de différentiabilité sur M et les coefficients de L , un développement asymptotique de $p(t,x,y)$ lorsque $t \rightarrow 0$.

4) Remarquons qu'ici l'hypothèse de différentiabilité C^k ($k \geq 4$) semble trop forte et que ce type de résultat devrait pouvoir se montrer sous l'hypothèse C^k ($k \geq 2$). Et, comme nous le verrons, c'est le passage aux coordonnées semi-géodésiques qui fait perdre trop de différentiabilité.

1.3. Aperçu de la démonstration.

L'étude de $p(t,x,y)$ pour des points voisins est en fait un problème local, et l'on peut, à l'aide d'une carte, transporter cette étude sur $\mathbb{R}^m$. Mais on ne peut raisonner si directement car on ne peut montrer que le problème est local qu'une fois l'étude sur $\mathbb{R}^m$ faite. Alors on procédera de la façon suivante : on considérera l'image L_ϕ de L par une carte semi-géodésique $\phi : U \rightarrow \mathbb{R}^m$, U ouvert de M , et on prolongera l'opérateur L_ϕ à l'espace $\mathbb{R}^m$ entier ; on cherchera sur $\mathbb{R}^m$, à l'aide de la paramétrix, un équivalent de la densité du semi-groupe de transition associé à l'opérateur prolongé L et on montrera que l'on peut, vu la forme de cet équivalent, en déduire un équivalent de $p(t,x,y)$ sur M .

Nous montrerons en fait le résultat suivant :

1.4. PROPOSITION. *On se place sous les hypothèses du théorème 1.1. Alors à tout point x_0 de M , on peut associer une boule ouverte centrée en x_0 et de rayon δ , notée $\mathcal{B}(x_0, \delta)$ telle que*

. il existe une unique géodésique de longueur minimale $\gamma_{x,y} : [0,1] \rightarrow M$ joignant deux points quelconques x et y de $\mathcal{B}(x_0, \delta)$,

. on ait, uniformément pour x et $y \in \mathcal{B}(x_0, \delta)$, lorsque $t \rightarrow 0$,

$$p(t, x, y) \sim (2\pi t)^{-\frac{m}{2}} H(x, y) \exp\left[\frac{-d^2(x, y)}{2t} + A(x, y)\right] \quad (0)$$

$$\text{où } A(x, y) = \int_0^1 \langle Y(\gamma_{x,y}(s)), \dot{\gamma}_{x,y}(s) \rangle ds$$

et où $H(x, y)$ s'obtient de la façon suivante :

Il existe une carte locale $\phi : \mathcal{B}(x_0, \delta) \rightarrow \mathbb{R}^m$ telle que si g est la matrice définissant la métrique riemannienne dans cette carte, on ait

$$H(x, y) = [\det g^{-1}(y) \det g^{-1}(x)]^{1/4} [\det \int_0^1 g^{-1}(\gamma_{x,y}(s)) ds]^{-1/2}$$

Nous renverrons au chapitre 6 (J. Granara) pour l'interprétation géométrique de $H(x, y)$ à l'aide des champs de Jacobi et la démonstration de l'égalité $H(x, y) = [\det \exp'_x(\dot{\gamma}_{x,y}(0))]^{-1/2}$.

Un argument de compacité immédiat permet alors de déduire le théorème 1.1 de la proposition 1.4, c'est-à-dire de prouver que (0) est encore vraie uniformément pour tous les points d'un compact arbitraire D situés à une distance inférieure à ε . En effet les boules $\mathcal{B}(x, \delta_x)$, $x \in D$, forment un recouvrement ouvert de D et si $\mathcal{B}(x_i, \delta_{x_i})$, $i \in I$, en désigne un recouvrement fini extrait, il suffit de poser $\varepsilon = \inf_{i \in I} \delta_{x_i}$.

2. Transport sur $\mathbb{R}^m$.

2.1. Carte semi-géodésique.

Nous savons (théorème 3.5 du chapitre 6) qu'à tout point x_0 de M , on peut associer un ouvert relativement compact U contenant x_0 tel que pour tout x et y de U , il existe une carte semi-géodésique $\phi_{x,y} : U \rightarrow \mathbb{R}^m$ associée à $(x, \dot{\gamma}_{x,y}(0))$ où $\gamma_{x,y}$ est l'unique géodésique de longueur minimale paramétrée par $[0,1]$ joignant x à y . En outre, l'application $\phi_{x,y}$ est de classe C^{k-1} et ses dérivées partielles jusqu'à l'ordre 2 sont bornées sur U uniformément en $(x, y) \in U \times U$. Lorsque ce ne sera pas ambigu, $\phi_{x,y}$ et $\gamma_{x,y}$ seront notés ϕ et γ . Alors, selon les notations du chapitre 6, pour tout $z \in U$, $\phi(z)$ s'écrira $(\tau, u) \in \mathbb{R} \times \mathbb{R}^{m-1}$ avec en particulier $\phi(\gamma(s)) = (s, 0)$,

d'où $\phi(x) = (0,0)$ noté 0 et $\phi(y) = (1,0)$ noté e_1 , et l'image de la géodésique γ est le segment $[0, e_1]$; de plus dans cette carte (U, ϕ) la métrique riemannienne s'écrit sous la forme

$$ds^2 = d\tau^2 + \sum_{1 \leq i, j \leq m-1} g_{i,j} du_i du_j$$

Comme l'application ϕ est de classe C^{k-1} sur U , les coefficients $g_{i,j}$ sont de classe C^{k-2} sur U . Il s'avère donc que le passage par la carte semi-géodésique fait perdre beaucoup de différentiabilité, d'autant plus que les géodésiques sont de classe C^k .

Dans cette carte, l'opérateur $L = \frac{1}{2} \Delta + Y$ s'exprimera alors, selon les notations du chapitre 3, sous la forme

$$L = \frac{\partial^2}{\partial \tau^2} + \frac{1}{2} \sum_{1 \leq i, j \leq m-1} g^{i,j} \frac{\partial^2}{\partial u_i \partial u_j} + b \frac{\partial}{\partial \tau} + \sum_{1 \leq i \leq m-1} b_i \frac{\partial}{\partial u_i}$$

où $[g^{i,j}]_{1 \leq i, j \leq m-1}$ désigne la matrice inverse de $g = [g_{i,j}]_{1 \leq i, j \leq m-1}$

$$b = \frac{1}{4} \left(\frac{\partial}{\partial \tau} \ln \det g \right) + \langle Y, \frac{\partial}{\partial \tau} \rangle$$

les coefficients $g^{i,j}$ sont C^{k-2} et les coefficients b et b_i sont C^{k-3} .

Et si on désigne par L_ϕ l'opérateur sur l'ouvert $\phi(U)$ de $\mathbb{R}^m$, image de L par ϕ , on pourra écrire

$$L_\phi = \frac{\partial^2}{\partial \tau^2} + \frac{1}{2} \sum_{1 \leq i, j \leq m-1} \bar{g}^{i,j} \frac{\partial^2}{\partial u_i \partial u_j} + \bar{b} \frac{\partial}{\partial \tau} + \sum_{1 \leq i \leq m-1} \bar{b}_i \frac{\partial}{\partial u_i}$$

avec $\bar{g}^{i,j} \circ \phi = g^{i,j}$ et $\bar{b}_i \circ \phi = b_i$.

On a imposé à k d'être ≥ 4 pour que les coefficients $\bar{b}$ et $\bar{b}_i$ soient C^1 et ceci dans le but d'utiliser la méthode de la paramétrix pour étudier L_ϕ . Comme les applications $\phi_{x,y}$ et leurs dérivées partielles sont bornées sur U , uniformément en $(x,y) \in U \times U$, il en est de même des coefficients $\bar{g}^{i,j}$, $\bar{b}$ et $\bar{b}_i$ (qui dépendent de (x,y)) et de leurs dérivées partielles d'ordre 1. De plus les opérateurs L_ϕ sont uniformément elliptiques sur $\phi(U)$ et ceci uniformément pour $(x,y) \in U \times U$, c'est-à-dire que le déterminant de $\bar{g}^{-1} = [\bar{g}^{i,j}]_{1 \leq i, j \leq m-1}$ est uniformément minoré.

2.2. Diffusions associées sur $\mathbb{R}^m$.

Pour tout $(x,y) \in U \times U$, on prolonge l'opérateur L_ϕ à $\mathbb{R}^m$ de façon à conserver toutes les propriétés de différentiabilité et d'uniformité ci-dessus. Plus précisément, si on note $\bar{L}$ l'opérateur prolongé, il s'écrit sous la forme

$$\frac{\partial^2}{\partial \tau^2} + \frac{1}{2} \sum_{1 \leq i, j \leq m-1} \bar{g}^{i,j} \frac{\partial^2}{\partial u_i \partial u_j} + \bar{b} \frac{\partial}{\partial \tau} + \sum_{1 \leq i \leq m-1} \bar{b}_i \frac{\partial}{\partial u_i}$$

et il existera deux réels C et λ strictement positifs, indépendants de $(x, y) \in U \times U$ tels que l'on ait :

$$(1) \quad \sup_{1 \leq i, j \leq m-1} [\|\bar{g}^{i,j}\|, \|\bar{b}\|, \|\bar{b}_i\|, \|\bar{d}\bar{g}^{i,j}\|, \|\bar{d}\bar{b}\|, \|\bar{d}\bar{b}_i\|] \leq C$$

où $\bar{d}b$ désigne l'application différentielle de b ,

$$(2) \quad \inf [\text{valeurs propres de } \bar{g}^{-1} = [\bar{g}^{i,j}]_{1 \leq i, j \leq m-1}] \geq \lambda \text{ sur } \mathbb{R}^m.$$

Il existe sur $\mathbb{R}^m$, à une équivalence près, une unique diffusion $\bar{\mathcal{Q}}$ admettant $\bar{L}$ comme générateur différentiel, et on notera $q(t, v, w)$ la version continue sur $\mathbb{R}^+ \times \mathbb{R}^m \times \mathbb{R}^m$ de la densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^m$ de son semi-groupe de transition.

Si V est un ouvert inclus dans U , les générateurs différentiels L_ϕ et $\bar{L}$ coïncident sur $\phi(V)$ et par suite les processus induits $\phi[\mathcal{P}_V]$ et $\bar{\mathcal{Q}}_{\phi(V)}$ sont équivalents (cf. Azencott, chapitre 2). Il en résulte que les densités de transition continues p_V de $\mathcal{P}_V$ et q_V de $\bar{\mathcal{Q}}_{\phi(V)}$ vérifient la relation

$$p_V(t, z_1, z_2) = (\det \bar{g}[\phi(z_2)])^{-1/2} q_V(t, \phi(z_1), \phi(z_2))$$

sur $\mathbb{R}^+ \times U \times U$ car la première densité est regardée relativement au volume riemannien π et la deuxième relativement à la mesure de Lebesgue sur $\mathbb{R}^m$ et que l'image de π par ϕ admet pour densité par rapport à la mesure de Lebesgue $\sqrt{\det \bar{g}}$.

En fait, à la densité p , on a associé toute une famille de densités q sur $\mathbb{R}^m$, indicées par $(x, y) \in U \times U$. Mais comme on s'est arrangé pour que les coefficients des générateurs différentiels associés aient des propriétés uniformes en (x, y) , on verra que les résultats obtenus sont uniformes en $(x, y) \in U \times U$ et on gardera seulement en mémoire le fait que q dépend de (x, y) .

2.3. La méthode utilisée pour étudier $p(t, x, y)$ sera la suivante :

i) Nous allons chercher un équivalent de $q(t, \phi(x), \phi(y)) = q(t, 0, e_1)$ uniformément pour x et $y \in U$ lorsque $t \rightarrow 0$. Ce sera l'objet du paragraphe 3.

ii) Nous montrerons au paragraphe 4 grâce à la forme de l'équivalent de q obtenu qu'il existe deux ouverts W_1 et V de U contenant x_0 tels que $\bar{W}_1 \subset V \subset \bar{V} \subset U$ et tels que l'on ait

$$q_V(t, \phi(x), \phi(y)) \sim q(t, \phi(x), \phi(y))$$

uniformément pour x et $y \in W_1$. On connaîtra donc un équivalent de $p_V(t, x, y)$ uniformément pour x et $y \in W_1$.

iii) Comme au ii), nous verrons à l'aide de la forme de l'équivalent de p_V qu'il existe un ouvert W_2 contenant x_0 tel que $\bar{W}_2 \subset W_1$ et tel que l'on ait

$$p(t, x, y) \sim p_V(t, x, y)$$

uniformément pour x et $y \in W_2$. La proposition 1.4 sera alors démontrée.

3. Etude sur $\mathbb{R}^m$ de $q(t, 0, e_1)$

L'outil de base pour cette étude est la méthode de la paramétrix, et nous allons donc utiliser les résultats obtenus au chapitre 4 (P. Baldi). Vu les conditions (1) et (2) que vérifie $\bar{L}$, on obtient, à l'aide du théorème 4.2 du chapitre 4, un majorant de $q(t, v, w)$ pour des points v et w quelconques de $\mathbb{R}^m$ et un équivalent lorsque $v = w$.

3.1. On cherche à étudier $q(t, 0, e_1)$ et on connaît donc seulement, puisque $0 \neq e_1$, un majorant de ce terme. L'idée est alors de considérer le processus $\mathcal{Q}_1$ image du processus $\mathcal{Q}$ par l'application qui à X_s , position du processus $\mathcal{Q}$ à l'instant s , associe $X_s - \frac{s}{t} e_1$. Ce nouveau processus est un processus de Markov inhomogène dont la version continue de sa densité q_1^t vérifie en particulier

$$q_1^t(0, t, 0, 0) = q(t, 0, e_1),$$

et compte tenu des résultats du chapitre 4, il est plus judicieux de rechercher un équivalent de $q_1^t(0, t, 0, 0)$.

Si on note L_1^t le générateur différentiel de $\mathcal{Q}_1$, l'indice t indiquant que le processus $\mathcal{Q}_1$ dépend non seulement du couple (x, y) de $U \in U$, mais aussi du paramètre t , on obtient, pour $s \in \mathbb{R}^+$ et $v = (\tau, u) \in \mathbb{R}^m$,

$$\begin{aligned} L_1^t(s, v) &= \frac{\partial^2}{\partial \tau^2} + \frac{1}{2} \sum_{1 \leq i, j \leq m-1} \bar{g}^{i, j}(v + \frac{s}{t} e_1) \frac{\partial^2}{\partial u_i \partial u_j} \\ &+ [\bar{b}(v + \frac{s}{t} e_1) - \frac{d(x, y)}{t}] \frac{\partial}{\partial \tau} + \sum_{1 \leq i \leq m-1} \bar{b}_i(v + \frac{s}{t} e_1) \frac{\partial}{\partial u_i} \end{aligned}$$

on remarque alors que le coefficient de $\frac{\partial}{\partial \tau}$ n'est pas uniformément borné en t lorsque $t \rightarrow 0$, ce qui entraîne que les conditions du théorème 4.2 du chapitre 4 ne sont pas

vérifiées uniformément en t pour les opérateurs L_1^t et par suite l'équivalent de $q_1^t(0, s, 0, 0)$ obtenu par ce théorème lorsque $s \rightarrow 0$, n'a aucune raison d'être uniforme en t ; et on ne peut rien dire de $q_1^t(0, t, 0, 0) = q(t, 0, e_1)$.

3.2. On cherche donc à supprimer le terme $\frac{d(x,y)}{t} \frac{\partial}{\partial t}$ et on pose

$$L_2^t = L_1^t + \frac{d(x,y)}{t} \frac{\partial}{\partial t}$$

On considère l'unique processus de Markov inhomogène Q_2 , défini à une équivalence près, admettant L_2^t comme générateur différentiel et on note q_2^t la version continue de sa densité de transition.

D'après les conditions (1) et (2), l'opérateur L_2^t vérifie alors les hypothèses du théorème 4.2 du chapitre 4 uniformément en $(x, y) \in U \times U$ et en $t \in \mathbb{R}^{++}$. De ce fait, pour tout $s_0 \in \mathbb{R}^+$, il existe deux constantes C_1 et C_2 indépendantes de $(x, y) \in U \times U$ et de $t \in \mathbb{R}^{++}$ telles que l'on ait

$$(3) \quad q_2^t(s, s_2, v, w) \leq C_1 (s_2 - s_1)^{-\frac{m}{2}} \exp\left[-\frac{C_2 |v-w|^2}{s_2 - s_1}\right]$$

pour $s_1 \leq s_2 \leq s_0$ et $(v, w) \in \mathbb{R}^m \times \mathbb{R}^m$.

De plus on a, lorsque $s \rightarrow 0$, uniformément en $(x, y) \in U \times U$ et $t \in \mathbb{R}^{++}$, en $v \in \mathbb{R}^m$, en $(s_1, s_2) \in \mathbb{R}^{+2}$ tels que $0 \leq s_1 < s_2 \leq s$,

$$q_2^t(s_1, s_2, v, v) \sim (2\pi)^{-\frac{m}{2}} \left[\det \int_{s_1}^{s_2} a^t(s, v) ds \right]^{-\frac{1}{2}}$$

avec

$$a^t(s, v) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & g^{-1}(v + \frac{s}{t} e_1) & & \\ 0 & & & \end{bmatrix}$$

En particulier, on obtient lorsque $t \rightarrow 0$, uniformément en $(x, y) \in U \times U$,

$$(4) \quad q_2^t(0, t, 0, 0) \sim (2\pi)^{-\frac{m}{2}} t^{-\frac{1}{2}} \left[\det \int_0^t g^{-1}(\frac{s}{t} e_1) ds \right]^{-\frac{1}{2}}.$$

3.3. La formule de Cameron-Martin (cf. chapitre 1 de M. Chaleyat-Maurel) va nous permettre de comparer q_2^t et q_1^t . En effet si P_1^V et P_2^V désignent respectivement les lois des trajectoires de $\mathcal{Q}_1$ et $\mathcal{Q}_2$, P_2^V est solution du problème des martingales associé à $(0, v, a^t, b^t)$ avec $b^t(s, v) = (\bar{b}, \dots, \bar{b}_1, \dots)(v + \frac{s}{t} e_1)$ et P_1^V est solution du problème des martingales associé à $(0, v, a^t, b^t + a^t c)$ avec $a^t c(s, v) = (\frac{-d(x, y)}{t}, 0, \dots, 0)$.

Il résulte de la forme particulière de a^t que $c(s, v) = (\frac{-d(x, y)}{t}, 0, \dots, 0)$. On en déduit (théorème 5.1 du chapitre 1) que

$$\frac{dP_1^V}{dP_2^V} / \mathcal{F}_t = \exp \left[\frac{-d(x, y)}{t} \left(X_t^1 - \int_0^t \bar{b}(X_s + \frac{s}{t} e_1) ds \right) - \frac{d^2(x, y)}{2t} \right]$$

où X_t^1 désigne la première coordonnée de X_t .

Alors si D est un voisinage de 0 dans $\mathbb{R}^m$, on obtient si P^V désigne la loi des trajectoires du processus $\mathcal{Q}$,

$$\begin{aligned} P^O[X_t \in e_1 + D] &= P_1^O[X_t \in D] = E_2^O[1_D(X_t) \frac{dP_1}{dP_2} / \mathcal{F}_t] \\ &= \exp \frac{-d^2(x, y)}{2t} E_2^O[1_D(X_t) \phi_{x, y}(X_{[0, t]})] \end{aligned}$$

$$\text{où } \phi(X_{[0, t]}) = \exp \left[\frac{d(x, y)}{t} \left(\int_0^t \bar{b}(X_s + \frac{s}{t} e_1) ds - X_t^1 \right) \right]$$

la variable $X_{[0, t]}$ indiquant que $\phi(X_{[0, t]})$ dépend de X_s pour $0 \leq s \leq t$.

De ce fait, on a uniformément en $(x, y) \in U \times U$ et en $t \leq s_0$,

$$\lim_{D \searrow 0} \frac{1}{\text{mes } D} P^O[X_t \in e_1 + D] = q(t, 0, e_1)$$

$$\text{et } \lim_{D \searrow 0} \frac{1}{\text{mes } D} P_2^O[X_t \in D] = q_2^t(0, t, 0, 0)$$

où $\text{mes } D$ désigne la mesure de Lebesgue de D .

Par suite,

$$\begin{aligned}
 q(t, 0, e_1) &= \exp \frac{-d^2(x, y)}{2t} \lim_{D \searrow 0} \frac{1}{\text{mes } D} E_2^O(1_D(X_t) \phi(X_{[0, t]})) \\
 &= \exp \frac{-d^2(x, y)}{2t} \lim_{D \searrow 0} \frac{1}{\text{mes } D} [P_2^O(X_t \in D) E_2^O(\phi(X_{[0, t]})/X_t \in D)] \\
 &= \exp \frac{-d^2(x, y)}{2t} q_2^t(0, t, 0, 0) \lim_{D \searrow 0} E_2^O(\phi(X_{[0, t]})/X_t \in D).
 \end{aligned}$$

3.4. Connaissant le comportement de $q_2^t(0, t, 0, 0)$, il s'agit d'étudier le dernier terme de l'égalité ci-dessus. Or ce dernier revient à conditionner le processus $\mathfrak{P}_2$ par $X_0 = X_t = 0$. Ces processus conditionnés ou "ponts" sont, lorsqu'ils sont associés à une diffusion, l'objet du paragraphe 5 de l'article de Molchanov et il y est déterminé leur densité de transition conditionnelle.

Dans notre cas particulier, le terme à étudier est simple et un calcul direct va nous permettre de prouver le lemme suivant :

3.5. LEMME. Lorsque $t \rightarrow 0$, nous avons, uniformément pour $(x, y) \in U \times U$,

$$\lim_{D \searrow 0} E_2^O(\phi(X_{[0, t]})/X_t \in D) \sim \exp\left(\frac{d(x, y)}{t}\right) \int_0^t \bar{b}\left(\frac{s}{t}e_1\right) ds.$$

Preuve : Posons $d = d(x, y)$. Nous pouvons écrire

$$E_2^O(\phi(X_{[0, t]})/X_t \in D) = \exp\left(\frac{d}{t} \int_0^t \bar{b}\left(\frac{s}{t}e_1\right) ds\right) E_2^O(\psi(X_{[0, r]})/X_t \in D)$$

$$\text{où } \psi(X_{[0, t]}) = \exp\left[\frac{d}{t}(-X_t^1 + \int_0^t (\bar{b}(X_s + \frac{s}{t}e_1) - \bar{b}(\frac{s}{t}e_1)) ds)\right].$$

Il va s'agir de montrer que si on considère

$$A(t, D) = |1 - E_2^O(\psi(X_{[0, t]})/X_t \in D)|,$$

alors $\lim_{t \searrow 0} (\lim_{D \searrow 0} A(t, D)) = 0$ uniformément pour $(x, y) \in U \times U$.

Comme d'après (1), $\|\bar{b}\|$ et $\|d\bar{b}\|$ sont bornés par C, il résulte de l'inégalité $|1 - e^a| \leq |a| e^{|a|}$, vraie pour $a \in \mathbb{R}$, que

$$a(t, D) \leq \frac{d}{t} E_2^0(|X_t^1| + C \int_0^t |X_s| ds) \exp(\frac{d}{t} |X_t^1| + 2Cd) / X_t \in D].$$

Puisque Q_2 est un processus de Markov, nous avons

$$\begin{aligned} E_2^0\left[\left(\int_0^t |X_s| ds\right) \exp\left(\frac{d}{t} |X_t^1|\right)\right] &= \int_0^t E_2^0(|X_s| \exp(\frac{d}{t} |X_t^1|) / X_t \in D) \\ &= \frac{\int_D h(z) \exp(\frac{d}{t} |z_1|) dz}{\int_D q_2^t(0, t, 0, z) dz} \end{aligned}$$

$$\text{avec } h(z) = \int_0^t \int_{\mathbb{R}^m} |v| q_2^t(0, s, 0, v) q_2^t(s, t, v, z) ds dv,$$

et z_1 la première coordonnée de z dans $\mathbb{R}^m$.

De plus,

$$\lim_{D \searrow 0} E_2^0[|X_t^1| \exp(\frac{d}{t} |X_t^1|) / X_t \in D] = 0.$$

Par suite, si h est continue en 0, on obtient

$$\lim_{D \searrow 0} A(t, D) \leq \frac{Cd}{t} \exp(2Cd) \frac{h(0)}{q_2^t(0, t, 0, 0)}.$$

En fait ce n'est pas h que l'on va étudier, mais la fonction h_1 suivante qui majore h , pour $t \leq t_0$, d'après (3),

$$h_1(z) = C_1^2 \int_0^t \int_{\mathbb{R}^m} |v| \exp(-C_2 \frac{|v|^2}{s}) \exp(-C_2 \frac{|z-v|^2}{t-s}) s^{-m/2} (t-s)^{-m/2} ds dv.$$

Il est facile de voir que

$$h_1(z) = C_1^2 t^{-(m+1)/2} \exp(-C_2 \frac{|z|^2}{t}) \int_0^t \int_{\mathbb{R}^m} \sqrt{s(t-s)} |v+z| \sqrt{\frac{s}{t(t-s)}} |\exp(-C_2 |v|^2)| ds dv$$

et si $|z| < \varepsilon$, alors

$$\sqrt{s(t,s)} |v+z \sqrt{\frac{s}{t(t-s)}}| \exp(-C_2 |v|^2) \leq \sqrt{s(t-s)} |v| \exp(-C_2 |v|^2) + \frac{\varepsilon s}{\sqrt{t}} \exp(-C_2 |v|^2)$$

et cette fonction majorante est intégrable sur $[0,t] \times \mathbb{R}^m$. Il résulte alors du théorème de Lebesgue que h_1 est continue en 0 et de plus

$$\begin{aligned} h_1(0) &= C_1^2 t^{-(m+1)/2} \left(\int_0^t \sqrt{s(t-s)} ds \right) \left(\int_{\mathbb{R}^m} |v| \exp(-C_2 |v|^2) dv \right) \\ &= K_1 t^{-(m+1)/2} t^2 = K_1 t^{-(m-3)/2}. \end{aligned}$$

D'où

$$D \lim_{\searrow 0} A(t,D) \leq C \frac{d}{dt} \exp(2Cd) \frac{h_1(0)}{q_2^t(0,t,0,0)},$$

et il existe donc une constante K_2 , indépendante de $(x,y) \in U \times U$, telle que

$$D \lim_{\searrow 0} A(t,D) \leq K_2 \frac{t^{-(m-1)/2}}{q_2^t(0,t,0,0)}$$

Or on sait d'après (4) que, lorsque $t \rightarrow 0$, uniformément en $(x,y) \in U \times U$,

$$q_2^t(0,t,0,0) \sim (2\pi)^{-\frac{m}{2}} t^{-\frac{1}{2}} \left[\det \int_0^t \bar{g}^{-1} \left(\frac{s}{t} e_1 \right) ds \right]^{-\frac{1}{2}}.$$

D'où

$$q_2^t(0,t,0,0) \sim (2\pi)^{-\frac{m}{2}} t^{-\frac{m}{2}} \left[\det \int_0^1 \bar{g}^{-1}(se_1) ds \right]^{-\frac{1}{2}}.$$

Comme les coefficients de $\bar{g}^{-1}$ sont uniformément bornés sur $\phi(U)$, pour $(x,y) \in U \times U$, on en déduit qu'il existe une constante K_3 , indépendante de $(x,y) \in U \times U$, telle que

$$q_2^t(0,t,0,0) \geq K_3 t^{-m/2}.$$

En conséquence, pour $t \leq t_0$,

$$D \lim_{\searrow 0} A(t,D) \leq \frac{K_2}{K_3} t^{1/2}$$

et ce dernier terme, indépendant de $(x,y) \in U \times U$, tend vers 0 lorsque $t \rightarrow 0$. D'où le lemme.

3.6. Equivalent de $q(t, 0, e_1)$ en temps petit :

Nous avons montré que

$$q(t, 0, e_1) \sim \exp\left(\frac{-d^2(x, y)}{2t}\right) q_2^t(0, t, 0, 0) \exp\left[\frac{d(x, y)}{t} \int_0^t \bar{b}\left(\frac{s}{t} e_1\right) ds\right]$$

et connaissant un équivalent de $q_2^t(0, t, 0, 0)$, nous en déduisons que

$$q(t, 0, e_1) \sim (2\pi)^{-m/2} t^{-m/2} [\det \int_0^1 \bar{g}^{-1}(se_1) ds]^{-\frac{1}{2}} \exp\left[\frac{-d^2(x, y)}{2t} + d(x, y) \int_0^1 \bar{b}(se_1) ds\right].$$

Comme la géodésique de longueur minimale γ joignant x à y est paramétrée par $[0, 1]$, l'image de $\gamma(s)$, $s \in [0, 1]$ par la carte locale semi-géodésique $\phi_{x, y}$ est se_1 et $\frac{\partial}{\partial t}[\gamma(s)] = \frac{1}{d(x, y)} \frac{\partial}{\partial s}(\gamma(s))$.

Par suite,

$$\bar{g}(se_1) = g(\gamma(s))$$

$$\text{et } \bar{b}(se_1) = b(\gamma(s)) = \frac{1}{d(x, y)} \left[\frac{1}{4} \frac{\partial}{\partial s} \ln \det g(\gamma(s)) + \langle Y(\gamma(s)), \dot{\gamma}(s) \rangle \right]$$

et donc, si nous posons $A(x, y) = \int_0^1 \langle Y(\gamma(s)), \dot{\gamma}(s) \rangle ds$, nous obtenons

$$\exp[d(x, y) \int_0^1 \bar{b}(se_1) ds] = [\det g(y)]^{1/4} [\det g(x)]^{-1/4} \exp(A(x, y))$$

D'où

$$(6) \quad q(t, \phi(x), \phi(y)) = q(t, 0, e_1) \sim (2\pi t)^{-m/2} \bar{H}(x, y) \exp\left[\frac{-d^2(x, y)}{2t} + A(x, y)\right]$$

si $\bar{H}(x, y)$ désigne $[\det g^{-1}(x)]^{1/4} [\det g^{-1}(y)]^{-1/4} [\det \int_0^1 g^{-1}(\gamma(s)) ds]^{-1/2}$,

et cet équivalent est uniforme pour $(x, y) \in U \times U$.

4. Retour à la variété.

4.1. LEMME. Soit M une variété et $\mathcal{P}$ une diffusion sur M de générateur différentiel elliptique à coefficients localement höldériens. On considère sur M deux ouverts V et W tels que $\bar{W} \subset V$ et tels que la frontière ∂W soit compacte. Si π est un volume riemannien, on désigne par p (respectivement p_V) la version continue de la densité

par rapport à π du semi-groupe de transition de $\mathcal{P}$ (respectivement du processus induit $\mathcal{P}_V$).

Alors pour tout $t_0 \in \mathbb{R}^+$, il existe une constante réelle positive B telle que l'on ait

$$\forall (t, x, y) \in [0, t_0] \times W \times W$$

$$(7) \quad 0 \leq 1 - \frac{p_V(t, x, y)}{p(t, x, y)} \leq B \sup_{\substack{0 \leq s \leq t \\ z \in \partial W}} \frac{p_V(s, z, y)}{p(t, x, y)}$$

Preuve : Elle découle immédiatement de la relation entre p et p_V obtenue par R. Azencott (ch.II, corollaire 4.3). En effet on a, pour tout $(t, x, y) \in \mathbb{R}^+ \times M \times W$,

$$p(t, x, y) = I_V(x) p_V(t, x, y) + \int_{[0, t] \times \partial W} p_V(t-s, z, y) d\nu_x(s, z)$$

où ν_x est une mesure de Radon positive sur $\mathbb{R}^+ \times \partial W$ telle que pour tout compact K de $\mathbb{R}^+ \times \partial W$, $\sup_{x \in M} \nu_x(K)$ soit fini.

On en déduit alors que, pour tout $(t, x, y) \in [0, t_0] \times W \times W$,

$$p_V(t, x, y) \leq p(t, x, y) \leq B \sup_{\substack{0 \leq s \leq t \\ z \in \partial W}} p_V(s, z, y)$$

où $B = \sup_{x \in M} \nu_x([0, t_0] \times \partial W)$. D'où le lemme.

A l'aide de ce lemme, on va pouvoir prouver selon le schéma donné en 2.3 l'équivalence de q et de q_V pour un certain V , puis de p_V et de p .

4.2. Equivalence de q et q_V .

Selon les notations du paragraphe 2, à un point x_0 quelconque de M , on a associé un ouvert U contenant x_0 , tel que, pour tout x et y de U , il existe une carte semi-géodésique $\phi_{x,y} : U \rightarrow \mathbb{R}^m$. Nous allons montrer qu'il existe deux boules ouvertes W_1 et V centrées en x_0 vérifiant $\bar{W}_1 \subset V \subset \bar{V} \subset U$, et trois réels positifs t_0, c, δ tels que, pour tout $(t, x, y) \in]0, t_0] \times W_1 \times W_1$, on ait

$$(8) \quad \left| \frac{q_V(t, \phi_{x,y}(x), \phi_{x,y}(y))}{q(t, \phi_{x,y}(x), \phi_{x,y}(y))} - 1 \right| \leq C \exp\left(-\frac{\delta}{t}\right).$$

D'après le lemme 4.1, il s'agit de majorer $q_V(s, \phi_{x,y}(z), \phi_{x,y}(y))$ pour $(s, z) \in]0, t_0] \times \partial W$, W étant un ouvert de frontière ∂W compacte et vérifiant $\bar{W} \subset V$, et de minorer $q(t, \phi_{x,y}(x), \phi_{x,y}(y))$.

Cette dernière minoration s'obtient immédiatement, à l'aide de l'équivalent de $q(t, \phi_{x,y}(x), \phi_{x,y}(y)) = q(t, 0, e_1)$ obtenue en 3.6, et pour $\varepsilon > 0$ donné, il existe donc, si V est une boule ouverte de centre x_0 telle que $\bar{V} \subset U$, deux constantes réelles positives t_1 et C_1 telles que, pour tout $(t, x, y) \in]0, t_1] \times V \times V$, on ait,

$$q(t, \phi_{x,y}(x), \phi_{x,y}(y)) \geq C_1 \exp\left[\frac{-d^2(x, y) - \varepsilon}{2t}\right].$$

En outre, il découle du théorème 4.2 du chapitre 4 (P. Baldi), que l'on connaît un majorant de $q(t, v, w), v, w \in \mathbb{R}^m$. Plus précisément, il existe deux constantes réelles positives C_2 et C_3 indépendantes de $(x, y) \in U \times U$ et un réel positif t_2 tels que, pour tout $(s, v, w) \in]0, t_2] \times \mathbb{R}^m \times \mathbb{R}^m$, on ait

$$q(s, v, w) \leq C_2 s^{-\frac{m}{2}} \exp\left[\frac{-C_3 |v-w|^2}{s}\right].$$

Par suite si W est une boule ouverte de M de centre x_0 telle que $\bar{W} \subset V$, et si note λ un minorant des valeurs propres de g^{-1} sur U , on en déduit que, pour $s \leq t_2$ et $(x, y) \in V \times V$,

$$\begin{aligned} \sup_{z \in W} q(s, \phi_{x,y}(z), \phi_{x,y}(y)) &\leq \sup_{z \in W} C_2 s^{-\frac{m}{2}} \exp\left[\frac{-\lambda C_3 d(z, y)^2}{s}\right] \\ &\leq C_2 s^{-\frac{m}{2}} \exp\left[\frac{-\lambda C_3 d(y, W)^2}{s}\right]. \end{aligned}$$

En conséquence, pour $\varepsilon > 0$ donné, il existe un réel > 0 suffisamment petit t_3 et une constante réelle positive C_4 tels que, pour $s \leq t_3$,

$$\sup_{z \in W} q(s, \phi_{x,y}(z), \phi_{x,y}(y)) \leq C_4 \exp\left[\frac{-\lambda C_3 d(y, W)^2 + \varepsilon}{t}\right]$$

Supposons que le terme $-\lambda C_3 d(y, \partial W)^2 + \varepsilon$ soit négatif, ce qui impose en particulier à y de ne jamais appartenir à ∂W , alors, pour $t < t_3$,

$$\sup_{\substack{0 \leq s \leq t \\ z \in W}} q(s, \phi_{x,y}(z), \phi_{x,y}(y)) \leq C_4 \exp\left[\frac{-\lambda C_3 d(y, \partial W)^2 + \varepsilon}{t}\right]$$

et par suite, si $t \leq \inf(t_1, t_3)$,

$$\sup_{\substack{0 \leq s \leq t \\ z \in \partial W}} \frac{q_V(s, \phi_{x,y}(z), \phi_{x,y}(y))}{q(t, x, y)} \leq \frac{C_4}{C_1} \exp\left[\frac{-2\lambda C_3 d(y, \partial W)^2 + d^2(x, y) + 3\varepsilon}{t}\right]$$

De ce fait, si on montre que pour x et y bien choisis, le terme $B(x, y) = -2\lambda C_3 d(y, W)^2 + d^2(x, y) + 3\varepsilon$ est majoré par un réel δ_1 négatif, la relation (8) sera prouvée pour $t \leq t_0 = \inf(t_1, t_3)$, $C = \frac{C_4}{C_1}$ et $\delta = -\delta_1$. On peut remarquer que si $B(x, y)$ est négatif, il en est de même du terme $-\lambda C_3 d(y, W)^2 + \varepsilon$ et la condition imposée plus haut est automatiquement vérifiée.

Notons 2β le rayon de la Boule ouverte V de centre x_0 et posons $W = \mathfrak{B}(x_0, \beta)$ boule ouverte de centre x_0 et de rayon β . Nous allons imposer à x et y d'appartenir à une boule $W_1 = \mathfrak{B}(x_0, \alpha\beta)$ strictement incluse dans W , c'est-à-dire que $\alpha < 1$. Alors $d(y, \partial W) \geq \beta(1-\alpha)$.

$$D'où $B(x, y) \leq -2\lambda C_3 \beta^2(1-\alpha)^2 + 4\alpha^2\beta^2 + 3\varepsilon$.$$

On peut alors choisir α suffisamment petit pour que $\delta_2 = -2\lambda C_3 \beta^2(1-\alpha)^2 + 4\alpha^2\beta^2$ soit strictement négatif, puis ε petit pour que $\delta_2 = \delta_2 + 3\varepsilon$ soit encore négatif. Alors la relation (8) est vérifiée pour x et y appartenant à $W_1 = \mathfrak{B}(x_0, \alpha\beta)$ et pour $t \leq t_0 = \inf(t_1, t_3)$, t_0 dépendant de ε .

4.3. Conséquence. Il découle de (8) que, uniformément pour $(x, y) \in W_1 \times W_1$,

$$q_V(t, \phi_{x,y}(x), \phi_{x,y}(y)) \sim q(t, \phi_{x,y}(x), \phi_{x,y}(y)) \text{ lorsque } t \rightarrow 0.$$

De plus, si p_V est la densité de transition par rapport au volume riemannien de $\mathfrak{P}_V$, processus induit du processus $\mathfrak{P}$ sur M de générateur différentiel L , on a vu (2.2) que

$$p_V(t, x, y) = \det g(y)^{-\frac{1}{2}} q_V(t, \phi_{x,y}(x), \phi_{x,y}(y))$$

Il résulte donc de l'équivalent de q obtenu en 3.6 que, uniformément pour $(x, y) \in W_1 \times W_1$, lorsque $t \rightarrow 0$,

$$p_V(t, x, y) \sim (2\pi t)^{-m/2} H(x, y) \exp\left[\frac{-d^2(x, y)}{2t} + A(x, y)\right]$$

avec $H(x, y) = \det g(y)^{-\frac{1}{2}} \bar{H}(x, y) = [\det g^{-1}(y) \det g^{-1}(x)]^{1/4} [\det \int_0^1 g^{-1}(\gamma(s)) ds]^{-1/2}$,
et $A(x, y) = \int_0^1 \langle Y(\gamma(s)), \dot{\gamma}(s) \rangle ds$.

Si nous montrons que, pour t petit, $p(t,x,y)$ est équivalent à $p_V(t,x,y)$, la proposition 1.4 sera prouvée.

4.4. Equivalence de p et p_V

Des arguments analogues à ceux utilisés ci-dessus permettent de prouver l'existence d'une boule ouverte W_2 de centre x_0 vérifiant $\bar{W}_2 \subset W_1$ et de trois réels positifs t'_0 , c' , δ' tels que, pour tout $(t,x,y) \in]0,t'_0] \times W_2 \times W_2$,

on ait

$$\frac{p_V(t,x,y)}{p(t,x,y)} - 1 \leq c' \exp\left(-\frac{\delta'}{t}\right).$$

La démonstration de la proposition est alors achevée.

BIBLIOGRAPHIE

- [1] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press (1978).
- [2] YU.I. Kifer, *On the asymptotics of the transition density of process with small diffusion*. th. of Probability, vol.21 n°3 (1976).
- [3] S.A. Molchanov, *Diffusion processes and riemannian geometry*. Russ. Math. Surveys 30 : 1 (1975) p.1 à 63.
- [4] S.R.S. Varadhan, *Diffusion Processes in a Small Time Interval*. Comm. on Pure and Applied Math. vol XX (1967) p.659-685.

CHAPITRE 6

CHAMPS DE JACOBI ET DÉRIVÉE DE L'APPLICATION EXPONENTIELLE

par

Jacques GRANARA

Le but du présent chapitre est de préciser quelques notions de géométrie riemannienne ; cela permet d'abord de définir la notion de coordonnées semi-géodésiques en fonction de conditions initiales ; ensuite de connaître leur comportement lors de variations de ces conditions dans de petits ouverts. Ces coordonnées sont utilisées dans le chapitre 5 pour obtenir un équivalent de $p(t,x,y)$; y intervient le déterminant de l'exponentielle à origine fixe : on le calcule ici à l'aide des champs de Jacobi, puis on en donne une interprétation géométrique intrinsèque.

1. Variations par géodésiques et champs de Jacobi

Considérons une variété riemannienne M de classe C^k avec $k \geq 4$; la variété différentielle sous-jacente est de classe C^{k+1} . Soit γ un chemin de classe C^n ($n \geq 1$) d'un intervalle I de $\mathbb{R}$ dans M . Soit enfin $\mathcal{V}$ un voisinage ouvert connexe de l'origine dans $\mathbb{R}^l$.

1.1. DEFINITION. Une l -variation de γ est une C^r -application α de $I \times \mathcal{V}$ dans M telle que $\alpha(.,0) = \gamma$ où $r \geq 1$.

On précisera que α est une variation par géodésiques lorsque, pour tout $u \in \mathcal{V}$, $\alpha(.,u)$ est une géodésique (en particulier γ) ; notons au passage que les géodésiques sont de classe C^{k+1} .

1.2. THÉORÈME. Soient γ une géodésique de M et α une ℓ -variation de γ de classe C^r avec $r \geq 3$; α est une variation par géodésiques si et seulement si pour tout i ,

$$\nabla_{\frac{\partial \alpha}{\partial \tau}} \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_i} = R\left(\frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial u_i}\right) \frac{\partial \alpha}{\partial \tau},$$

où R est le tenseur de courbure de M , conformément aux notations de Helgason, et en opposition avec celle de Milnor : X, Y , étant des champs de classe C^1 , Z un champ de classe C^2 , on a $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{(X, Y)} Z$; au point a cette quantité ne dépend que de X_a, Y_a et Z_a . Si x est une coordonnée de $\mathcal{V} \times I$ (u_i ou τ) $\frac{\partial \alpha}{\partial x}$ désigne $d\alpha \cdot \frac{\partial}{\partial x}$.

Démonstration. D'après la définition 1.1, α est une variation par géodésiques si et seulement si $\nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial \tau} = 0$ or $\nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial \tau} (0, \cdot) = 0$; donc une condition équivalente est que pour tout i , $\nabla_{\frac{\partial \alpha}{\partial u_i}} \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial \tau} = 0$, ou encore que pour tout i ,

$$\nabla_{\frac{\partial \alpha}{\partial \tau}} \nabla_{\frac{\partial \alpha}{\partial u_i}} \frac{\partial \alpha}{\partial \tau} + R\left(\frac{\partial \alpha}{\partial u_i}, \frac{\partial \alpha}{\partial \tau}\right) \frac{\partial \alpha}{\partial \tau} = 0, \text{ ou enfin que pour tout } i,$$

$$\nabla_{\frac{\partial \alpha}{\partial \tau}} \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_i} + R\left(\frac{\partial \alpha}{\partial u_i}, \frac{\partial \alpha}{\partial \tau}\right) \frac{\partial \alpha}{\partial \tau} = 0.$$

Ce résultat conduit à la

1.3. DÉFINITION. Soit γ une géodésique de M et W de champ le long de γ . On dit que W est un champ de Jacobi lorsque $\nabla_{\frac{d\gamma}{dt}} \nabla_{\frac{d\gamma}{dt}} W = R\left(\frac{d\gamma}{dt}, W\right) \frac{d\gamma}{dt}$.

Pour d'évidentes raisons de commodité, ϕ étant un chemin C^r de M , de domaine I , f étant une fonction dérivable sur I , X étant un champ de vecteurs dérivable le long de ϕ , nous conviendrons de noter $\dot{\phi}, \dot{f}, \dot{X}, \ddot{\phi}, \dots$ les dérivées respectives $d\phi, df, \nabla_{d\phi} X, \nabla_{d\phi} d\phi, \dots$ et ainsi de suite pour les dérivées d'ordre supérieur ; notons au passage qu'ici $d\phi = d\phi$. $\frac{d}{dt} = \frac{d\phi}{dt}$ et $df = \frac{df}{dt}$. L'équation définissant les champs de Jacobi s'appelle "l'équation de Jacobi" ; avec les présentes notations, elle devient $\ddot{W} = R(\dot{\gamma}, W)\dot{\gamma}$. Cette équation peut être exprimée sur les composantes des champs dans une base orthonormée de champs parallèles le long de γ : ce peut être utile dans le cas d'espaces symétriques. Notons que les champs de Jacobi sont de classe C^k .

1.4. PROPOSITIONS. M étant une C^k variété riemannienne (selon la convention du début de ce paragraphe) de dimension m , soit γ une géodésique de M :

a) L'ensemble des champs de Jacobi le long de γ est un espace vectoriel de dimension $2m$, un tel champ W étant déterminé par ses valeurs initiales $W(\tau_0), \dot{W}(\tau_0)$

arbitraires dans $T_{\gamma(\tau_0)}M$.

b) W est partout normal à γ (orthogonal à $\dot{\gamma}$) si et seulement si $W(\tau_0)$ et $\dot{W}(\tau_0)$ sont normaux à γ .

c) L'espace des champs de Jacobi normaux à γ est de dimension $2(m-1)$.

d) Tout champ de Jacobi tangent à γ est de la forme $(A+B)\dot{\gamma}(\tau)$ où A, B sont des constantes. (démonstrations triviales).

Il importe de voir que tout champ de Jacobi dérive d'une variation par géodésiques. Plus généralement, I étant un intervalle compact, γ une géodésique $I \rightarrow M$ et $W_1, \dots, W_l$ des champs C^k le long de γ ; on a :

1.7. PROPOSITION. $W_1, \dots, W_l$ sont des champs de Jacobi si et seulement si il existe une variation par géodésiques α de γ telle que pour tout i

$$\frac{\partial \alpha}{\partial u_i}(\cdot, 0) = W_i$$

Démonstration. La partie directe est conséquence immédiate de (1.2); montrons la réciproque: pour cela choisissons un élément s de I et appelons β l'application de B dans M telle que $\beta(u) = \exp(\sum u_i W_i(s))$ où B est une boule de centre O de $\mathbb{R}^l$ et de rayon assez petit pour que β ait un sens; β paramètre une surface dégénérée ou non qui nous servira de surface de départ de la variation.

Soit z un champ le long de β tel que $z(0) = \dot{\gamma}(s)$ et $\nabla_{\frac{\partial \beta}{\partial u_i}(0)} z = \dot{W}_i(s)$. Posons

$\alpha(\tau, u) = \exp((\tau-s)z(u))$ pour $u \in [-\varepsilon, \varepsilon]^l$ et $\tau \in I$; cela est possible pour ε suffisamment petit car I est compact. Il est clair que :

1) α est une variation par géodésiques de γ .

$$2) \frac{\partial \alpha}{\partial u_i}(s, 0) = \frac{\partial \beta}{\partial u_i}(0) = W_i(s)$$

$$3) \alpha(\tau, 0) = \gamma(\tau) (= \exp((\tau-s)\dot{\gamma}(s)).)$$

$$4) \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_i}(s, 0) = \nabla_{\frac{\partial \alpha}{\partial u_i}} \frac{\partial \alpha}{\partial \tau}(s, 0) = \nabla_{\frac{\partial \alpha}{\partial u_i}} z = \dot{W}_i(s)$$

5) d'après 1), 2) et 4), $\frac{\partial \alpha}{\partial u_i}(\cdot, 0)$ le champ de Jacobi de valeurs initiales $W_i(s)$ et $\dot{W}_i(s)$: donc $\frac{\partial \alpha}{\partial u_i}(\cdot, 0) = W_i$.

2. Dérivée de l'exponentielle

Soient M une variété riemannienne, a un point de M , x un vecteur de $T_a M$ et y

un vecteur de $T_x(T_a M)$. L'application α de $[-\varepsilon, \varepsilon] \times [0, 1]$ dans M définie par $\alpha(\tau, u) = \exp \tau(x+uy)$, où ε est assez petit est une variation par géodésiques : $\alpha(., u)$ est la géodésique issue de a et tangente à $x + uy$ (élément de l'espace vectoriel $T_a M$) donc $\frac{D}{du} \frac{\partial \alpha}{\partial \tau}(0, 0) = \nabla_{\frac{\partial \alpha}{\partial u}} \frac{\partial \alpha}{\partial \tau}(0, 0) = y$. Ainsi $\frac{\partial \alpha}{\partial u}(., 0)$ est le champ de Jacobi W le long de $\alpha(., 0)$ tel que $W(0) = 0$ et $W'(0) = y$. Par ailleurs $\alpha(1, u) = \exp(x+uy)$ donc $\frac{\partial \alpha}{\partial u}(1, 0) = \exp'_x y$.

Soit $\exp'_x y = W'(1)$.

3. Coordonnées semi-géodésiques

Considérons une variété riemannienne M de classe C^k ($k \geq 4$), une géodésique γ de M injective et une hypersurface S de M non dégénérée et normale à γ en $\gamma(s)$. Soit $\mathcal{O}$ un voisinage ouvert de γ (dans M). Supposons donné un système de coordonnées C^{k-1} de $\mathcal{O}$, notées $\tau, u_1, \dots, u_{m-1}$. Notons α la carte ainsi définie.

3.1. DÉFINITION. On dit que les coordonnées sont semi-géodésiques autour de γ lorsque sont réalisées les conditions suivantes :

- 1) pour tout élément τ de I , domaine de γ , $\gamma(\tau) = \alpha(\tau, 0)$.
- 2) pour tout u , $\alpha(., u)$ est une géodésique.
- 3) $\|\frac{\partial \alpha}{\partial \tau}\|$ est constant (ne dépend ni de τ ni de u).
- 4) $\alpha(s, .)$ est une immersion d'image $S \cap \mathcal{O}$
- 5) $\frac{\partial \alpha}{\partial \tau}(s, .)$ est normal à $S \cap \mathcal{O}$ en $\alpha(s, .)$.

On voit que, si le domaine de α est de la forme $J \times \mathcal{V}$, où $\mathcal{V}$ est un voisinage ouvert connexe de 0 dans $\mathbb{R}^{m-1}$ et J un intervalle ouvert, α est une variation par géodésiques pour un prolongement de γ (conditions 1) et 2)). Si I est compact, on pourra se ramener à ce cas en diminuant $\mathcal{O}$. Dans ce cas on a de surcroît :

$$\left\langle \frac{\partial \alpha}{\partial u_i}, \frac{\partial \alpha}{\partial \tau} \right\rangle = 0 \quad (i=1, \dots, m-1)$$

$$\text{preuve : } \frac{\partial}{\partial \tau} \left\langle \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial u_i} \right\rangle = \left\langle \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial u_i} \right\rangle + \left\langle \frac{\partial \alpha}{\partial \tau}, \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_i} \right\rangle$$

$$\nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial \tau} = 0 \text{ car les } \alpha(., u) \text{ sont géodésiques}$$

$$\left\langle \frac{\partial \alpha}{\partial \tau}, \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_i} \right\rangle = \left\langle \frac{\partial \alpha}{\partial \tau}, \nabla_{\frac{\partial \alpha}{\partial u_i}} \frac{\partial \alpha}{\partial \tau} \right\rangle = \frac{1}{2} \frac{\partial}{\partial u_i} \left\langle \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial \tau} \right\rangle = 0$$

car $\left\langle \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial \tau} \right\rangle$ est constant (condition 3))

donc $\frac{\partial}{\partial \tau} \left\langle \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial u_i} \right\rangle = 0$; comme on a aussi $\left\langle \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial u_i} \right\rangle (s, u) = 0$ et que le domaine

est de la forme $\mathcal{V} \times J$, $\langle \frac{\partial \alpha}{\partial \tau}, \frac{\partial \alpha}{\partial u_i} \rangle$ est partout nul. Notons que les hypersurfaces $\alpha(\tau, \cdot)$ sont ici parallèles, au sens où sont parallèles des sphères concentriques. Voyons comment on peut construire une telle application α ; l'analyse nous fournit la

3.2. PROPOSITION. Soient M une C^k -variété riemannienne de dimension m (cf. §1), une géodésique $\gamma : J \rightarrow M$ et $\alpha : J \times \mathcal{V} \rightarrow M$ vérifiant les conditions de la définition 3.1. Si $W_i = \frac{\partial \alpha}{\partial u_i}(\cdot, 0)$, W_i est normal à γ et $\langle W_i, \dot{W}_j \rangle = \langle W_j, \dot{W}_i \rangle$.

Preuve : $\langle \frac{\partial \alpha}{\partial u_i}, \frac{\partial \alpha}{\partial \tau} \rangle = 0$: en particulier $\langle W_i, \dot{\gamma} \rangle = 0$. D'autre part

$$\frac{\partial}{\partial u_j} \langle \frac{\partial \alpha}{\partial u_i}, \frac{\partial \alpha}{\partial \tau} \rangle = 0 \text{ soit encore : } \langle \nabla_{\frac{\partial \alpha}{\partial u_j}} \frac{\partial \alpha}{\partial u_i}, \frac{\partial \alpha}{\partial \tau} \rangle + \langle \frac{\partial \alpha}{\partial u_i}, \nabla_{\frac{\partial \alpha}{\partial u_j}} \frac{\partial \alpha}{\partial \tau} \rangle = 0 ; \text{ on}$$

échange les dérivations dans chaque terme, on élimine les premiers termes ce qui donne $\langle \frac{\partial \alpha}{\partial u_i}, \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_j} \rangle = \langle \frac{\partial \alpha}{\partial u_j}, \nabla_{\frac{\partial \alpha}{\partial \tau}} \frac{\partial \alpha}{\partial u_i} \rangle$.

Donnons nous donc une C^k -variété riemannienne M de dimension m , une géodésique γ de domaine $[0, s]$ compact et des champs de Jacobi $W_1, \dots, W_{m-1}$ normaux à γ tels que, pour tout τ , $W_1(\tau), \dots, W_{m-1}(\tau)$ soient indépendants (ils constituent donc une base de l'espace normal à γ en $\gamma(\tau)$) ; supposons vérifiées pour tous i, j les conditions $\langle W_i, \dot{W}_j \rangle = \langle W_j, \dot{W}_i \rangle$. Donnons-nous enfin un nombre $\varepsilon > 0$ assez petit pour pouvoir poser, pour tout u dans $] -\varepsilon, \varepsilon[^{m-1}$,

$$\beta(u) = \exp\left(\sum_i u_i W_i(0) - \frac{1}{2} \left[\sum_{i,j} u_i u_j \langle W_i(0), \dot{W}_j(0) \rangle \right] \frac{\dot{\gamma}(0)}{\|\dot{\gamma}(0)\|^2}\right).$$

Un tel ε peut être choisi assez petit pour que β soit une immersion. Remarquons que β est C^k ; on prendra ε assez petit pour que soit défini un difféomorphisme α de $] -\varepsilon, s + \varepsilon[\times] -\varepsilon, \varepsilon[^{m-1}$ par $\alpha(\tau, u) = \exp(\tau z(u))$ où z est l'unique champ continu normal à β , de norme $\|\dot{\gamma}(0)\|$, valant $\dot{\gamma}(0)$ en $\beta(0)$; ceci est possible grâce à la compacité de $[0, s]$. α est solution du problème dès que $k \geq 4$ comme le montre la

3.3. PROPOSITION. α est un C^{k-1} -difféomorphisme de $] -\varepsilon, s + \varepsilon[\times] -\varepsilon, \varepsilon[^{m-1}$ dans un voisinage ouvert de γ dans M ; les coordonnées ainsi définies sont semi-géodésiques autour de γ dès que $k \geq 4$.

Démonstration. Le champ z est de classe C^{k-1} ; le reste ne présente aucune difficulté.

Remarquons que $\gamma(0)$ n'a aucun point conjugué le long de γ prolongé à $] -\varepsilon, s + \varepsilon[$ (voir chapitre 7).

De plus α est entièrement déterminé par les données suivantes :

$\gamma(0)$, $\dot{\gamma}(0)$, s , $w_1(0), \dots, w_{m-1}(0)$, $\dot{w}_1(0), \dots, \dot{w}_{m-1}(0)$, ε ; si l'on fait varier ces données de façon C^r ($4 \leq r \leq k$) ; alors α est fonction C^{r-1} de ces données donc C^3 au moins.

Par ailleurs, α étant au moins C^3 , l'expression de g dans α est au moins C^2 .

Dans une optique de cartes on a le

3.4. THÉORÈME. *On peut recouvrir M par des ouverts U munis chacun d'une application ψ de $TU \times U$ dans $\mathbb{R}^m$ telle que les applications $u \mapsto \psi((x, X), u)$ (notée $\psi_{x, X}$) sont des cartes semi-géodésiques sur U de classe C^{k-1} , $\psi_{x, X}$ vérifiant, avec les notations de la proposition 3.3 $\gamma(0) = x$, $\dot{\gamma}(0) = X$, et les dérivées partielles de $\psi_{x, X}$ par rapport à $u = (u_1, \dots, u_{m-1})$ sont, jusqu'à l'ordre $k-1$, uniformément bornées en $((x, X), u)$ sur $B \times U$ où $B = \{(x, X) \mid x \in U, X \in T_x M, \|X\| \leq 1\}$.*

Preuve : il suffit de choisir pour x_0 fixé, un ouvert relativement compact U possédant x_0 ; le reste s'ensuit trivialement par argument de compacité de $\bar{U}$ et des domaines des paramètres auxiliaires figurant au 3.3. (ces domaines pouvant être rendus arbitrairement petits.

Notons que si m est pair, on peut rendre ψ de classe C^{k-1} globalement (par rapport à $((x, X), u)$) ; pour m impair, ψ ne peut être continue : on s'en convainc aisément en prenant pour M l'espace euclidien $\mathbb{R}^3$.

3.5. THÉORÈME. *On peut recouvrir M par des ouverts U munis chacun d'une application ϕ de U^3 dans $\mathbb{R}^m$ telle que l'application $u \mapsto \phi(x, y, u)$, notée $\phi_{x, y}$ est une carte semi-géodésique de classe C^{k-1} vérifiant, avec les notations de la proposition 3.3 :*

$\gamma(0) = x$, $\gamma(1) = y$, et les dérivées partielles de $\phi_{x, y}$ par rapport aux composantes de $(u_1, \dots, u_{m-1})$ sont, jusqu'à l'ordre $k-1$ uniformément bornées sur U^3 .

Preuve : on se ramène au théorème précédent pour U assez petit en déterminant X tel que $\dot{\gamma}(0) = X$ et γ soit la géodésique minimale allant de x à y ; les cartes sont alors les mêmes, quitte à diminuer uniformément l'ouvert U .

4. Expression des champs de Jacobi normaux le long de γ lorsqu'est donné un système de coordonnées semi-géodésiques autour de γ

Donnons-nous une variété riemannienne M de classe C^k et une géodésique γ de M . Soient X, Y deux champs de Jacobi le long de γ ; on a :

4.1. LEMME. *La différence $\langle \dot{X}, Y \rangle - \langle X, \dot{Y} \rangle$ est constante le long de γ*

Preuve : $\frac{d}{d\tau} \langle \dot{X}, Y \rangle = \langle \ddot{X}, Y \rangle + \langle \dot{X}, \dot{Y} \rangle = \langle R(\dot{Y}, X) \dot{Y}, Y \rangle + \langle \dot{X}, \dot{Y} \rangle$ or $\langle R(\dot{Y}, X) \dot{Y}, Y \rangle = \langle R(\dot{Y}, Y) \dot{Y}, X \rangle$ car R a la propriété de symétrie suivante dont le lecteur se convaincra aisément en se plaçant dans une carte

$$\langle R(U, X) V, Y \rangle = \langle R(V, Y) U, X \rangle$$

$$\text{donc } \frac{d}{d\tau} \langle \dot{X}, Y \rangle = \frac{d}{d\tau} \langle \dot{Y}, X \rangle$$

Soit α un système de coordonnées semi-géodésiques autour de γ ; notons W_i le champ $\frac{\partial \alpha}{\partial u_i}(\cdot, 0)$: c'est un champ de Jacobi normal à γ , et les $W_i(\tau)$ constituent une base de l'espace normal à γ en $\gamma(\tau)$. Soit X un champ normal à γ ; décomposons-le sur cette base : $X = \sum_i x_i W_i$. X est-il un champ de Jacobi ? des équivalences sont données par la

4.2. PROPOSITION. *Les conditions suivantes sont équivalentes :*

- i) X est un champ de Jacobi
- ii) pour tout j , $\langle \dot{X}, W_j \rangle - \langle X, \dot{W}_j \rangle$ est constant
- iii) pour tout j , $\sum_i \dot{x}_i \langle W_i, W_j \rangle$ est constant
- iv) pour tout j , $\frac{d}{d\tau} [\sum_i \dot{x}_i \langle W_i, W_j \rangle] = 0$

Preuve : l'équivalence (iii) $\iff$ (iv) étant triviale, montrons que

$$\langle \dot{X}, W_j \rangle - \langle X, \dot{W}_j \rangle = \sum_i \dot{x}_i \langle W_i, W_j \rangle$$

$$\dot{X} = \sum_i \dot{x}_i W_i + \sum_i x_i \dot{W}_i \text{ donc}$$

$$\langle \dot{X}, W_j \rangle = \sum_i \dot{x}_i \langle W_i, W_j \rangle + \sum_i x_i \langle \dot{W}_i, W_j \rangle$$

$$\langle X, \dot{W}_j \rangle = \sum_i x_i \langle W_i, \dot{W}_j \rangle = \sum_i x_i \langle \dot{W}_i, W_j \rangle$$

d'après 3, d'où le résultat par différence ; cela prouve que (ii) $\iff$ (iii)

(i) $\implies$ (ii) résulte de 4.1.

Il reste à montrer que (iv) $\implies$ (i) : les équations de (iv) forment un système différentiel linéaire d'ordre 1 ; cherchons-en les solutions : la première intégration ramène à (iii), soit :

$$\sum_i \dot{x}_i \langle W_i, W_j \rangle = C_j \quad (j=1, \dots, m-1)$$

Notons $\mathcal{M}(\tau)$ la matrice des $\langle W_i(\tau), W_j(\tau) \rangle$: elle est définie positive car les $W_i(\tau)$ sont indépendants ; les solutions sont donc de la forme :

$$x_i(\tau) = x_i(0) + \sum_j \int_0^\tau [\mathcal{M}(s)^{-1}]_{ij} C_j ds$$

Les solutions forment un espace vectoriel de dimension $2(m-1)$ et tout champ de Jacobi normal est solution d'après le schéma (i) $\implies$ (ii) $\iff$ (iii) $\iff$ (iv). Comme l'espace des champs de Jacobi normaux est de même dimension, il y a coïncidence des espaces, d'où le résultat. Remarquons au passage que, de ce fait, le système (iv) exprime l'équation de Jacobi en coordonnées semi-géodésiques autour de γ .
Autre remarque : si $W_i(0)$ est orthonormée, $\dot{x}_i(0) = C_i$.

5. Expression en coordonnées semi-géodésiques du Jacobien de $\exp_a$ en x , pour un point a de M et un point x de $T_a M$.

Soit $(z_1, \dots, z_m)$ une base orthonormée de $T_x(T_a M)$ assimilé naturellement à $T_a M$, avec $x = \|x\| z_m$. Posons $\phi(\tau, u_1, \dots, u_{m-1}) = \tau(x + u_1 z_1 + \dots + u_{m-1} z_{m-1})$ et $\alpha(\tau, u) = \exp \phi(\tau, u)$; considérons la géodésique de M $\alpha(\cdot, 0)$ notée ici γ et dont α est une variation par géodésiques ; on a : $\frac{\partial \phi}{\partial u_i}(\tau, 0) = \tau z_i$ et $\phi(\tau, 0) = x$ donc

$$\frac{\partial \alpha}{\partial u_i}(\tau, 0) = [\exp_a]_x' \cdot \tau z_i ; \text{ notons } Y_i \text{ le champ de Jacobi } \frac{\partial \alpha}{\partial u_i}(\cdot, 0) \text{ (le long de } \gamma) :$$

$$Y_i(0) = [\exp_a]_0' \cdot 0 z_i = 0 ; \text{ calculons } \dot{Y}_i(0)$$

$$Y_i'(0) = \nabla_{\frac{\partial}{\partial \tau}} \frac{\partial \alpha}{\partial u_i}(0, 0) = \nabla_{\frac{\partial}{\partial \tau}} \frac{\partial \alpha}{\partial \tau}(0, 0) ; \text{ or } \frac{\partial \alpha}{\partial \tau}(0, u) \in T_a M \text{ qui est un espace vectoriel ;}$$

$$\text{donc } \nabla_{\frac{\partial}{\partial u_i}} \frac{\partial \alpha}{\partial \tau}(0, 0) = \frac{\partial}{\partial u_i} \frac{\partial \alpha}{\partial \tau}(0, 0).$$

$$\frac{\partial \alpha}{\partial \tau}(0, u) = [\exp_a]_0' \cdot \frac{\partial \phi}{\partial \tau}(0, u) = [\exp_a]_0' \cdot (x + u_1 z_1 + \dots + u_{m-1} z_{m-1})$$

$$\frac{\partial \alpha}{\partial \tau}(0, u) = x + u_1 z_1 + \dots + u_{m-1} z_{m-1} \text{ car } [\exp_a]_0' = \text{id}_{T_a M} \text{ donc } \frac{\partial}{\partial u_i} \frac{\partial \alpha}{\partial \tau}(0, u) = z_i \text{ soit}$$

$$\dot{Y}_i(0) = z_i.$$

Ainsi Y_i est le champ de Jacobi le long de γ déterminé par les conditions initiales $Y_i(0) = 0$ et $\dot{Y}_i(0) = z_i$.

5.1. PROPOSITION. Le Jacobien au point $(\tau, 0)$ de ϕ est $\|x\| \tau^{m-1}$ et celui de α au même point est $\|x\| [\det \mathcal{A}(\tau)]^{1/2}$ où $\mathcal{A}(\tau)$ est la matrice carrée symétrique constituée par les $\langle Y_i(\tau), Y_j(\tau) \rangle$. Le Jacobien de $\exp_\alpha$ au point $\phi(\tau, 0)$ est le rapport de ceux-ci, soit $\tau^{1-m} [\det \mathcal{A}(\tau)]^{1/2}$.

$$\text{Preuve : 1) pour } \phi : \frac{\partial \phi}{\partial u_i}(\tau, u) = \tau z_i \text{ et } \frac{\partial \phi}{\partial \tau}(\tau, u) = x + u_1 z_1 + \dots + u_{m-1} z_{m-1}$$

$$\frac{\partial \phi}{\partial \tau} \wedge \frac{\partial \phi}{\partial u_1} \wedge \dots \wedge \frac{\partial \phi}{\partial u_{m-1}} = x \wedge \tau z_1 \wedge \dots \wedge \tau z_{m-1} = \|x\| \tau^{m-1} z_m \wedge z_1 \wedge \dots \wedge z_{m-1}$$

$$\text{donc } \left\| \frac{\partial \phi}{\partial \tau} \wedge \frac{\partial \phi}{\partial u_1} \wedge \dots \wedge \frac{\partial \phi}{\partial u_{m-1}} \right\| = \|x\| \tau^{m-1} \quad (z_1, \dots, z_m \text{ étant orthonormée})$$

2) pour α :

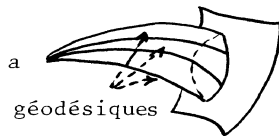
$$\frac{\partial \alpha}{\partial \tau}(\tau, 0) = \dot{\gamma}(\tau) \text{ et } \|\dot{\gamma}(\tau)\| = \|\dot{\gamma}(0)\| = \|x\|$$

$$\frac{\partial \alpha}{\partial u_i}(\tau, 0) = Y_i(\tau) \perp \dot{\gamma}(\tau) \text{ d'où le résultat.}$$

Soit à calculer $[\det \mathcal{A}]^{1/2}$ dans un système de coordonnées semi-géodésiques autour de γ ; notons β la carte de ce système, et W_i le champ de Jacobi $\frac{\partial \beta}{\partial u_i}(\cdot, 0)$; $(W_i(\tau))_i$ étant une base de l'espace à γ en $\gamma(\tau)$, décomposons les $Y_i(\tau)$ sur cette base : notons $y(\tau)$ la matrice exprimant les $Y_i(\tau)$ dans cette base ; on a $\mathcal{A}(\tau) = y(\tau)^* \mathcal{M}(\tau) y(\tau)$ avec la notation du 4 ($\mathcal{M}(\tau)$ est la matrice des $\langle W_i(\tau), W_j(\tau) \rangle$) donc $\det[\mathcal{A}(\tau)] = [\det \mathcal{M}(\tau)] [\det y(\tau)]^2$ soit $[\det \mathcal{A}(\tau)]^{1/2} = [\det \mathcal{M}(\tau)]^{1/2} |\det y(\tau)|$; d'après 4, et sachant que $y(0) = 0$, on a $y(\tau) = \left[\int_0^\tau \mathcal{M}^{-1}(s) ds \right] \mathcal{C}$ où $\mathcal{C}$ est constante et $\dot{y}(0) = \mathcal{M}^{-1}(0) \mathcal{C}$; comme $(\dot{Y}_1(0), \dots, \dot{Y}_{m-1}(0))$ est orthonormée $\dot{y}(0)^* \mathcal{M}(0) \dot{y}(0) = \text{Id}$ d'où, après un calcul élémentaire, $\det \mathcal{C} = [\det \mathcal{M}(0)]^{1/2}$; cela fournit le résultat suivant :

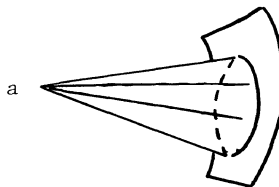
$$5.2. [\det \mathcal{A}(\tau)]^{1/2} = [\det \mathcal{M}(0)]^{1/2} [\det \mathcal{M}(\tau)]^{1/2} \left[\det \int_0^\tau \mathcal{M}^{-1}(s) ds \right]$$

Remarquons que $[\det \mathcal{A}(\tau)]^{1/2}$ est un rapport d'hypersurfaces infinitésimales : $[\det \mathcal{A}(\tau)]^{1/2} du_1 \dots du_{m-1}$ représente l'élément d'"aire" d'hypersphère éclairée dans M par l'élément d'angle solide (généralisé en dimension m) $\|x\|^{-m+1} du_1 \dots du_{m-1}$



← sphère de centre a
et de rayon $\|x\| \tau$
dans M

De même $\tau^{m-1} du_1 \dots du_{m-1}$ représente l'élément d'aire d'hypersphère éclairée dans $T_a M$ par le même élément d'angle solide.



← sphère de centre a
et de rayon $\|x\| \tau$
dans $T_a M$

6. Passage à la limite dans la formule du chapitre 5, intervenant dans l'équivalent de la densité $p(t,a,b)$.

Dans cet équivalent apparaît un facteur $H(a,b)$ valant

$[\det g(b) \cdot \det g(a)]^{-1/4} = [\det \int_0^1 g^{-1}(\gamma_{xy}(s)ds)]^{-1/2}$. Or, avec les présentes notations introduites au paragraphe 4, $g(\gamma_{xy}(s))$ n'est autre que $\mathcal{M}(s)$, d'où $g(a) = \mathcal{M}(0)$, $g(b) = \mathcal{M}(1)$, donc $H(a,b) = [(\det \mathcal{M}(0))^{1/2} (\det \mathcal{M}(1))^{1/2} \det \int_0^1 \mathcal{M}(s)^{-1} ds]^{-1/2}$ soit $H(a,b) = [\det \mathcal{M}(1)]^{-1/4}$, c'est-à-dire l'inverse de la racine carrée du jacobien de l'exponentielle de base a .

Que se passe-t-il lorsque l'on déforme la surface de départ $\beta(.,0)$? Les calculs gardent-ils un sens ? Le terme $H(a,b)$ ne dépendant pas de cette surface (normale à γ en $\gamma(0)$), le passage à la limite y est trivial même lorsqu'on fait dégénérer en $\gamma(0)$ les coordonnées semi-géodésiques.

BIBLIOGRAPHIE

- [1] J. Cheeger, D.G. Ebin, *Comparison Theorems in Riemannian Geometry*, North Holland Publishing Company, 1975.
- [2] J. Minor, *Morse Theory*, Annals of Math. Studies. Princeton University Press. 1963.
- [3] Helgason (Sigurdur), *Differential geometry and symmetric spaces*, New-york and London Academic Press. 1962.

CHAPITRE 7

PROPRIÉTÉS EXTRÊMALES DES GÉODÉSQUES

par

André BELLAÏCHE

1. Les géodésiques et l'énergie

1.1. Les géodésiques d'une variété riemannienne M ont été définies au ch. 3 (Ph. Bougerol) comme les solutions de l'équation différentielle du second ordre $\nabla_{\dot{c}} \dot{c} = 0$. Rappelons que cette équation exprime que le vecteur vitesse $\dot{c}$ est transporté parallèlement à lui-même le long de c .

On a aussi vu au ch. 3 les liens qui existent entre cette notion et celle de ligne de plus court chemin entre deux points. Et, de fait, c'est en raison de leurs propriétés de minimum que les géodésiques interviennent dans la théorie des diffusions. On verra au ch. 9 le rôle que jouent dans l'évaluation asymptotique de la densité $p(t, x, y)$ d'une diffusion les géodésiques de longueur minimale joignant x et y - dans le cas d'une diffusion à générateur elliptique, la métrique riemannienne étant celle définie au ch. 4. L'examen des démonstrations du ch. 9, ainsi que d'un autre côté celui des résultats de Ventsell et Freidlin [9] et d'Azencott [2], nous montre que l'intervention de ces géodésiques est surtout due au fait qu'elles réalisent le minimum de la fonctionnelle "énergie"

$$E(c) = \frac{1}{2} \int_0^1 \|\dot{c}(t)\|^2 dt,$$

parmi tous les chemins $c : [0, 1] \longrightarrow M$ d'origine x et d'extrémité y .

C'est ce qui justifie l'existence de ce long chapitre consacré à l'étude des géodésiques du point de vue du calcul des variations.

Voici une brève description de ce qu'il contient. On prouve d'abord, ce qui vient d'être affirmé : l'identité entre minimums de E et géodésiques minimisantes (1.7).

On introduit ensuite une structure de variété différentiable de dimension infinie sur l'espace Ω_{xy} des chemins absolument continus $c : [0,1] \longrightarrow M$ tels que $c(0) = x$, $c(1) = y$ et $E(c) < +\infty$. La fonctionnelle E devient une fonction différentiable sur Ω_{xy} et on retrouve les géodésiques non nécessairement minimisantes comme points critiques de E sur Ω_{xy} (2.3, 2.7, 3.2).

On peut alors établir un véritable dictionnaire reliant les propriétés géométriques des géodésiques joignant deux points x et y aux propriétés de la fonction énergie sur Ω_{xy} : le tableau suivant contient la plupart des résultats utiles et des définitions des §§1 à 5 et §7. On y suppose la variété M complète, ce qui assure l'existence dans Ω_{xy} d'au moins une géodésique minimisante (§6).

Propriété de la fonction $E : \Omega_{xy} \rightarrow \mathbb{R}$	Propriété géométrique équivalente	Réf
γ point critique de E	γ géodésique	3.3
γ minimum de E	γ géodésique de longueur $d(x,y)$	1.7
γ point critique non dégénéré de E	pour (x',y') voisin de (x,y) il existe une géodésique γ' joignant x' à y' t.q. l'application $(x',y') \mapsto \gamma'$ soit différentiable (stabilité du point critique)	$\Leftrightarrow y$ est valeur régulière de $\exp_x$ au point $(\gamma(0), \dot{\gamma}(0))$ 5.3 5.4
γ point critique dégénéré de E		on dit que x et y sont <u>conjugués</u> suivant γ . 5.1
E possède au moins un point critique dégénéré		on dit que x et y sont <u>conjugués</u> ; les points <u>conjugués</u> d'un point x donné forment un ensemble de mesure nulle 5.5

γ minimum unique et non dégénéré de E	pour (x', y') voisin de (x, y) il existe γ' géodésique minimisante joignant x' à y' t.q. l'application $(x', y') \mapsto \gamma'$ soit différentiable (stabilité du minimum)	$\Leftrightarrow \gamma$ est la seule géodésique minimisante joignant x et y et x et y ne sont pas conjugués suivant $\Leftrightarrow$ prolongée un peu au delà de y , γ est encore minimisante	voir 7.3 7.12
E atteint plusieurs fois son minimum ou possède un minimum dégénéré		on dit que (x, y) appartient au <u>cut-locus</u> de M ; le cut-locus est une partie fermée de mesure nulle de $M \times M$	7.1

On prouve ensuite (8.3), et ceci est absolument essentiel pour les applications qui suivront, que les assertions contenues dans ce tableau restent toutes valables si l'on y remplace la fonction E par sa restriction à la sous-variété de dimension finie $\Omega_{xy}^{n,h}$, formée de géodésiques brisées aux points $t_i = \frac{i}{n+1}$, $i = 1, \dots, n$ (cf. 8.2 et 8.3 pour la définition de $\Omega_{xy}^{n,h}$) et γ par un point de $\Omega_{xy}^{n,h}$.

1.2. Soit M une variété. Un chemin dans M est par définition une application continue d'un intervalle $[a, b]$ dans M. On notera $C(I, M)$ l'espace des chemins de M paramétrés par $I = [a, b]$ et on munira cet espace de la topologie de la convergence uniforme.

Supposons maintenant M de classe C^1 et munie d'une métrique riemannienne (de classe C^0). La longueur

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

d'un chemin $c : I \rightarrow M$ absolument continu est toujours finie. L'énergie de c est définie comme

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

et peut prendre la valeur $+\infty$. Pour $c \in C(I, M)$ non absolument continu on pose $E(c) = +\infty$.

1.4. On a entre énergie et longueur la relation

$$(1) \quad E(c) \geq \frac{L(c)^2}{2(b-a)}$$

qui découle de l'inégalité de Schwarz $(\int_a^b \|\dot{c}(t)\|^2 dt)^2 \leq \int_a^b \|\dot{c}(t)\|^2 dt \int_a^b dt$.

On a égalité dans (1) si et seulement si $\|\dot{c}\|$ est constant, c'est-à-dire si le paramètre t est proportionnel à la longueur d'arc.

1.5. La distance $d(x,y)$ de deux points x et y de M est égale à la borne inférieure des longueurs des chemins absolument continus joignant x et y (comparer avec la définition, ch.3, 1.1). Des exemples évidents montrent que cette borne inférieure n'est pas nécessairement atteinte. Pour qu'elle le soit il faut et il suffit - si la métrique est de classe C^1 - qu'il existe des géodésiques de longueur $d(x,y)$ joignant x et y (géodésiques minimisantes) et les chemins qui réalisent la distance $d(x,y)$ entre x et y sont alors ceux qui coïncident avec les géodésiques minimisantes au paramétrage près (ch.3, th.4.6).

1.6. Étant donnés deux points x et y de M et $t > 0$ on note $E(t,x,y)$ la borne inférieure de $E(c)$ quand c décrit l'ensemble des chemins $c : [0,t] \rightarrow M$ d'origine x et d'extrémité y (énergie minimale en temps t entre x et y).

On pose aussi $E(x,y) = E(1,x,y)$

1.7. PROPOSITION. Soit M une variété riemannienne de classe C^1 et soient x et y deux points de M . Pour $t > 0$ on a

$$(2) \quad E(t,x,y) = \frac{d^2(x,y)}{2t}.$$

Sur l'ensemble des chemins $c : [0,t] \rightarrow M$ joignant x à y la borne inférieure $E(t,x,y)$ est atteinte exactement pour ceux d'entre eux qui sont des géodésiques minimisantes.

Démonstration. Compte tenu du fait que tout chemin peut être reparamétrisé proportionnellement à la longueur la relation (2) est une conséquence de 1.4.

Les mêmes remarques nous montrent que pour réaliser le minimum de l'énergie entre x et y un chemin $c : [0,t] \rightarrow M$ doit être paramétrisé uniformément et réaliser le minimum de la longueur, c'est-à-dire (cf. 1.5) être une géodésique minimisante entre x et y (et que cela suffit).

2. La variété des chemins d'une variété

2.1. Soit M une variété de classe C^k ($k \geq 1$), de dimension m et soit $I = [a,b]$ un intervalle compact de $\mathbb{R}$.

Considérons sur M une métrique riemannienne g (de classe C^0 , au moins). On dit qu'un chemin $c : I \rightarrow M$ est de classe H^1 , ou encore d'énergie finie s'il est absolument continu et vérifie la condition

$$\int_a^b g(\dot{c}(t), \dot{c}(t)) dt < +\infty.$$

Cette notion est indépendante de la métrique riemannienne choisie sur M .

On notera $\Omega(I, M)$ l'espace des chemins d'énergie finie $c : I \rightarrow M$.

L'espace $\Omega(I, \mathbb{R}^m)$, en particulier, est un espace de Hilbert séparable pour la norme

$$\|c\|_H^2 = |c(a)|^2 + \int_a^b |\dot{c}(t)|^2 dt,$$

la notion $|\cdot|$ désignant dans le second membre la norme euclidienne de $\mathbb{R}^m$. La norme H^1 est plus fine que la norme de la convergence uniforme. On a, en effet, d'après l'inégalité de Schwarz

$$\|c\|_\infty \leq |c(a)| + \int_a^b |\dot{c}(t)| dt \leq |c(a)| + (b-a)^{1/2} \left(\int_a^b |\dot{c}(t)|^2 dt \right)^{1/2} \leq \text{const.} \|c\|_H.$$

Il s'ensuit que si U est un ouvert de $\mathbb{R}^m$ l'ensemble $\Omega(I, U)$ est ouvert dans $\Omega(I, \mathbb{R}^m)$.

2.2. Nous allons maintenant introduire sur $\Omega(I, M)$ une structure de variété différentielle. Rappelons qu'une variété différentielle de dimension infinie modelée sur un espace de Banach B se définit de la même façon qu'une variété de dimension n à la différence près que l'on remplace les cartes à valeurs dans $\mathbb{R}^n$ par des cartes à valeurs dans B . Un coup d'oeil sur Lang [7] ou Bourbaki [3] permet de se rendre compte que pour ce qui est des fondements de la théorie, il n'y a pas de différence essentielle entre les variétés de dimension finie et les variétés "banachiques". En particulier si B est un espace de Hilbert on n'a aucune précaution à prendre (grâce au fait que tout sous-espace vectoriel fermé de B admet un supplémentaire) et toutes les définitions habituelles ainsi que les théorèmes basés sur le théorème des fonctions implicites se transposent sans problème.

La notion de submersion, par exemple : soient X et Y deux variétés "hilbertiennes"

de classe C^r ($r \geq 0$). On dit qu'une application $f : X \rightarrow Y$ de classe C^r est une submersion si pour tout $x \in X$ il existe des cartes ϕ et ψ au voisinage de x et $f(x)$ telles que l'application $\psi \circ f \circ \phi^{-1}$ soit une projection $U \times V \rightarrow V$. Si f est une submersion, l'ensemble $f^{-1}(y)$ est une sous-variété de X , de classe C^r , quel que soit $y \in Y$. Si $r \geq 1$, pour que f soit une submersion il faut et il suffit que pour tout $x \in X$ l'application $df_x : T_x X \rightarrow T_{f(x)} Y$ soit surjective et l'espace tangent en x à la sous-variété $f^{-1}(y)$ (où $y = f(x)$) est dans ce cas le noyau de df_x .

2.3. PROPOSITION. *Soit M une variété de classe C^k ($k \geq 1$), de dimension finie m , et soit $I = [a, b]$ un intervalle compact de $\mathbb{R}$. L'espace $\Omega(I, M)$ peut être muni d'une structure naturelle de variété de classe C^{k-1} , modélisée sur un espace de Hilbert séparable.*

Remarque. Dans cet énoncé l'adjectif naturel a les sens de naturel, canonique et fonctoriel.

Démonstration. Pour chaque suite $\sigma = (t_i)_{0 \leq i \leq n}$ de points de I telle que $a = t_0 \leq t_1 \leq \dots \leq t_n = b$ et chaque suite $\theta = (U_i, \phi_i)_{1 \leq i \leq n}$ de cartes de M , soit $W_{\sigma, \theta}$ l'ensemble (éventuellement vide) des chemins $c \in C(I, M)$ tels que $c([t_{i-1}, t_i]) \subset U_i$ pour $1 \leq i \leq n$. Pour $c \in W_{\sigma, \theta}$ posons $c_i = c|_{[t_{i-1}, t_i]}$. Alors pour que $c \in \Omega(I, M)$ il faut et il suffit que pour $1 \leq i \leq n$ le chemin $\phi_i \circ c_i$ appartienne à $\Omega([t_{i-1}, t_i], \mathbb{R}^m)$. Posons alors $\Omega_{\sigma, \theta} = W_{\sigma, \theta} \cap \Omega(I, M)$ et soit

$$\theta_{\sigma, \theta} : \Omega_{\sigma, \theta} \rightarrow \Omega([t_0, t_1], \mathbb{R}^m) \times \Omega([t_{n-1}, t_n], \mathbb{R}^m)$$

l'application qui à c associe la suite $(\phi_1 \circ c_1, \dots, \phi_n \circ c_n)$.

Les couples $(\Omega_{\sigma, \theta}, \theta_{\sigma, \theta})$ obtenus de toutes les façons possibles forment sur $\Omega(I, M)$ un atlas de classe C^{k-1} . Pour le montrer il suffit de prouver que les changements de cartes sont de classe C^{k-1} , ce qui revient à prouver le lemme suivant :

soit $\phi : \mathbb{R}^m \rightarrow \mathbb{R}^p$ une application différentiable de classe C^k ($k \geq 1$) ; alors l'application $\phi_* : \Omega(I, \mathbb{R}^m) \rightarrow \Omega(I, \mathbb{R}^p)$ définie par $\phi_* c = \phi \circ c$ est de classe C^{k-1} (lemme de Palais).

Par récurrence on se ramène à prouver la continuité de ϕ_* lorsque $k = 1$. Supposons donc ϕ de classe C^1 .

Soit $c \in \Omega(I, \mathbb{R}^m)$. On doit montrer que $\phi \circ (c+h) \rightarrow \phi \circ c$ dans $\Omega(I, \mathbb{R}^p)$ lorsque $h \rightarrow 0$ dans $\Omega(I, \mathbb{R}^m)$. Observons d'abord que $\sup_{t \in I} |h(t)| \leq \text{const.} \|h\|_{H^1}$. Puisque ϕ est continue, et même uniformément continue sur un voisinage compact de $c(I)$, il en résulte que $\phi \circ (c+h) \rightarrow \phi \circ c$ uniformément lorsque $h \rightarrow 0$. En outre on a

$$\begin{aligned} \frac{d}{dt}(\phi \circ (c+h)) - \frac{d}{dt}(\phi \circ c) &= (\phi' \circ (c+h)) \cdot (\dot{c} + \dot{h}) - (\phi' \circ c) \cdot \dot{c} \\ &= (\phi' \circ (c+h) - \phi' \circ c) \cdot \dot{c} + (\phi' \circ (c+h)) \cdot \dot{h} \end{aligned}$$

d'où

$$\begin{aligned} \left\| \frac{d}{dt}(\phi \circ (c+h)) - \frac{d}{dt}(\phi \circ c) \right\|_{L^2(I, \mathbb{R}^p)} &\leq \sup_{t \in I} |\phi'(c(t)+h(t)) - \phi'(c(t))| \cdot \|\dot{c}\|_{L^2(I, \mathbb{R}^m)} \\ &\quad + \sup_{t \in I} |\phi'(c(t)+h(t))| \cdot \|\dot{h}\|_{L^2(I, \mathbb{R}^m)} \end{aligned}$$

en notant $|\cdot|$ la norme habituelle d'une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^p$. Comme ϕ' est continue, on voit que $\phi' \circ (c+h)$ tend vers $\phi' \circ c$ uniformément lorsque $h \rightarrow 0$. Comme d'autre part $\dot{h}$ tend vers 0 dans L^2 lorsque h tend vers zéro en norme H^1 la continuité de ϕ en découle.

Il reste à vérifier que la topologie définie (en même temps que la structure de variété) sur $\Omega(I, M)$ par l'atlas considéré est séparée. Pour cela il suffit de montrer que deux chemins distincts peuvent être considérés comme appartenant à des ouverts $W_{\sigma, \theta}$ disjoints, ce qui est évident (puisque M est séparée).

2.4. Soit $c \in \Omega(I, M)$ et soit V un champ de vecteurs le long de c (c'est-à-dire, rappelons-le, une application $V : I \rightarrow TM$ telle que $V(t) \in T_{c(t)}M$ pour tout $t \in I$). Soient $(t_i)_{0 \leq i \leq n}$ une suite de points de I et $(U_i, \phi_i)_{1 \leq i \leq n}$ une suite de cartes de M telles que

$$a = t_0 \leq t_1 \leq \dots \leq t_n = b \text{ et } c(t_{i-1}, t_i) \subset U_i \quad (1 \leq i \leq n).$$

Posons $V_i = V|_{[t_{i-1}, t_i]}$. On dit que V est un champ de vecteurs de classe H^1 le long de c si, pour $1 \leq i \leq n$, l'application $d\phi_i \circ V_i : [t_{i-1}, t_i] \rightarrow \mathbb{R}^m$ est de classe H^1 . Cette définition est indépendante de la subdivision de I et des cartes choisies.

2.5. PROPOSITION. (On suppose $k \geq 2$). L'espace tangent en c à $\Omega(I, M)$ s'identifie naturellement à l'espace des champs de vecteurs de classe H^1 le long de c .

Démonstration. Soit $c \in \Omega(I, M)$, soit ω un élément de $T_c \Omega(I, M)$ et soit $v :]-\varepsilon, \varepsilon[\rightarrow \Omega(I, M)$ une courbe C^1 telle que $v(0) = c$ et $\dot{v}(0) = \omega$. Associons à v l'application $\tilde{v} :]-\varepsilon, \varepsilon[\times I \rightarrow M$ définie par $\tilde{v}(s, t) = v(s)|_t$ (une telle application s'appelle d'habitude une variation à un paramètre du chemin c cf. ch.6, 1.1) et définissons un champ de vecteurs $\tilde{V}$ le long de c en posant

$$\tilde{V}(t) = \frac{\partial}{\partial s} v(s, t) \Big|_{s=0}.$$

On montre aisément que $\tilde{V}$ (qu'on appelle variation infinitésimale dérivée de la variation $\tilde{v}$) est un champ de vecteurs de classe H^1 et ne dépend que de ω . L'application $\omega \mapsto \tilde{V}$ est un isomorphisme de $T_c \Omega(I, M)$ sur l'espace des champs de vecteurs de classe H^1 le long de c et réalise l'identification annoncée.

On s'autorisera de la proposition 2.5. pour noter $T_c \Omega(I, M)$ l'espace des champs de vecteurs de classe H^1 le long de c . Remarquons que sur cet espace, il n'y a pas de produit scalaire canonique : la structure d'espace de Hilbert de $T_c \Omega(I, M)$ n'est définie qu'à isomorphisme linéaire continu près.

2.6. Soient x et y deux points de M . On notera $\Omega_{xy}(I, M)$ (ou brièvement Ω_{xy}) l'espace des chemins $c \in \Omega(I, M)$ d'origine x et d'extrémité y .

On notera encore $\pi : \Omega(I, M) \rightarrow M \times M$ l'application $c \mapsto (c(a), c(b))$.

2.7. PROPOSITION. L'application $\pi : c \mapsto (c(a), c(b))$ est une submersion de $\Omega(I, M)$ sur $M \times M$, de classe C^{k-1} . Si $k \geq 2$ sa dérivée est donnée par $d\pi_c(V) = (V(a), V(b))$.

Pour $x \in M$ et $y \in M$, l'espace $\Omega_{xy}(I, M)$ est une sous-variété de classe C^{k-1} et de codimension $2m$ dans $\Omega(I, M)$. Si $k \geq 2$ l'espace tangent à cette sous-variété en c est défini dans $T_c \Omega(I, M)$ par les équations $V(a) = 0$ et $V(b) = 0$.

Démonstration. Avec les notations de 2.2 l'application π se lit dans les cartes $(\Omega_{\sigma, \theta}, \Phi_{\sigma, \theta})$ et $(U_1 \times U_n, \phi_1 \times \phi_n)$ comme l'application linéaire surjective $(c_1, \dots, c_n) \mapsto (c_1(a), c_n(b))$. La proposition en découle.

3. Les géodésiques sont les points critiques de l'énergie (formule de la variation première).

3.1. On désigne par M une variété riemannienne de classe C^r avec $r \geq 1$ (la variété sous-jacente est donc de classe C^{r+1} , la métrique riemannienne de classe C^r) et on fixe un intervalle $I = [a, b]$.

Les notations employées pour le produit scalaire et la connexion riemannienne sont celles du ch.3.

3.2. PROPOSITION. (i) La fonction énergie $E(c) = \frac{1}{2} \int_a^b \|\dot{c}(t)\|^2 dt$ est une fonction différentiable de classe C^r sur $\Omega(I, M)$.

(ii) Pour $c \in \Omega(I, M)$ et $V \in T_c \Omega(I, M)$ on a

$$(3) \quad dE_c(V) = \int_a^b \langle \dot{c}, \nabla_{\dot{c}} V \rangle dt ;$$

si le chemin c est de classe C^2 on a encore

$$(3') \quad dE_c(V) = \langle \dot{c}(b), V(b) \rangle - \langle \dot{c}(a), V(a) \rangle - \int_a^b \langle \nabla_{\dot{c}} \dot{c}, V \rangle dt.$$

(formule de la variation première).

Démonstration. (i) Considérons d'abord le cas où M est un ouvert U de $\mathbb{R}^m$, muni d'une métrique riemannienne g de classe C^r . On a

$$E(c) = \frac{1}{2} \int_a^b \sum_{k,l=1}^m g_{kl}(c(t)) \dot{c}_k(t) \dot{c}_l(t) dt$$

pour tout chemin $c \in \Omega(I, U)$. L'application

$$(\lambda, \mu, \nu) \mapsto \frac{1}{2} \int_a^b \sum_{k,l=1}^m \lambda_{kl}(t) \mu(t) \nu(t) dt$$

est une forme trilinéaire continue sur $C(I, \mathbb{R}^{m^2}) \times L^2(I, \mathbb{R}^m) \times L^2(I, \mathbb{R}^m)$. On est donc ramené à montrer que les applications $c \mapsto g \circ c$ et $c \mapsto \dot{c}$ à valeurs respectivement dans $C(I, \mathbb{R}^{m^2})$ et $L^2(I, \mathbb{R}^m)$ sont de classe C^r . Pour l'application $c \mapsto \dot{c}$ il n'y a pas de problème puisque c'est la restriction à un ouvert d'une application linéaire continue. L'application $c \mapsto g \circ c$ est, quant à elle, la composée de l'injection $\phi(I, U) \rightarrow C(I, U)$ (linéaire continue) et de l'application $g \mapsto g \circ c$ définie sur $C(I, U)$. Tout se ramène donc à montrer que si $\phi_* : U \rightarrow \mathbb{R}^p$ est une application de classe C^r

l'application $\phi_* : C(I, U) \rightarrow C(I, \mathbb{R}^P)$ définie par $\phi_* c = \phi \circ c$ est aussi de classe C^r , ce qu'on fait facilement en raisonnant par récurrence sur r .

Le cas général se ramène au précédent en utilisant les cartes définies en 2.2.

(ii) Soit $V \in T_V \Omega$. Considérons une variation $v(s, t)$ de c admettant V comme champ de vecteurs dérivé : $V(t) = \frac{\partial}{\partial s} v(s, t) \big|_{s=0}$. Définissons des champs de vecteurs X, W le long de v en posant

$$X = dv\left(\frac{\partial}{\partial t}\right), \quad W = dv\left(\frac{\partial}{\partial s}\right).$$

On a (le point désignant toujours la dérivée par rapport à la variable t)

$$\begin{aligned} \frac{d}{ds} E(v_s) &= \frac{1}{2} \int_a^b \frac{d}{ds} \langle \dot{v}_s, \dot{v}_s \rangle dt \\ &= \frac{1}{2} \int_a^b W \langle X, X \rangle dt \\ &= \int_a^b \langle X, \nabla_W X \rangle dt. \end{aligned}$$

(On a utilisé la formule donnant la dérivée d'un produit scalaire. Rappelons (cf. ch. 3, 2.2) qu'on peut calculer avec des champs le long de v de la même façon qu'avec des champs de vecteurs sur M). La symétrie de la connexion riemannienne (ibid., 2.4) donne

$$\nabla_W X - \nabla_X W = [X, W] = \left[dv\left(\frac{\partial}{\partial t}\right), dv\left(\frac{\partial}{\partial s}\right) \right] = dv\left(\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial s}\right]\right) = 0.$$

On obtient donc

$$(4) \quad \frac{d}{ds} E(v_s) = \int_a^b \langle X, \nabla_X W \rangle dt,$$

d'où on tire (2) en faisant $s = 0$ puisque $X(0, t) = \dot{c}(t)$ et $W(0, t) = V(t)$.

Si le chemin c est de classe C^2 on passe de (2) à (3) en utilisant l'identité

$$\frac{d}{dt} \langle \dot{c}, V \rangle = \langle \nabla_{\dot{c}} \dot{c}, V \rangle + \langle \dot{c}, \nabla_{\dot{c}} V \rangle$$

(ce qui revient à intégrer par parties).

3.3. THÉOREME. Soit M une variété riemannienne de classe C^1 et soient x et y deux points de M . Soit $I = [a, b]$ un intervalle compact de $\mathbb{R}$.

Les géodésiques $\gamma : I \rightarrow M$ d'origine x et d'extrémité y sont exactement les points critiques de la fonction énergie E sur la variété $\Omega_{xy}(I, M)$.

Démonstration. Soit $\gamma \in \Omega_{xy} = \Omega_{xy}(I, M)$ une géodésique. Par définition γ est un chemin de classe C^2 et vérifie l'équation différentielle $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ (ch.3, 3). On a donc d'après (3)

$$dE_{\gamma}(V) = \langle \dot{\gamma}(b), V(b) \rangle - \langle \dot{\gamma}(a), V(a) \rangle$$

pour tout champ $V \in T_{\gamma} \Omega$. Si $V \in T_{\gamma} \Omega_{xy}$ on a $V(a) = 0$ et $V(b) = 0$ donc $dE_{\gamma}(V) = 0$ ce qui montre que γ est un point critique pour la restriction de E à Ω_{xy} .

Réciproquement soit γ un point critique de E sur Ω_{xy} . Considérons une base orthonormée $V_1, \dots, V_m$ de $T_x M$. Les champs de vecteurs $X_1, \dots, X_m$ le long de γ obtenus en transportant parallèlement $v_1, \dots, v_m$ sont de classe H^1

$$\dot{X}^k = - \sum_{i,j=1}^m \Gamma_{ij}^k(\gamma(t)) X^i \dot{\gamma}_j(t)$$

et forment pour tout $t \in I$ une base orthonormée de $T_{\gamma(t)} M$.

Pour toute suite $f_1, \dots, f_m$ de fonctions C^∞ sur I , à support compact dans $[a, b]$, on a, en posant $V(t) = \sum_{i=1}^m f_i(t) X_i(t)$,

$$\int_a^b \langle \dot{\gamma}, \sum_j V_j \rangle = dE_{\gamma}(V) = 0$$

(puisque $V \in T_c \Omega_{xy}$) soit

$$\sum_{i=1}^m \int_a^b \langle \dot{\gamma}, X_i \rangle \frac{df_i}{dt} = \sum_{i=1}^m \int_a^b \frac{d}{dt} \langle \dot{\gamma}, X_i \rangle f_i = 0$$

où la dérivée $\frac{d}{dt} \langle \dot{\gamma}, X_i \rangle$ de la fonction $\langle \dot{\gamma}, X_i \rangle$ est prise au sens des distributions (sur $]a, b[$). On a donc $\frac{d}{dt} \langle \dot{\gamma}, X_i \rangle = 0$ ce qui entraîne que la fonction $\langle \dot{\gamma}, X_i \rangle$ est égale presque partout à une constante, soit λ_i . On a par conséquent

$$\dot{\gamma} = \sum \lambda_i X_i \quad \text{p.p.}$$

On a ainsi montré que le champ de vecteurs $\dot{\gamma}$ est de classe H^1 et parallèle le long de γ

$$\nabla_{\dot{\gamma}} \dot{\gamma} = \sum \lambda_i \nabla_{\dot{\gamma}} X_i = 0 \quad \text{p.p.}$$

Pour terminer, montrons que le chemin γ est de classe C^2 , ce qui prouvera du même coup que $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ partout et donc que γ est une géodésique : puisque $\dot{\gamma}$ est de classe H^1 (donc continu) le chemin γ est de classe C^1 . Comme le champ $\dot{\gamma}$ est parallèle le long de γ il est lui aussi de classe C^1 - considérer encore une fois l'équation différentielle $\dot{X}^k = - \sum_{i,j=1}^m \Gamma_{ij}^k (\gamma(t)) X^i \dot{\gamma}_j(t)$ - et γ est donc de classe C^2 .

4. Le Hessien de l'énergie en un point critique (ou la formule de la variation seconde)

4.1. La notion de dérivée seconde en un point x d'une fonction deux fois différentiable définie sur un espace normé ne possède pas d'analogue naturel pour une fonction définie sur une variété, sauf en un point critique de cette fonction.

LEMME. Soient X une variété de classe C^2 , $f : X \rightarrow \mathbb{R}$ une fonction de classe C^2 et x un point critique de f . Soit ϕ une carte de X définie au voisinage de x .

Alors la forme bilinéaire sur $T_x X$ définie par

$$\text{Hess}_x f(u, v) = (f \circ \phi^{-1})''_{\phi(x)} (d\phi_x(u), d\phi_x(v))$$

ne dépend pas de la carte ϕ choisie.

Démonstration. Soit ψ une autre carte définie au voisinage de x et soit $\Phi = \phi \circ \psi^{-1}$.

Introduisons les fonctions $g = f \circ \phi^{-1}$ et $h = f \circ \psi^{-1}$. La relation $h = g \circ \Phi$ entraîne

$$h' = (g' \circ \Phi) \cdot \Phi'$$

et

$$h'' = (g'' \circ \Phi) \cdot \Phi'^2 + (g' \circ \Phi) \Phi''$$

d'où on tire

$$\begin{aligned} (f \circ \psi^{-1})''_{\psi(x)} (d\psi_x(u), d\psi_x(v)) &= (f \circ \phi^{-1})''_{\phi(x)} (d\phi_x(u), d\phi_x(v)) \\ &\quad + (f \circ \phi^{-1})'_{\phi(x)} (\Phi''_{\psi(x)} (d\psi_x(u), d\psi_x(v))). \end{aligned}$$

Si x est un point critique de f c'est-à-dire si $df_x = 0$, on a

$$(f \circ \phi^{-1})'_{\phi(x)} = df_x \circ d(\phi^{-1})_{\phi(x)} = 0, \text{ d'où le lemme.}$$

4.2. DÉFINITION. On dit que $\text{Hess}_x f$ est le Hessien de f au point critique x . Si X est un espace de Banach le Hessien coïncide avec la dérivée seconde ordinaire.

4.3. DÉFINITION. Supposons maintenant que X soit une variété hilbertienne. On dit que x est un point critique non dégénéré de f si la forme bilinéaire symétrique $\text{Hess}_x f$ est non dégénérée.

Précisons ce que nous entendons par "non dégénérée". Soit H un espace de Hilbert et soit b une forme bilinéaire symétrique continue sur H . À b on fait correspondre une application linéaire $\beta : H \rightarrow H$ caractérisée par l'identité $\langle \beta(u), v \rangle = b(u, v)$. On dit que b est non dégénérée si β est un isomorphisme. En dimension finie il suffit pour que b soit non dégénérée que son noyau (qui est par définition le noyau de β) soit réduit à 0.

4.4. Nous utiliserons pour énoncer la formule de la variation seconde un résultat un peu plus précis que le lemme 4.2.

LEMME. Soient X et Y deux variétés (hilbertiennes) de classe C^r ($r \geq 2$), $\pi : X \rightarrow Y$ une submersion et $f : X \rightarrow \mathbb{R}$ une fonction de classe C^r . Soit $y \in Y$ et soit x un point critique de la restriction de f à la sous-variété $\pi^{-1}(y)$. Soient enfin ϕ et ψ des cartes de X et Y définies au voisinage de x et de y respectivement, telles que $\psi \circ \pi \circ \phi^{-1}$ soit une projection $U \times V \rightarrow V$

Alors, la forme bilinéaire

$$\text{Hess}_x^{\pi} f : T_x X \times T_x(\pi^{-1}(y)) \rightarrow \mathbb{R}$$

définie par

$$\text{Hess}_x^{\pi} f(u, v) = (f \circ \phi^{-1})''_{\phi(x)}(d\phi_x(u), d\phi_x(v))$$

est indépendante des cartes choisies.

Démonstration. Analogue à celle de 4.2.

4.5. La formule de la variation seconde. Considérons une variété riemannienne M de classe C^r ($r \geq 2$) et fixons un intervalle $I = [a, b]$. On a vu que $\Omega = \Omega(I, M)$ est une variété de classe C^r et que l'application $\pi : \Omega \rightarrow M \times M$ qui associe à un chemin c le couple $(c(a), c(b))$ est une submersion.

Soient x et y deux points de M et soit $\gamma : I \rightarrow M$ une géodésique joignant x à y .

Puisque γ est sur Ω_{xy} un point critique de la fonction énergie E le lemme 4.4 donne un sens à la forme bilinéaire

$$\text{Hess}_{\gamma}^{\pi} E : T_{\gamma} \Omega \times T_{\gamma} \Omega_{xy} \rightarrow \mathbb{R}.$$

Pour calculer $\text{Hess}_{\gamma}^{\pi}$ nous supposons provisoirement $r \geq 3$ et nous utiliserons une carte de Ω définie à partir de l'application exponentielle. Puisque le domaine de définition D de l'application exponentielle est ouvert dans TM , l'ensemble D_{γ} des champs $V \in T_{\gamma} \Omega$ tels que pour tout $t \in I$ le vecteur $V(t) \in T_{\gamma(t)} M$ appartienne à D est dans $T_{\gamma} \Omega$ un voisinage ouvert de l'origine. Les propriétés de l'application $\exp$ (ch. 3 (Ph. Bougerol)) assurent que l'application

$$\text{EXP}_{\gamma} : D_{\gamma} \rightarrow \Omega$$

définie par

$$(\text{EXP}_{\gamma} V)(t) = \exp_{\gamma(t)} V(t)$$

est de classe C^{r-1} et que c'est un difféomorphisme local au voisinage de D . D'autre part l'application

$$\exp_x \times \exp_y : D_x \times D_y \rightarrow M \times M$$

définie par

$$(\exp_x \times \exp_y)(v, w) = (\exp_x(v), \exp_y(w))$$

est un difféomorphisme local au voisinage de 0 . Les applications inverses locales de EXP_{γ} et $\exp_x \times \exp_y$ fournissent donc un système de cartes au voisinage de γ et de (x, y) respectivement dans lequel l'application $\pi : \Omega \rightarrow M \times M$ est lue comme l'application linéaire surjective $d\pi_{\gamma} : T_{\gamma} \Omega \rightarrow T_x M \times T_y M$.

Puisqu'à l'origine $d\text{EXP}$ est l'identité on obtiendra en vertu de (10) la forme bilinéaire $\text{Hess}_{\gamma}^{\pi} E$ en restreignant à $T_{\gamma} \Omega \times T_{\gamma} \Omega_{xy}$ la forme bilinéaire $(E \circ \text{EXP})''$.

Posons, pour $t \in I$ et s assez petit, $v(s, t) = \exp_{\gamma(t)}(s V(t))$. Il nous suffit de calculer $(E \circ \text{EXP})''(V, V) = \frac{d^2}{ds^2} E(v_s) \Big|_{s=0}$. Posons $X = dv(\frac{\partial}{\partial t})$ et $W = dv(\frac{\partial}{\partial s})$ (cf. démonstration de 3.1). On a

$$\frac{d}{ds} E(v_s) = \int_a^b \langle X, \nabla_X W \rangle dt,$$

$$\begin{aligned} \frac{d^2}{ds^2} E(v_s) &= \int_a^b \frac{\partial}{\partial s} \langle X, \nabla_X W \rangle dt = \int_a^b X \langle X, \nabla_X W \rangle dt \\ &= \int_a^b [\langle \nabla_W X, \nabla_X W \rangle + \langle X, \nabla_W \nabla_X W \rangle] dt. \end{aligned}$$

Observons maintenant que par définition du tenseur de courbure R (ch. 6 (J. Granara)) on a

$$(\nabla_W \nabla_X - \nabla_X \nabla_W)W = R(W, X)W = -R(X, W)W$$

et que

$$\nabla_W W = 0$$

puisque pour t fixé la courbe $s \rightarrow v(s, t)$ est une géodésique. On obtient

$$\frac{d^2}{ds^2} E(v_s) = \int_a^b [\langle \nabla_X W, \nabla_X W \rangle - \langle R(X, W)W, X \rangle] dt$$

soit, en faisant $s = 0$

$$(11) \quad \frac{d^2}{ds^2} E(v_s) \Big|_{s=0} = \int_a^b [\langle \nabla_{\dot{\gamma}} V, \nabla_{\dot{\gamma}} V \rangle - \langle \dot{\gamma}, R(\dot{\gamma}, V)V \rangle] dt$$

(formule de la variation seconde).

On en tire par polarisation $(E \circ \text{EXP})''(V, W)$, d'où

$$\text{Hess } E(V, W) = \int_a^b [\langle \nabla_{\dot{\gamma}} V, \nabla_{\dot{\gamma}} W \rangle - \langle \dot{\gamma}, R(\dot{\gamma}, V)W \rangle] dt.$$

Cette formule qui ne fait intervenir que les dérivées premières et secondes de la métrique riemannienne g (en coordonnées locales, des fonctions g_{ij}) est encore vraie si on suppose seulement g de classe C^2 (et la variété sous-jacente de classe C^3), auquel cas Ω est une variété de classe C^2 et E une fonction de classe C^2 sur Ω . Mais l'application exponentielle est alors de classe C^1 et EXP_{γ} est seulement continue. La démonstration ci-dessus n'est donc plus valable.

Un examen attentif montre que pour prouver (11) nous n'avons utilisé essentiellement comme propriété de l'exponentielle que la suivante (on pose $F_t = \exp_{\gamma}(t)$)

(i) pour tout $t \in I$ et pour tout $v \in T_{\gamma(t)}$ la courbe $s \rightarrow F_t(s, v)$ satisfait au point $s = 0$ l'équation différentielle des géodésiques.

Il nous suffira donc pour prouver que (11) est encore vraie pour une métrique C^2 de construire une application $F : D \rightarrow M$ de classe C^2 , l'ensemble D étant un voisinage dans TM de l'ensemble des points $(\gamma(t), 0)$ avec $t \in I$, telle que la restriction F_t de F à $D \cap T_{\gamma(t)}M$ vérifie (i). Supposons pour simplifier que M soit un ouvert de $\mathbb{R}^m$ (on peut toujours si γ est sans point double se ramener à ce cas au moyen d'une carte). On peut alors définir F en prenant pour $F(t, v)$ la valeur en $s = 1$ de la solution de

$$\frac{d^2 X_k}{ds^2} + \sum_{i,j=1}^m \Gamma_{ij}^k \frac{dX_i}{ds} \frac{dX_j}{ds} = 0$$

qui satisfait aux conditions initiales

$$X_k(0) = 0, \quad \frac{dX_k}{ds}(0) = v_k.$$

4.6. THÉOREME. Soit M une variété riemannienne de classe C^2 . Soit $I = [a, b]$ et soit $\gamma : I \rightarrow M$ une géodésique d'extrémités x et y . On a, pour $V \in T_{\gamma} \Omega(I, M)$ et $W \in T_{\gamma} \Omega_{xy}(I, M)$

$$(12) \text{ Hess}_{\gamma}^{\pi} E(V, W) = \int_a^b [\langle \nabla_{\dot{\gamma}} V, \nabla_{\dot{\gamma}} W \rangle - \langle \dot{\gamma}, R(\dot{\gamma}, V)W \rangle] dt.$$

Le Hessien de la restriction de E à la sous-variété $\Omega_{xy}(I, M) = \Omega_{xy}$ en son point critique γ , s'obtient en restreignant à $T_{\gamma} \Omega_{xy} \times T_{\gamma} \Omega_{xy}$ la forme bilinéaire $\text{Hess}_{\gamma}^{\pi} E$. Pour alléger l'écriture on le notera I_{γ} (au lieu de $\text{Hess}_{\gamma}(E|_{\Omega_{xy}})$).

4.7. Soit $b : H_1 \times H_2 \rightarrow \mathbb{R}$ une forme bilinéaire continue définie sur un produit d'espaces de Hilbert réels. On définit deux applications linéaires continues $\beta_g : H_1 \rightarrow H_2$ et $\beta_d : H_2 \rightarrow H_1$, appelées respectivement applications linéaires gauche et droite associées à b , en posant

$$b(u, v) = \langle \beta_g(u), v \rangle = \langle u, \beta_d(v) \rangle.$$

On dira que b est non dégénérée à gauche si β_d est surjective. Cette condition implique la condition " β_g injective" puisque β_d et β_g sont adjoints mais elle ne lui

est équivalente qu'en dimension finie. On définit de même la notion de forme non dégénérée à droite. Le noyau de β_g (resp. β_d) est appelé noyau à gauche (resp. noyau à droite) de b .

4.8. LEMME. (Notations et hypothèses du lemme 4.4). Soit C_f l'ensemble des points x de X qui sont sur la sous-variété $\pi^{-1}(\pi(x))$ points critiques de la fonction f . Supposons que pour tout $x \in C_f$ la forme bilinéaire $\text{Hess}_x^\pi f : T_x V \times \ker d\pi_x \rightarrow \mathbb{R}$ soit non dégénérée à droite.

Alors C_f est une sous-variété de classe C^{r-1} de X et l'espace tangent en un point $x \in C_f$ est le noyau à gauche de $\text{Hess}_x^\pi f$.

Démonstration. Il suffit de traiter le cas où $X = U \times V$, $Y = V$ les ensembles U et V étant des ouverts dans des espaces de Hilbert H et K et où $\pi : U \times V \rightarrow V$ est la projection sur le second facteur.

L'ensemble C_f est alors défini par l'équation $\frac{\partial f}{\partial v}(u,v) = 0$. Or l'application $x = (u,v) \rightarrow \frac{\partial f}{\partial v}(u,v)$ est au voisinage de C_f une submersion de X dans K puisque sa dérivée au point $x \in C_f$ est l'application linéaire gauche associée à $\frac{\partial^2 f}{\partial x \partial v}$, surjective par hypothèse. Il en résulte que C_f est une sous-variété, l'espace tangent en un point $x \in C_f$ étant le noyau de cette application linéaire.

4.9. PROPOSITION. La forme bilinéaire $\text{Hess}_Y^\pi E$ fournie par la relation (12) est non dégénérée à droite et son noyau à gauche est l'espace des champs de vecteurs J le long de γ , de classe C^r , solutions de l'équation différentielle, dite de Jacobi

$$(13) \quad \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J + R(\dot{\gamma}, J)\dot{\gamma} = 0$$

Démonstration. Faisons sur $T_Y \Omega$ le choix du produit scalaire

$$\langle V, W \rangle_{H^1} = \langle V(a), W(a) \rangle + \int_a^b \langle \nabla_{\dot{\gamma}} V, \nabla_{\dot{\gamma}} W \rangle dt.$$

L'application linéaire gauche β_g associée à la forme $\text{Hess}_Y^\pi E = b$ fait correspondre à $V \in T_Y \Omega$ l'unique $\tilde{V} \in T_Y \Omega_{xy}$ tel que l'on ait

$$(14) \quad \int_a^b [\langle \nabla_{\dot{\gamma}} V, \nabla_{\dot{\gamma}} W \rangle - \langle \dot{\gamma}, R(\dot{\gamma}, V)W \rangle] dt = \int_a^b \langle \nabla_{\dot{\gamma}} \tilde{V}, \nabla_{\dot{\gamma}} W \rangle dt$$

pour tout $W \in T_{\gamma}^{\Omega} \text{xy}$. Considérons des champs de vecteurs X_i parallèles le long de γ formant en chaque point $\gamma(t)$ une base orthonormée de $T_{\gamma(t)}^M$. Si on pose

$V(t) = \sum_i f_i(t) X_i(t)$, $V(t) = \sum_i f_i(t) X_i(t)$ la relation (14) devient

$$(15) \quad \sum_i \int_a^b \dot{f}_i \langle X_i, \nabla_{\dot{\gamma}} W \rangle - \sum_i \int_a^b f_i \langle \dot{\gamma}, R(\dot{\gamma}, X_i)W \rangle = \sum_i \int_a^b f_i \langle X_i, \nabla_{\dot{\gamma}} W \rangle.$$

Posons $a_{ij}(t) = \langle \dot{\gamma}, R(\dot{\gamma}, X_i)X_j \rangle$. Si on prend W de la forme $W(t) = \sum_j g_j(t) X_j(t)$, avec g_j de classe C^∞ à support compact dans $[a, b]$, la relation (15) s'écrit encore après intégration par parties

$$(16) \quad \sum_i \int_a^b \ddot{f}_i g_i + \sum_{i,j} \int_a^b f_i a_{ij} g_j = \sum_i \int_a^b \ddot{f}_i g_i$$

Comme les champs W de la forme précédente sont denses dans $T_{\gamma}^{\Omega} \text{xy}$, pour que l'on ait $V = \Lambda(V)$ il faut et il suffit que l'on ait la relation (16) pour tout choix des fonctions g_i , c'est-à-dire que l'on ait

$$(17) \quad \ddot{f}_i + \sum_j f_j a_{ij} = \ddot{f}_i \quad i=1, \dots, n$$

au sens des distributions. Il est clair, maintenant, que l'équation différentielle (17) admet une solution (f_i) de classe H^1 quelles que soient les fonctions $\tilde{f}_i$ de classe H^1 . Ceci montre que β_g est surjective et que Hess_γ^π est une forme non dégénérée à droite. Pour que $V = \sum f_i X_i$ appartienne au noyau de β_g il faut et il suffit que l'on ait

$$(18) \quad \ddot{f}_i + \sum_j f_j a_{ij} = 0$$

Comme les fonctions a_{ij} sont de classe C^{r-2} on voit que les fonctions f_i sont de classe C^r . L'équation (18) peut donc s'écrire

$$\nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} V + R(\dot{\gamma}, V)\dot{\gamma} = 0.$$

Les solutions de l'équation (13) sont appelées champs de Jacobi le long de (cf. ch.6 (J. Granara)). Elles forment un espace vectoriel de dimension $2m$.

En rassemblant les résultats de 4.8 et 4.9 on obtient

4.10. THÉOREME. Soit M une variété riemannienne de classe C^r ($r \geq 2$). L'ensemble $G(I, M)$ des géodésiques de M paramétrées par $I = [a, b]$ est une sous-variété de classe C^{r-1} et de dimension $2m$ de $\Omega(I, M)$. L'espace tangent en un point γ à $G(I, M)$ est l'espace des champs de Jacobi le long de γ .

4.11. Remarque. On vérifie aisément que l'application $\gamma \rightarrow (\gamma(a), (b-a)\dot{\gamma}(a))$ réalise un difféomorphisme de classe C^{r-1} de $G(I, M)$ sur le domaine de définition de l'application exponentielle.

5. Points conjugués

On suppose que M est une variété riemannienne de classe C^r ($r \geq 2$). On pose $I = [0, 1]$. Tous les chemins considérés sont paramétrés par I , sauf indication contraire.

5.1. DÉFINITION. Soient x et y deux points de M joints par une géodésique γ . On dit que x et y sont conjugués le long de γ si γ est un point critique dégénéré de la fonction énergie E définie sur $\Omega_{xy}(I, M)$ par

$$E(c) = \frac{1}{2} \int_0^1 \|\dot{c}(t)\|_{c(t)}^2 dt.$$

5.2. PROPOSITION. Soient x et y deux points de M joints par une géodésique γ et soit $v = \dot{\gamma}(0)$ le vecteur-vitesse en x de γ . Les deux espaces vectoriels suivants coïncident:

(i) le noyau de la forme I_γ , Hessien au point critique γ de la fonction $E : \Omega_{xy} \rightarrow \mathbb{R}$;

(ii) l'espace des champs de Jacobi le long de γ nuls en x et en y .
Ils sont, de façon naturelle, isomorphes à

(iii) le noyau de l'application dérivée de l'exponentielle $d(\exp_x)_v : T_x^M \rightarrow T_y^M$.

La dimension commune de ces trois espaces est $\leq m-1$.

Les points x et y sont conjugués le long de γ quand cette dimension est ≥ 1 . On dit que c'est la multiplicité de la conjugaison.

Démonstration. Rappelons (th.4.6) que I est la restriction à $T_\gamma \Omega_{xy} \times T_\gamma \Omega_{xy}$ de la forme bilinéaire

$$\text{Hess}_\gamma^\pi(U, V) = \int_0^1 [\langle \nabla_{\dot{\gamma}} U, \nabla_{\dot{\gamma}} V \rangle - \langle \dot{\gamma}, R(\dot{\gamma}, U)V \rangle] dt$$

définie sur $T_\gamma \Omega \times T_{\gamma xy} \Omega$ et (prop. 4.9) que le noyau à gauche de cette forme bilinéaire est l'espace vectoriel des champs de Jacobi le long de γ . Le noyau de I_γ s'obtient évidemment en prenant l'intersection du noyau à gauche de Hess E et de $T_{\gamma xy} \Omega$. Il est donc constitué des champs de Jacobi J le long de γ tels que $J(0) = 0$ et $J(1) = 0$, ce qui prouve l'égalité de (i) et de (ii).

$$\begin{array}{ccccc} & & G & & \\ & \nearrow & & \searrow \pi^G & \\ TM \supset D & \xrightarrow{\exp} & M & & M \end{array}$$

Soit D le domaine de définition de l'application exponentielle. L'application $v \rightarrow \gamma_v$ qui à $v \in T_x M$ associe la géodésique $\gamma_v : t \rightarrow \exp_x(tv)$ ($0 \leq t \leq 1$) est un difféomorphisme de D sur G. L'application exponentielle $\exp : D \rightarrow M \times M$ s'obtient en composant cette application $v \rightarrow \gamma_v$ établit donc un isomorphisme entre le noyau de $d \exp_v : T_v(TM) \rightarrow T_x M \times T_y M$ et celui de $d\pi^G : T_\gamma G \rightarrow T_x M \times T_y M$ lequel se trouve être, puisque $d\pi(J) = (J(0), J(1))$, l'espace vectoriel des champs de Jacobi nuls en x et y. L'isomorphisme de (iii) sur (ii) s'obtient en composant l'isomorphisme ci-dessus avec celui que la différentielle de l'inclusion $T_x M \rightarrow TM$ induit entre le noyau de $d(\exp_x)_v$ et le noyau de $d \exp_v$.

Cet isomorphisme de (iii) sur (ii) peut se décrire de façon plus explicite : soit $v \in T_x M$. En associant à $w \in T_x M$ le champ de Jacobi $J_w(t) = \frac{\partial}{\partial s} \exp_x(t(v+sw))$ on définit une bijection linéaire entre $T_x M$ et l'espace des champs de Jacobi le long de γ_v nuls pour $t = 0$. Il est évident que cette bijection met en correspondance le noyau de $d(\exp_x)_v$ et les champs de Jacobi le long de γ_v nuls pour $t = 0$ et $t = 1$.

La dernière assertion de l'énoncé résulte de ce que $\exp_x$ est partout de rang ≥ 1 , puisque $d(\exp_x)_v \cdot v = \frac{d}{ds}(\exp_x(v+sv))|_{s=0} = \dot{\gamma}_v(1) \neq 0$.

5.3. COROLLAIRE. Soient x et y deux points de M joints par une géodésique γ et soit $v = \dot{\gamma}(0)$. Les propriétés suivantes sont équivalentes :

- (i) x et y sont conjugués le long de γ ;
- (ii) (x,v) est point critique de l'application exponentielle ;
- (iii) v est point critique de $\exp_x$;

(iv) il existe un champ de Jacobi non nul le long de γ tel que $J(0) = 0$ et $J(1) = 0$.

5.4. La condition "x et y ne sont pas conjugués le long de γ " exprime la stabilité de la géodésique γ - par stabilité, on entend ici la stabilité de γ , en tant que solution d'un problème de Calcul des variations, en fonction des variations des données x et y : on dit que γ est stable si pour x' et y' assez voisins de x et y il existe une géodésique $\gamma(x', y')$ joignant x' à y' , et si $\gamma(x', y')$ peut être choisie de manière que l'application $(x', y') \mapsto \gamma(x', y')$ soit différentiable. En effet, pour que x et y soient non conjugués, il faut et il suffit que $T_{\gamma} \Omega_{xy} \cap T_{\gamma} G = 0$. Comme $T_{\gamma} \Omega_{xy}$ est de codimension 2m et $T_{\gamma} G$ de dimension 2m cela revient à dire que Ω_{xy} et G coupent transversalement, ou que

$$T_{\gamma} \Omega_{xy} + T_{\gamma} G = T_{\gamma} \Omega.$$

Comme $d\pi_{\gamma} : T_{\gamma} \Omega \rightarrow T_x^M \times T_y^M$ est surjective et que $\ker d\pi_{\gamma} = T_{\gamma} \Omega_{xy}$ il revient encore au même de dire que $d\pi_{\gamma}^G : T_{\gamma} G \rightarrow T_x^M \times T_y^M$ est bijective ce qui, d'après le théorème des fonctions implicites exprime exactement la stabilité de γ telle qu'elle a été définie ci-dessus.

5.5. PROPOSITION. *L'ensemble des points conjugués d'un point donné x est de mesure nulle dans M.*

Démonstration. Les points conjugués de x sont les valeurs critiques de $\exp_x$. Il suffit d'appliquer le théorème de Sard.

5.6. Exemples de points conjugués (sans démonstration)

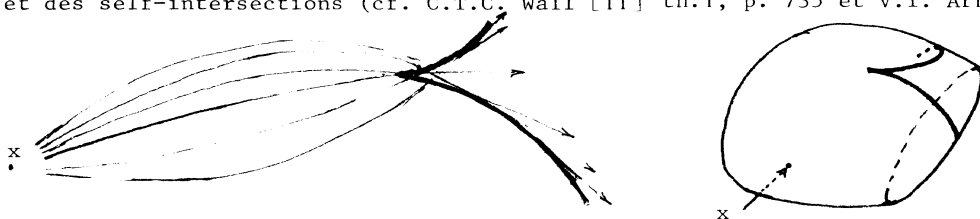
1) Soient x et y deux points de M joints par une géodésique γ . Supposons que γ appartienne à une variété de dimension k de géodésiques joignant x et y. Alors x et y sont conjugués, avec une multiplicité $\geq k$, le long de γ . Par exemple, deux points opposés de la sphère S^m sont conjugués avec multiplicité m-1 le long de tout demi-grand cercle les joignant. Les pôles d'une surface de révolution sont conjugués avec multiplicité 1 le long d'une méridienne quelconque.

2) La situation précédente où l'ensemble des points conjugués (lieu conjugué) d'un certain point x est réduit à un point (ou deux) sur lequel se localisent toutes les géodésiques issues de x est loin d'être le cas général.

Le lieu conjugué de x est en général de dimension $m-1$ (mais n'est pas une sous-variété même immergée, car il peut présenter des singularités) et en général, un point conjugué de x n'est conjugué de x que suivant une seule géodésique.

En dimension 2, le lieu conjugué de x est l'enveloppe des géodésiques issues de x et on démontre sur la structure de ce lieu conjugué le résultat "générique" suivant:

Soit M une surface C^∞ . Il existe dans l'espace des métriques riemanniennes g de classe C^∞ sur M un ensemble exceptionnel ε tel que si $g \notin \varepsilon$ la variété riemannienne (M, g) possède la propriété suivante : pour tout $x \in M$ le lieu conjugué de x est une courbe différentiable ne présentant comme singularités que des points de rebroussement et des self-intersections (cf. C.T.C. Wall [11] th.1, p. 735 et V.I. Arnold [1])



cor. 11.8, p.271, ou peut être Thom, mais où ?).

3) On montre à partir de l'équation différentielle des champs de Jacobi que si M est à courbure ≤ 0 (c'est-à-dire si pour tout $x \in M$ et pour tout u et tout v dans $T_x M$ on a $\langle R(u, v)u, v \rangle \leq 0$) il n'y a pas de points conjugués sur M (cf. Milnor [8]).

6. Existence de géodésiques minimisantes.

6.1. PROPOSITION. Soit M une variété riemannienne de classe C^0 et soit $I = [a, b]$.

(i) La fonction énergie E est semi-continue inférieurement sur l'espace $C(I, M)$.

(ii) Soit $\lambda < +\infty$ et soit K un compact de M . L'ensemble des chemins $c \in C(I, M)$ tels que $c(I) \subset K$ et $E(c) \leq \lambda$, noté $\Omega^\lambda(I, K)$, est compact dans $C(I, M)$.

Démonstration. Il s'agit de prouver que pour toute suite de chemins $c_1, c_2, \dots, c_p, \dots$ convergeant uniformément vers un chemin c on a

$$\liminf_{p \rightarrow \infty} E(c_p) \geq E(c).$$

On se ramène immédiatement à prouver que si, de plus, $E(c_p)$ converge vers une limite λ finie on a $\lambda \geq E(c)$.

Considérons une suite $(t_i)_{0 \leq i \leq n}$ de points de I et une suite $(U_i, \phi_i)_{1 \leq i \leq n}$ de cartes de M telles que

$$a = t_0 \leq t_1 \leq \dots \leq t_n = b \text{ et } c([t_{i-1}, t_i]) \subset U_i.$$

On a $c_p([t_{i-1}, t_i]) \subset U_i$ pour $i=1, \dots, n$ dès que p est assez grand. En tronçonnant les chemins c et c_p on est ainsi ramené au cas où M est un ouvert U de $\mathbb{R}^m$ muni d'une métrique riemannienne g . On notera $\langle u, v \rangle$ le produit scalaire ordinaire dans $\mathbb{R}^m$ et $\langle g(x)u, v \rangle$ le produit scalaire défini par g en x , en considérant $g(x)$ comme un endomorphisme symétrique.

Comme les chemins c_p sont tous situés dans un même compact de U on a une majoration

$$\int_a^b \langle \dot{c}_p(t), \dot{c}_p(t) \rangle dt \leq \text{const.} \int \langle g(c_p(t)) \dot{c}_p(t), \dot{c}_p(t) \rangle dt = \text{const. } E(c_p)$$

qui montre que l'intégrale $\int_a^b \langle \dot{c}_p(t), \dot{c}_p(t) \rangle dt$ reste bornée quand $p \rightarrow +\infty$. Quitte à remplacer la suite (c_p) par une sous-suite on peut donc supposer que la suite des dérivées $(\dot{c}_p)$ converge faiblement dans $L^2(I, \mathbb{R}^m)$ vers une fonction ϕ . On a alors

$$\begin{aligned} c(t) &= \lim_{p \rightarrow \infty} c_p(t) = \lim_{p \rightarrow \infty} [c_p(a) + \int_a^b 1_{[a, t]}(s) \dot{c}_p(s) ds] \\ &= c(a) + \int_0^t \phi(s) ds \end{aligned}$$

pour tout $t \in I$ ce qui montre déjà que c est absolument continu et que $\dot{c} = \phi$.

Remarquons maintenant que l'intégrale

$$J_p = \frac{1}{2} \int_a^b \langle g(c(t)) \dot{c}_p(t), \dot{c}_p(t) \rangle dt$$

a même limite λ que $E(c_p)$ quand $p \rightarrow \infty$. La quantité

$$\frac{1}{2} \int_a^b \langle g(c(t)) (\dot{c}_p(t) - \phi(t)), \dot{c}_p(t) - \phi(t) \rangle dt$$

est toujours ≥ 0 . En développant on voit qu'elle s'écrit

$$J_p - \int_a^b \langle g(c(t)) \phi(t), \dot{c}_p(t) \rangle dt + E(c)$$

et qu'elle a pour limite $\lambda - E(c)$ quand $p \rightarrow +\infty$.

(ii) L'ensemble $\Omega^\lambda(I, K)$ est d'après (i) une partie fermée de $C(I, M)$. Comme on a d'après l'inégalité de Schwarz

$$d(c(s), c(t)) \leq \int_s^t \|\dot{c}(u)\| du \leq \left(\int_s^t \|\dot{c}(u)\|^2 du \right)^{1/2} (t-s)^{1/2} \leq (2E(c))^{1/2} (t-s)^{1/2}$$

pour tout chemin c absolument continu, c'est aussi une partie équicontinue, et en vertu du théorème d'Ascoli une partie compacte de $C(I, M)$.

6.2. PROPOSITION. *Soit M une variété riemannienne de classe C^0 et soient x un point de M et r un nombre réel > 0 . Si la boule fermée $\bar{B}(x, r)$ est compacte, tout point y de cette boule peut être joint à x par une géodésique minimisante.*

Démonstration. Si $d(x, y) < r$ l'ensemble des chemins $c : [0, 1] \rightarrow M$ joignant x à y et d'énergie $\leq \frac{1}{2}r^2$ est compact et non vide (prendre pour c un chemin de longueur $< r$ paramétré uniformément). Sur cet ensemble la fonction E atteint son minimum. Si $d(x, y) = r$ le point y est limite d'une suite de points y_p , chaque point y_p pouvant être joint à x par une géodésique minimisante c_p . Puisque $E(c_p) = \frac{1}{2} d^2(x, y_p) \leq \frac{1}{2} r^2$ et que l'ensemble des chemins situés dans $\bar{B}(x, r)$ et d'énergie $\leq \frac{1}{2} r^2$ est compact la suite (c_p) possède un point adhérent qui est la géodésique minimisante cherchée.

6.3. THÉORÈME. *Soit M une variété riemannienne de classe C^0 . Les propriétés suivantes sont équivalentes :*

(i) *Munie de la distance riemannienne d , la variété M est un espace métrique complet. (On dit dans ce cas que M est complète).*

(ii) *Dans M , les boules fermées sont compactes.*

Si elles sont vérifiées, deux points quelconques de M peuvent être joints par une géodésique minimisante.

Démonstration. Soit $x \in M$ et soit R la borne supérieure des $r \geq 0$ tels que la boule fermée $\bar{B}(x, r)$ soit compact. Puisque M est localement compacte on a $R > 0$.

Supposons $R < +\infty$. Soit $(x_n)_{n \geq 0}$ une suite de points de $\bar{B}(x, R)$. Il existe pour tout n , et pour tout p entier $0 \leq p < n$ un chemin joignant x à x_n de longueur $\leq R + \frac{1}{p}$. Sur ce chemin on peut trouver un point x_{np} tel que

$$d(x, x_{np}) \leq R - \frac{1}{p}, \quad d(x_{np}, x_n) \leq \frac{2}{p}$$

Pour chaque p la suite $(x_{np})_{n \geq 0}$ est dans un compact et on peut en extraire une sous-suite convergente $(x_{i(n,p),p})_{n \geq 0}$ de façon à ce que la suite d'indices $(x_{i(p,p),p})_{n \geq 0}$ soit extraite de la suite $(i(n,p-1))_{n \geq 0}$. La suite diagonale $(x_{i(p,p),p})_{p \geq 1}$ est une suite de Cauchy, convergente en vertu de (i) ce qui montre que la suite (x_n) admet un point adhérent et que $\overline{B}(x, R)$ est compact.

Mais alors la boule $\overline{B}(x, R)$ possède dans M un voisinage compact K lequel contient nécessairement une boule $B(x, r)$ de rayon $r > R$; ce qui contredit la définition de R . On a donc $R = +\infty$ et (i) $\Leftrightarrow$ (ii).

L'assertion (ii) $\Leftrightarrow$ (i) est évidente et la seconde partie de l'énoncé résulte de 6.2.

6.4. Remarque. Si la variété riemannienne M est de classe C^1 on peut montrer que les assertions (i) et (ii) du th. 6.3 sont équivalentes à

(iii) toute géodésique peut être prolongée indéfiniment ; ou, ce qui revient au même, à

(iii') l'application exponentielle est définie sur TM tout entier (théorème de Hopf-Rinow, [5]).

7. Le cut-locus et la distance riemannienne.

Soit M une variété riemannienne complète de classe C^r ($r \geq 2$) et soit $I = [0, 1]$.

7.1. DÉFINITION. Soit x et y deux points de M . On dit que le couple (x, y) appartient au cut-locus de M si l'une des deux conditions suivantes est réalisée :

- (i) x et y sont joints par au moins deux géodésiques minimales
- (ii) x et y sont conjugués le long d'une géodésique minimale.

On note $C(M)$ le cut-locus de M . On note aussi $C(x)$, et on appelle cut-locus de x , l'ensemble des points $y \in M$ tels que $(x, y) \in C(M)$.

La condition $(x, y) \notin C(M)$ exprime alors le fait que sur Ω_{xy} la fonction énergie E

possède un minimum unique et non dégénéré.

7.2. L'hypothèse M complète assure l'existence d'un minimum γ sur Ω_{xy} pour l'énergie. La condition $(x,y) \notin C(M)$ traduit l'unicité et la stabilité de ce minimum :

7.3. PROPOSITION. Soient x et y deux points de M . Les propriétés suivantes sont équivalentes :

(i) (x,y) n'appartient pas au cut-locus de M

(ii) il existe une application $(x',y') \rightarrow \gamma(x',y')$ de classe C^1 , définie sur un voisinage $U \times V$ de (x,y) et à valeurs dans G , telle que pour $x' \in U$ et $y' \in V$ la géodésique $\gamma(x',y')$ soit une géodésique minimale joignant x' et y' .

7.4. COROLLAIRE. Le cut-locus $C(M)$ est une partie fermée de $M \times M$.

Démonstration. La condition (ii) définit un ouvert de $M \times M$.

7.5. LEMME. Soit $\gamma : W \rightarrow G$ une application de classe C^1 définie sur un ouvert W de $M \times M$, associant à tout couple (x,y) une géodésique $\gamma(x,y)$ joignant x et y . On pose pour $(x,y) \in W$

$$E^\gamma(x,y) = E(\gamma(x,y)) \text{ et } d^\gamma(x,y) = L(\gamma(x,y)).$$

Alors, la fonction E^γ (resp. d^γ) est différentiable de classe C^{r-1} sur W (resp. sur W privé des points de la forme (x,x)) et on a

$$(19) \quad d(E^\gamma) = \langle \dot{\gamma}(1), dy \rangle - \langle \dot{\gamma}(0), dx \rangle$$

$$(20) \quad d(d^\gamma) = \langle \frac{\dot{\gamma}(1)}{\|\dot{\gamma}(1)\|}, dy \rangle - \langle \frac{\dot{\gamma}(0)}{\|\dot{\gamma}(0)\|}, dx \rangle.$$

Démonstration. L'application γ est en fait de classe C^{r-1} puisqu'elle est par hypothèse un inverse local de la restriction à la sous-variété G , de classe C^{r-1} , de l'application $\pi : \gamma \rightarrow (\gamma(1), \gamma(0))$. La fonction E^γ est la composée de γ et de E .

La formule (17) est une conséquence de la formule de la variation première (3). La formule (18) se déduit de (17), en remarquant que

$E^\gamma = \frac{1}{2}(d^\gamma)^2$ et que pour $\gamma = \gamma(x,y)$ on a $\|\dot{\gamma}(0)\| = \|\dot{\gamma}(1)\| = d^\gamma(x,y)$.

7.6. Démonstration de la proposition 7.3.

(i) $\Leftrightarrow$ (ii). Supposons que $(x,y) \notin C(M)$. Soit γ_0 l'unique géodésique minimale joignant x et y . Puisque x et y ne sont pas conjugués suivant γ_0 , il existe (cf. 5.3) une application $\gamma : U \times V \rightarrow G$ de classe C^1 , où U (resp. V) est un voisinage de x (resp. y) telle que pour $x' \in U$ et $y' \in V$ la géodésique (x',y') admette x' et y' pour extrémités.

Montrons que (en restreignant si besoin est U et V) la géodésique $\gamma(x',y')$ est l'unique géodésique minimale d'extrémités x' et y' . Supposons qu'il existe une suite (x_n) (resp. (y_n)) convergeant vers x (resp. y) et pour chaque n une géodésique minimale c_n joignant x_n et y_n , distincte de $\gamma(x_n, y_n)$. La suite c_n admet un point adhérent c (d'après le théorème d'Ascoli, ou d'après 6.1) qui est, en vertu de la semi-continuité de la longueur, ou de l'énergie (6.1) une géodésique minimale d'extrémités x et y . Or c_n n'appartient pas à $\gamma(U \times V)$ et comme $\gamma(U \times V)$ est dans G un voisinage de γ_0 on a $c \neq \gamma_0$. Ce qui est absurde.

(ii) $\Leftrightarrow$ (i). Soit γ_0 une géodésique minimale d'extrémités x et y . La condition (ii) entraîne déjà que x et y ne sont pas conjugués le long de γ_0 . Supposons qu'il existe une seconde géodésique minimale c joignant x et y . On a alors pour $0 \leq t \leq 1$

$$d(x, c(t)) = L(c|_{[0,t]}) = td(x,y)$$

tandis que la formule (18) donne, puisque par hypothèse on a $d^\gamma = d$ sur $U \times V$,

$$\frac{d}{dt} d(x, c(t)) \Big|_{t=1} = \cos(\dot{\gamma}_0(1), \dot{c}(1)) d(x,y)$$

Il en résulte que $\dot{c}(1) = \dot{\gamma}_0(1)$ donc que c coïncide avec γ_0 .

7.7. PROPOSITION. (i) Pour tout point x de M , l'ensemble $C(x)$ est fermé et de mesure nulle dans M .

(ii) Pour tout point x de M , l'ensemble $C'(x)$ des points conjugués de x suivant au moins une géodésique minimale est fermé (et de mesure nulle) dans M .

Démonstration. La trace de $C'(x)$ sur la boule de centre x et de rayon ρ est exactement l'ensemble des valeurs critiques de $\exp_x$ correspondant aux points critiques de norme $\|v\| \leq \rho$. C'est donc l'image continue d'un compact, donc un fermé.

Un argument élémentaire de topologie montre alors que $C'(x)$ est fermé.

Nous savons déjà que $C'(x)$ est de mesure nulle (5.5) et que $C(x)$ est fermé. Il reste à montrer que $C(x) - C'(x)$ est de mesure nulle. Soit $y \in C(x) - C'(x)$. On a $y \neq x$ et x et y sont joints par un nombre fini de géodésiques minimales $\gamma_1, \dots, \gamma_k$. Pour $i = 1, \dots, k$ on peut définir, sur un voisinage convenable de x, y une application $(x', y') \rightarrow \gamma_i(x', y')$ de classe C^1 où $\gamma_i(x', y')$ est une géodésique d'extrémités x' et y' . Si x' et y' sont assez voisins, respectivement, de x et y toute géodésique allant de x' à y' distincte de $\gamma_1(x', y'), \dots, \gamma_k(x', y')$ est de longueur $> d(x, y)$. Le cut-locus de x est donc défini au voisinage de y par l'égalité d'au moins deux des fonctions $E^{\gamma_1}(x, y'), \dots, E^{\gamma_k}(x, y')$. Pour $i \neq j$ le gradient de $E^{\gamma_i} - E^{\gamma_j}$ au point y est $\dot{\gamma}_i(1) - \dot{\gamma}_j(1) \neq 0$.

Il en résulte que l'égalité $E^{\gamma_i}(x, y') = E^{\gamma_j}(x, y')$ définit au voisinage de y une hypersurface et que $C(x)$ est toujours dans un voisinage de y , une réunion finie d'hypersurfaces, donc de mesure nulle. On achève la démonstration au moyen d'un recouvrement dénombrable de $M - C'(x)$.

7.8. THÉOREME. Soit M une variété riemannienne de classe C^r ($r \geq 2$). La fonction $E(x, y) = \frac{1}{2} d^2(x, y)$ est différentiable, de classe C^{r-1} , sur le complémentaire du cut-locus $C(M)$. Sa différentielle est donnée par

$$(21) \quad dE = \langle \dot{\gamma}(1), dy \rangle - \langle \dot{\gamma}(0), dy \rangle$$

où γ désigne l'unique géodésique minimale d'extrémités x et y

Démonstration. Il suffit d'appliquer le lemme 7.5 puisque sur le complémentaire de $C(M)$ la géodésique minimale est unique et dépend de façon C^1 de ses extrémités.

7.9. COROLLAIRE. Soient $x, y \in M$ tels que $(x, y) \notin C(M)$. La dérivée seconde croisée $\frac{\partial^2 E}{\partial x \partial y}$ existe.

En posant $u = \dot{\gamma}(0)$ on a

$$\frac{\partial^2 E}{\partial x \partial y}(v, w) = - \langle (d(\exp_x)_u)^{-1}(v), w \rangle$$

pour tout $v \in T_x M$ et tout $w \in T_y M$.

On peut encore dire que $\frac{\partial^2 E}{\partial x \partial y}$ correspond à $(d(\exp_x)_u)^{-1}$ dans l'identification que fournit le produit scalaire entre formes bilinéaires sur $T_x^*M \times T_y^*M$ et applications linéaires de T_y^*M dans T_x^*M .

7.10. Rappelons que si deux espaces vectoriels F_1 et F_2 de même dimension sont munis de formes volumes ω_1 et ω_2 on peut donner un sens à la notion de déterminant d'une application linéaire $f : F_1 \rightarrow F_2$. De même on peut définir le déterminant d'une forme bilinéaire $b : F_1 \times F_2 \rightarrow \mathbb{R}$: on munit F_2^* de la forme volume duale ω_2^* et on prend pour $\det(b)$ le déterminant de l'application linéaire $\beta_b : F_1 \rightarrow F_2^*$ associée à b . Si F_1 et F_2 sont des espaces euclidiens on prendra pour ω_1 et ω_2 les formes volumes euclidiennes canoniques. Notons que l'identification entre F_2 et F_2^* identifie aussi ω_2 et ω_2^* .

7.11. Avec les notation de 7.9 on pose

$$(23) \quad \theta(x, y) = \det(d(\exp_x)_u)$$

Du corollaire 7.9 et de 7.10 on déduit aussitôt

$$(24) \quad \theta(x, y) = \det\left(\frac{\partial^2 E}{\partial x \partial y}\right)^{-1}.$$

En coordonnées locales la relation (22) s'écrit

$$\left(\frac{\partial^2 E}{\partial x_i \partial y_j}\right) = \left(\frac{\partial y_i}{\partial u_j}\right)^{-1} (g_{ij}(y)) \quad (\text{produit de matrices})$$

si l'on convient de noter $\frac{\partial y_i}{\partial u_j}$ la matrice jacobienne de l'application $\exp_x : u \mapsto y$. Les relations (23) et (24) deviennent

$$\begin{aligned} \theta(x, y) &= \left| \det\left(\frac{\partial y_i}{\partial u_j}\right) \right| \det(g_{ij}(x))^{1/2} \det(g_{ij}(y))^{-1/2} \\ &= \left| \det\left(\frac{\partial^2 E}{\partial x_i \partial y_j}\right) \right|^{-1} \det(g_{ij}(x))^{1/2} \det(g_{ij}(y))^{1/2} \end{aligned}$$

7.12. PROPOSITION. Soient x un point de M , v un vecteur de norme 1 dans $T_x M_x$ et soit $\gamma : [0, +\infty[\rightarrow M$ la géodésique définie par $\dot{\gamma}(0) = v$. Soit $\tau = \tau(v)$ la borne supérieure des $t \in [0, +\infty[$ pour lesquels la géodésique γ réalise la distance entre $\gamma(0)$ et $\gamma(t)$.

Alors on a $\gamma(t) \notin C(x)$ pour $0 \leq t < \tau$ et $\gamma(\tau) \in C(x)$ si $\tau < +\infty$.

Démonstration. Montrons tout d'abord que si $\tau < +\infty$ la géodésique $\gamma|_{[0, \tau]}$ est minimisante : on a $d(x, \gamma(t)) = L(\gamma|_{[0, t]}) = t$ pour $0 \leq t < \tau$ d'où on tire, par passage à la limite

$$d(x, \gamma(t)) = L(\gamma|_{[0, t]}) = t.$$

Montrons ensuite que $\gamma(t) \in C(x)$. Pour n entier > 0 , il existe, puisque γ n'est pas minimisante entre $\gamma(0)$ et $\gamma(t + \frac{1}{n})$ un vecteur $w_n \in T_x M$ tel que

$$\exp_x((t + \frac{1}{n})v) = \exp_x(w_n) \text{ et } \|w_n\| < t + \frac{1}{n}.$$

Soit w_0 une valeur d'adhérence de la suite (w_n) . On a nécessairement

$$\exp_x(tv) = \exp_x(w_0) \text{ et } \|w_0\| = t.$$

Si $w_0 \neq tv$ le point $\gamma(t) = \exp_x(tv)$ est joint à x par deux géodésiques de longueur minimale. Si $w_0 = tv$ l'application $\exp_x$ n'est injective sur aucun voisinage de tv . Le point tv est alors, en vertu du théorème d'inversion locale, un point critique de $\exp_x$ et les points x et $\gamma(t)$ sont conjugués le long d'une géodésique minimale.

Montrons maintenant que si $0 \leq t < \tau$ la restriction de γ à l'intervalle $[0, t]$ est l'unique géodésique minimale joignant x et $\gamma(t)$. Soit $\gamma_1 : [0, t] \rightarrow M$ une seconde géodésique minimale joignant x et $\gamma(t)$. Le chemin $\gamma_2 : [0, \tau] \rightarrow M$ défini par

$$\begin{aligned} \gamma_2(s) &= \gamma_1(s) & \text{si } 0 \leq s \leq t \\ \text{et } \gamma_2(s) &= \gamma(s) & \text{si } t \leq s \leq \tau \end{aligned}$$

minimise l'énergie entre x et $\gamma(\tau)$. C'est donc aussi une géodésique. Or deux géodésiques qui coïncident sur un intervalle coïncident sur tout leur domaine de définition puisqu'elles satisfont à la même équation différentielle du second ordre. On a donc $\gamma_2 = \gamma|_{[0, \tau]}$ et $\gamma_1 = \gamma|_{[0, t]}$.

Montrons enfin que si $0 \leq t < \tau$ les points $\gamma(0)$ et $\gamma(t)$ ne sont pas conjugués suivant γ . Soit $J_1 : [0, t] \rightarrow TM$ un champ de Jacobi le long de $\gamma|_{[0, t]}$ nul en $\gamma(0)$ et en $\gamma(t)$. Prolongeons J_1 par 0 sur l'intervalle $[t, \tau]$ de manière à obtenir un champ J_2 . La formule de la variation seconde (12) montre que

$$I_{\gamma|_{[0, \tau]}}(J_2, J_2) = I_{\gamma|_{[0, t]}}(J_1, J_1) = 0$$

Or, puisque $\gamma|_{[0, \tau]}$ est une géodésique minimale de la forme $I_{\gamma|_{[0, \tau]}}$ est positive. Un vecteur isotrope de cette forme appartient donc à son noyau, ce qui montre que J_2 est un champ de Jacobi. Comme $J_2(s)$ est nul pour $t \leq s < \tau$ et que les champs de Jacobi satisfont à une équation différentielle du second ordre on a $J_2 = 0$ et donc $J_1 = 0$.

Remarque. Le cut-locus de x est donc l'ensemble des points de la forme $\exp_x(\tau(v)v)$ avec $\|v\| = 1$ et $\tau(v) < +\infty$.

On peut montrer que la fonction $v \rightarrow \tau(v)$ qui prend ses valeurs dans l'intervalle $]0, +\infty]$ est continue. On déduit de là que, pour tout x , la variété M est réunion de $C(x)$ et d'un ouvert difféomorphe à une boule de $\mathbb{R}^m$ (cf. [6]).

7.13. Le rayon d'injectivité.

On appelle rayon d'injectivité de l'application exponentielle en x et on note $i(x)$ la borne supérieure des nombres $\rho \geq 0$ tels que la restriction de $\exp_x$ à la boule $\|v\| < \rho$ soit injective.

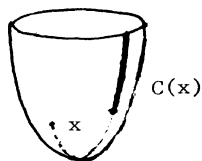
Si $d(x, y) < i(x)$ il existe donc une seule géodésique minimale d'extrémités x et y . Il résulte en fait de 7.12 que $i(x) = d(x, C(x))$. Il est facile de montrer que la fonction $x \rightarrow i(x)$ est continue.

7.14. Exemples de cut-locus (sans démonstration)

1) $M = S^m$. Le cut-locus d'un point x de M se réduit à un point : l'antipode de x .

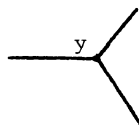
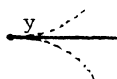
2) M est une surface de révolution et x est un pôle de M . Le cut-locus de x est composé d'un seul point : le pôle opposé.

3) M est dans $\mathbb{R}^3$ le parabolôïde d'équation $x_3 = x_1^2 + x_2^2$. Il existe un nombre $\lambda > 0$ tel que si $x_3 < \lambda$ le cut-locus de x est vide et si $x_3 \geq \lambda$ le cut-locus de x est une portion de la méridienne de x



4) Pour une métrique générique sur une surface donnée M , le cut-locus de tout point x de M a, au voisinage de chacun de ses points y l'une des formes suivantes, à difféomorphisme près :

- (i) une droite passant par y ;
- (ii) une demi-droite issue de y ;
- (iii) trois demi-droites issues de y , non situées dans un même demi-plan



Les points du type (ii) sont ceux de $C'(x)$ (on a tracé en pointillé le lieu conjugué de x (Buchner, [4])

5) Si M est compacte, tout point y réalisant le maximum de la distance à x appartient à $C(x)$.

6) Conséquence de la remarque qui suit le lemme 7.12 (cf. [6]) : si M n'est pas homéomorphe à $\mathbb{R}^m$ le cut-locus d'un point x n'est jamais vide (on rappelle que M est supposée complète).

Comme exemple de variétés dont le cut-locus est vide, citons les variétés simplement connexes à courbure négative (cf. Cheeger-Ebin [5]).

7.15. Le cut-locus d'une variété non complète.

Soit M une variété riemannienne non complète de classe C^2 au moins. Soit $\hat{M}$ l'espace métrique complété de M . On pose $\partial M = \hat{M} - M$.

On peut prolonger aux chemins continus $c : [a, b] \rightarrow \hat{M}$ les fonctionnelles longueur et énergie : on pose

$$L(c) = \sup \sum d(c(t_i), c(t_{i+1}))$$

et

$$E(c) = \sup \sum \frac{d(c(t_i), c(t_{i+1}))^2}{2(t_{i+1} - t_i)}$$

les bornes supérieures étant prises pour toutes les suites finies $(t_i)_{0 \leq i \leq n}$ telles que $a = t_0 < t_1 < \dots < t_n = b$.

Le minimum de l'énergie est encore réalisé par les chemins de longueur $d(x, y)$ parcourus uniformément (géodésiques minimales). On montre que deux points quelconques x et y de $\hat{M}$ peuvent être joints par une géodésique minimale : il suffit de prouver que dans $\hat{M}$ les boules fermées sont compactes et d'appliquer la proposition 6.2 qui reste valable.

Introduisons l'ensemble $B(M) \subset M \times M$ formé des couples (x, y) qui satisfont à l'une des conditions équivalentes :

(i) toutes les géodésiques minimales qui joignent les points x et y dans l'espace $\hat{M}$ sont situées dans M ;

(ii) pour un $\varepsilon > 0$ au moins, l'ensemble des chemins $c : [0, 1] \rightarrow M$ d'énergie

$\frac{1}{2} d^2(x,y) + \varepsilon$ est compact dans $C(I,M)$.

Définissons maintenant le cut-locus de M comme l'ensemble des couples $(x,y) \times B(M)$ satisfaisant l'une des conditions (i) ou (ii) de 7.1. L'ensemble $B(M)$ est ouvert dans $M \times M$ et toutes les propriétés du cut-locus subsistent à condition de remplacer $M \times M$ par $B(M)$. Par exemple $C(M)$ est fermé dans $B(M)$, l'énergie $E(x,y)$ est différentiable sur $B(M) - C(M)$, etc...

Pour $x \in M$ on note $\rho(x)$ la borne supérieure des $r > 0$ tels que l'application exponentielle soit définie sur la boule de rayon r dans $T_x M$. On peut dire que $\rho(x)$ est le rayon d'existence de l'application exponentielle. On a

$$\rho(x) = d(x, \partial M) = \liminf_{y \rightarrow \infty} d(x,y)$$

- ce dernier nombre est noté $d(x, \infty)$ dans les ch. 8 et 9 - et on voit que la condition

$$d(x,y) < \sup(\rho(x), \rho(y))$$

est une condition suffisante pour que $(x,y) \in B(M)$.

On peut encore définir si M n'est pas complète le rayon d'injectivité $i(x)$ comme la borne supérieure des $r < \rho(x)$ tels que la restriction de $\exp_x$ à la boule de rayon r soit injective.

8. Utilisation d'espaces de chemins de dimension finie.

8.1. Notre principal instrument dans l'étude de la géométrie des géodésiques d'une variété riemannienne M a été jusqu'à présent la caractérisation des géodésiques joignant deux points donnés x et y comme points critiques de la fonction énergie E sur la variété de dimension infinie Ω_{xy} . Nous allons montrer (dans le cas des géodésiques minimales) qu'on peut dans cette étude remplacer Ω_{xy} par une sous-variété de dimension finie, formée de géodésiques brisées.

L'intérêt de cette réduction à la dimension finie apparaîtra dans l'utilisation qui sera faite des notions introduites dans ce chapitre : en effet dans l'évaluation asymptotique de $p(t,x,y)$ telle qu'elle est faite au ch.9 la fonction énergie et les géodésiques minimales n'interviennent pas directement en tant que telles ⁽¹⁾ : on est

(1) Par contre, dans l'exposé d'Azencott [2], l'énergie intervient de façon naturelle comme une fonctionnelle sur l'espace des chemins absolument continus. (Les résultats obtenus ne sont pas de même ordre que ceux du ch.9).

amené à considérer sur M^n la fonction

$$\eta(y_1, \dots, y_n) = \frac{n+1}{2} [d^2(x, y_1) + d^2(y_1, y_2) + \dots + d^2(y_n, y)]$$

et à étudier le comportement (dégénérescence, ...) de η au voisinage de ses minimums. Or (on va le montrer) si $(y_1, \dots, y_n) \in W$ où W est un voisinage convenable de l'ensemble des minimums de η il existe un chemin unique $\gamma(x, y_1, \dots, y_n, y) \in \Omega_{xy}$ composé de géodésiques minimales successives joignant (en temps $\frac{1}{n+1}$) le point x à y_1 , y_1 à $y_2, \dots, y_n$ à y . L'énergie de ce chemin est exactement

$$\frac{n+1}{2} [d^2(x, y_1) + d^2(y_1, y_2) + \dots + d^2(y_n, y)].$$

On ne change donc rien en considérant au lieu de η la restriction de E à la sous-variété formée par les chemins $\gamma(y_1, \dots, y_n)$ où $(y_1, \dots, y_n) \in W$. Si maintenant on prouve que remplacer cette sous-variété par Ω_{xy} ne change rien à certaines propriétés de E (localisation des minimums, dégénérescence, etc...) on aura obtenu une relation simple entre les propriétés des minimums de η et les propriétés des géodésiques joignant x et y .

8.2. On suppose que M est une variété riemannienne de classe C^r ($r \geq 2$). Etant donnés $x, y \in M$ et un entier $n \geq 0$, on note Ω_{xy}^n la partie de Ω_{xy} formée par les chemins $c : [0, 1] \rightarrow M$ tels que pour $i = 0, 1, \dots, n$ la restriction de γ à l'intervalle

$[\frac{i}{n+1}, \frac{i+1}{n+1}]$ soit une géodésique. Nous appellerons géodésiques brisées n fois les éléments de Ω_{xy}^n .

L'espace Ω_{xy}^n contient - c'est évident - toutes les géodésiques joignant x à y et les minimums de la fonction E sont donc les mêmes sur Ω_{xy} et sur Ω_{xy}^n .

On note encore $\Omega_{xy}^{n,h}$, pour $h > 0$, l'ouvert défini dans Ω_{xy}^n par la condition $E(\gamma) < \frac{1}{2} d^2(x, y) + h$. L'ensemble Ω_{xy}^r n'est pas en général une sous-variété de Ω_{xy} , mais pour n et h convenables, l'ensemble $\Omega_{xy}^{n,h}$ en est une - ce qui permet de considérer comme des points critiques les minimums de E sur Ω_{xy}^n - et de plus peut être mis naturellement en bijection avec un ouvert W de M^n :

8.3. THEOREME. Soit M une variété riemannienne de classe C^r ($r \geq 3$) et soient x et y deux points de M . Supposons $(x, y) \in B(M)$ - ce qui est toujours le cas si M est complète.

Alors il existe $H > 0$ et N entier > 0 tels que les propriétés suivantes soient vérifiées si $0 < h \leq H$ et $n \geq N$:

(i) l'espace $\Omega_{xy}^{n,h}$ est une sous-variété de Ω_{xy} , de classe C^{r-1} , et l'application $c \rightarrow (c(\frac{i}{n+1}))_{1 \leq i \leq n}$ est un difféomorphisme de cette sous-variété sur l'ouvert $M_{xy}^{n,k}$ de M^n défini par la condition

$$(25) \quad \frac{n+1}{2} \sum_{i=0}^n d^2(y_i, y_{i+1}) < \frac{1}{2} d^2(x, y) + h \quad (y_0 = x, y_{n+1} = y).$$

(ii) Les géodésiques minimales joignant x à y sont les minimums de la restriction de E à $\Omega_{xy}^{n,h}$.

(iii) Pour que x et y soient conjugués suivant une géodésique minimale γ , il faut et il suffit que γ soit un minimum dégénéré de E sur $\Omega_{xy}^{n,h}$ et l'ordre de la conjugaison est égal au corang du Hessien au point γ de $E|_{\Omega_{xy}^{n,h}}$. Pour que y appartienne au cut-locus de x il faut et il suffit que la fonction $E|_{\Omega_{xy}^{n,h}}$ possède un minimum dégénéré ou atteigne sa valeur minimale en plusieurs points.

Démonstration de (i). (On suppose M complète). Il existe un compact K de M et $H > 0$ tels que tout chemin $c \in \Omega_{xy}$ d'énergie $< \frac{1}{2} d^2(x, y) + H$ soit contenu dans K . Soit $i(K)$ la borne inférieure, quand x décrit K , du rayon d'injectivité $i(x)$. Soit alors h tel que $0 < h \leq H$ et soit n un entier > 0 . Soit $(y_1, \dots, y_n)$ un point de $M_{xy}^{n,h}$. On peut

trouver pour chaque entier i ($i=0, \dots, n$) un chemin $c_i : [\frac{i}{n+1}, \frac{i+1}{n+1}] \rightarrow M$ joignant y_i à y_{i+1} , d'énergie $< \frac{1}{2} d^2(y_i, y_{i+1}) + \frac{h}{n+1}$. En mettant bout à bout de tels chemins, on peut obtenir un chemin c tel que $c(\frac{i}{n+1}) = y_i$ ($i=0, 1, \dots, n+1$) et

$$\frac{n+1}{2} \sum_{i=0}^n d^2(y_i, y_{i+1}) < E(c) < \frac{1}{2} d^2(x, y) + h.$$

Il en résulte que la condition $(y_1, \dots, y_n) \in M_{xy}^{n,h}$ entraîne $y_i \in K$ et

$$d(y_i, y_{i+1}) < \left(\frac{d^2(x, y) + 2h}{n+1} \right)^{1/2}.$$

Soit N le plus petit entier tel que

$$\left(\frac{d^2(x, y) + 2h}{N+1} \right)^{1/2} \leq i(K).$$

On suppose désormais $n \geq N$. Pour toute suite de points $y_1, \dots, y_n$ tels que

$(y_1, \dots, y_n) \in M_{xy}^{n,h}$ il existe donc une unique géodésique minimale

$\gamma_i : [\frac{i}{n+1}, \frac{i+1}{n+1}] \rightarrow M$ joignant y_i à y_{i+1} ($i=0, \dots, n$). En juxtaposant $\gamma_0, \gamma_1, \dots, \gamma_n$ on obtient une géodésique brisée joignant x à y que nous noterons $\gamma(y_1, \dots, y_n)$. L'application $\phi : M_{xy}^{n,h} \rightarrow \Omega_{xy}$ est injective, puisque $d(y_i, y_{i+1}) < i(K)$, et a pour image $\Omega_{xy}^{n,h}$. C'est une immersion, de classe C^{r-1} : on sait en effet que l'application définie sur le complémentaire du cut-locus qui associe à un couple de points l'unique géodésique minimale qui les joints est une immersion de classe C^{r-1} de cet ouvert dans M et il est clair que mettre $(n+1)$ chemins bout à bout est une opération différentiable. Pour montrer que l'image $\Omega_{xy}^{n,h}$ de ϕ est une sous-variété et ϕ un difféomorphisme sur celle-ci, il suffit de montrer que ϕ définit un homéomorphisme de $M_{xy}^{n,h}$ sur son image ce qui est évident, car l'application $c \rightarrow (c(\frac{i}{n+1}))_{1 \leq i \leq n}$ est une application continue de Ω_{xy} dans M^n .

8.4. LEMME. (*Lemme de Morse pour un minimum*). Soit H un espace de Hilbert et soit f une fonction de classe C^r ($r \geq 2$) définie au voisinage de l'origine et admettant en ce point un minimum non dégénéré.

Il existe des voisinages U et V de 0 dans H , un difféomorphisme $\phi : U \rightarrow V$ de classe C^{r-1} , tel que $\phi(0) = 0$ et que

$$(f \circ \phi)(x) = \|x\|^2 + f(0)$$

pour tout $x \in U$.

Démonstration. L'opérateur auto-adjoint $A : H \rightarrow H$ défini par

$$\frac{1}{2} f''(0)(u, v) = (u | Av)$$

est positif et inversible. En remplaçant au besoin f par $f \circ A^{-1/2}$ on peut se ramener au cas où A est l'identité. On peut donc supposer pour démontrer le lemme que

$$\frac{1}{2} f''(0).v^2 = \|v\|^2$$

on peut aussi supposer $f(0) = 0$. On a alors, puisque f est de classe C^2 ,

$$(26) \quad f(x) = \|x\|^2 + o(\|x\|^2)$$

quand x tend vers zéro et, en identifiant l'espace de Hilbert H avec son dual,

$$(27) \quad f'(x) = 2x + o(\|x\|)$$

L'idée de la construction du difféomorphisme ϕ est simple : pour $\lambda > 0$ assez petit l'hypersurface $f^{-1}(\lambda)$ et la sphère $\|x\|^2 = \lambda$ sont proches, y compris au sens C^1 , en vertu de (26) et (27) et les rayons de la sphère coupent donc transversalement l'hypersurface $f^{-1}(\lambda)$. Si $x \neq 0$ est assez proche de 0 il existe donc sur le rayon passant par x un seul point y tel que $f(y) = \|x\|^2$. On pose alors $\phi(x) = y$ ⁽¹⁾.

Pour être plus précis, fixons $\varepsilon > 0$ de manière que pour $\|x\| < \varepsilon$ on ait

$$|f(x) - \|x\|^2| < \frac{1}{2} \|x\|^2$$

et

$$\|f'(x) - 2x\| < \|x\|$$

Si $0 < \|x\| < \frac{\varepsilon}{2}$ l'équation (en t) $f(tx) = \|x\|^2$ admet une seule solution $t \in]0, 2[$ en effet on a $f(0) = 0$, $f(2x) > 2\|x\|^2$ et pour $t \in]0, 2[$

$$\frac{d}{dt} f(tx) = (f'(tx) | x) \geq (2tx | x) - t(x | x) \geq 0$$

On note $\theta(x)$ cette solution, et on pose en prenant pour U la boule $\|x\| < \varepsilon/2$, $\phi(x) = \theta(x)x$ pour $x \neq 0$ et $\phi(0) = 0$.

Le théorème des fonctions implicites nous permet de montrer que l'application $x \rightarrow \theta(x)$ est de classe C^r sur $U - \{0\}$: la différentielle de $F(t, x) = f(tx) - \|x\|^2$ est en effet

$$\langle f'(tx), x \rangle dt + t \langle f'(tx), dx \rangle - 2 \langle x, dx \rangle$$

et pour $0 < \|x\| < \varepsilon/2$ on a $(f'(tx) | x) \neq 0$.

On a, sur $U - \{0\}$,

$$(28) \quad d\theta = \frac{1}{\langle f'(\theta(x)x), x \rangle} \langle 2x - \theta(x)f'(\theta(x)x), dx \rangle.$$

L'application $x \rightarrow \phi(x) = \theta(x)x$ est donc elle aussi de classe C^r sur $U - \{0\}$ et sa différentielle est

$$\phi'(x) = \theta(x).Id + x \otimes \theta'(x)$$

(1) Cette méthode est due à N. Kuiper et son emploi m'a été suggéré par N. Desolneux-Moulis.

Nous savons déjà (puisque θ est bornée) que ϕ est continue au point 0. Pour montrer que ϕ est C^1 sur U il nous suffit de prouver que $d\phi$ a une limite quand x tend vers 0. Il résulte de (26) que $\theta(x) \rightarrow 1$ quand $x \rightarrow 0$. On en déduit, en utilisant (27) et (28) que

$$\theta(x) g'(\theta(x)x) - 2x = O(\|x\|)$$

et

$$x \theta'(x) = O(\|x\|)$$

Par conséquent $\phi'(x) \rightarrow \text{Id}$ quand $x \rightarrow 0$. Ceci montre en même temps que ϕ est un difféomorphisme local en 0, de classe C^1 . Une démonstration analogue permet de prouver que ϕ est de classe C^{r-1} .

8.5. Le lemme de Morse, tel qu'il est énoncé en 8.4, ne suffit pas à nos besoins. Il nous faut encore savoir que si dans la construction précédente la fonction f dépend continûment (resp. différenciablement) de certains paramètres le difféomorphisme dépend continûment (resp. différenciablement) des mêmes paramètres.

DÉFINITIONS. Soit H un espace normé, Λ un espace topologique et soit f une fonction définie sur un ouvert U de $H \times \Lambda$. On dit que f est de classe $C^{r,0}$ sur les dérivées $D_x^p f(x, \lambda)$ existent et sont continues pour tout $p \leq r$. De même si Λ est un espace normé on dit que f est de classe $C^{r,s}$ si les dérivées $D_x^p D_\lambda^q f(x, \lambda)$ existent et sont continues sur U pour tout $p \leq r$ et tout $q \leq s$. Cette dernière définition s'étend bien sur au cas où Λ est une variété.

8.6. LEMME. (Lemme de Morse avec paramètres). Soit U un voisinage de l'origine dans un espace de Hilbert et soit Λ une variété de classe C^s (resp. un espace topologique). Soit f une fonction définie sur $U \times \Lambda$, de classe $C^{r,s}$ avec $r \geq 2$ et $0 \leq s \leq r$ (resp. $s = 0$), telle que pour tout $\lambda \in \Lambda$ la fonction $f_\lambda : x \rightarrow f(x, \lambda)$ admette au point 0 un minimum non dégénéré.

Alors il existe une application ϕ de classe $C^{r-1,s}$ définie sur un voisinage de $\{0\} \times \Lambda$ dans $U \times \Lambda$ telle que pour tout λ on ait $\phi(\lambda, 0) = 0$, que $\phi_\lambda : x \rightarrow \phi(\lambda, x)$ soit un difféomorphisme local en 0 et que l'on ait

$$f_\lambda(\phi_\lambda(x)) = f_\lambda(0) + \|x\|^2$$

dès que le premier membre est défini.

Démonstration. On recopie, avec quelques précautions, la démonstration du lemme 8.4.

8.7. LEMME. Soit X une variété hilbertienne de classe C^k ($k \geq 2$). Soient Y une sous-variété de X et $p : X \rightarrow Y$ une application différentiable de classe C^k telle que $p(y) = y$ pour tout $y \in Y$.

(i) Soit $y_0 \in Y$. Il existe une carte $\phi : W \rightarrow H' \times H''$ de classe C^r , où W est un voisinage ouvert de y_0 dans X et où H' et H'' sont des espaces de Hilbert convenables, telle que l'on ait

$$\phi(y_0) = (0, 0),$$

$$\phi(Y \cap W) = (H' \times \{0\}) \cap \phi(W),$$

et $\phi(p^{-1}(y) \cap W) = (\{\phi(y)\} \times H'') \cap \phi(W)$

pour tout $y \in Y \cap W$. En particulier l'ensemble $p^{-1}(y)$ est une sous-variété au voisinage du point y .

(ii) Soit $f : X \rightarrow \mathbb{R}$ une fonction de classe C^r telle que pour tout $y \in Y$ la restriction de f à $p^{-1}(y)$ admette au point y un minimum unique non dégénéré.

Alors on peut pour $y_0 \in Y$ trouver une carte $\tilde{\phi} : W \rightarrow H' \times H''$ de classe C^{r-1} , dont la restriction à $Y \cap W$ soit de classe C^r , satisfaisant aux mêmes conditions que ϕ ci-dessus et telle que de plus

$$(f \circ \tilde{\phi}^{-1})(u, v) = f(\tilde{\phi}^{-1}(u, 0)) + \|v\|^2$$

pour tout $u \in U$ et tout $v \in V$.

Démonstration. (i) L'application linéaire $dp_{y_0} : T_{y_0} X \rightarrow T_{y_0} Y$ est un projecteur. Elle

est donc surjective et on peut trouver des cartes $\psi : W \rightarrow T_{y_0} X$ et $\theta : W' \rightarrow T_{y_0} Y$, définies dans X et Y respectivement, au voisinage de y_0 telles que $\psi(y_0) = \theta(y_0) = 0$,

$p(W) = W'$ et $\theta \circ p \circ \psi^{-1} = dp_{y_0}$. Posons $H' = T_{y_0} Y$ et $H'' = \ker dp_{y_0}$ de sorte que

$T_{x_0} X = H' \oplus H'' \xrightarrow{\sim} H' \times H''$. Si $y \in Y$ on a $\psi(y) - \theta(y) \in H''$. L'application $\phi : W \rightarrow H' \times H''$

définie par $\phi(x) = \psi(x) - \psi(p(x)) + \theta(p(x))$ répond à la question.

Pour prouver (ii) il suffit d'appliquer le lemme 8.5 à la fonction $(f \circ \phi^{-1})(u', u'')$ où u' est le paramètre et u'' la variable.

8.8. Démonstration de l'assertion (iii) du théorème 8.3.

On applique le lemme 8.6 en prenant pour X l'ouvert défini dans Ω_{xy} par la condition $E(c) < \frac{1}{2} d^2(x, y) + h$, pour Y la sous-variété $\Omega_{xy}^{n, h}$, pour $p : X \rightarrow Y$ l'application qui à un chemin $c \in X$ associe l'unique géodésique brisée $\gamma \in \Omega_{xy}^{n, h}$ telle que $\gamma(\frac{i}{n+1}) = c(\frac{i}{n+1})$ ($i=1, \dots, n$) et pour f la fonction énergie. L'application p est de classe C^{r-1} en vertu de 8.3.i). Pour prouver que γ est minimum non dégénéré de l'énergie sur $p^{-1}(\gamma)$ on remarque que l'espace tangent en γ à cette sous-variété est l'espace des champs de vecteurs le long de γ nuls pour $t = \frac{i}{n+1}$ ($i=0, 1, \dots, n+1$). Posons $\gamma_i = \gamma|[\frac{i}{n+1}, \frac{i+1}{n+1}]$ et pour $V \in T_{\gamma} p^{-1}(\gamma)$ posons $V_i = V|[\frac{i}{n+1}, \frac{i+1}{n+1}]$. Puisque chacune des formes I_{γ_i} est non dégénérée on obtient en appliquant la formule de la variation seconde (12)

$$\text{Hess}_{\gamma}(E|p^{-1}(\gamma))(V, V) = \sum_{i=0}^n I_{\gamma_i}(V_i, V_i) > 0$$

si $V \neq 0$. Les conclusions de 8.7 montrent alors que les Hessiens en γ des restrictions de E à Ω_{xy} et à $\Omega_{xy}^{n, h}$ ont le même noyau et iii) en découle, en tenant compte des définitions 5.1 et 7.1.

8.9. Nous venons de voir que pour étudier la dégénérescence de la fonction énergie au voisinage de ses minimums sur Ω_{xy} on peut remplacer Ω_{xy} par une sous-variété, contenant les minimums, de dimension finie, en l'occurrence mn (n suffisamment grand). Nous allons montrer maintenant qu'on peut abaisser cette dimension et pour faire l'étude de la dégénérescence de E au voisinage d'une géodésique minimale donnée se ramener à une sous-variété de dimension $m-1$.

Soient x et y deux points de M et soit $\gamma : [0, 1] \rightarrow M$ une géodésique minimale joignant ces deux points. Considérons une hypersurface Z de M orthogonale à la géodésique γ en son milieu $z = \gamma(\frac{1}{2})$. On notera Ω_{xy}^Z l'ensemble des chemins $c : [0, 1] \rightarrow M$ tels que $c(\frac{1}{2}) \in Z$ et dont les restrictions à $[0, \frac{1}{2}]$ et $[\frac{1}{2}, 1]$ sont des géodésiques.

8.10. THÉOREME. (Suite de 8.3). Soit $\gamma : [0, 1] \rightarrow M$ une géodésique minimale joignant x et y . Alors il existe une hypersurface Z de M orthogonale en son milieu à la géodésique γ telle que les propriétés suivantes soient vérifiées si $0 < h \leq H$ et si n

est un entier impair $\geq N$:

(i) l'ensemble Ω_{xy}^Z est une sous-variété de classe C^{r-1} de $\Omega_{xy}^{n,k}$ et l'application $c \rightarrow c(\frac{1}{2})$ est un difféomorphisme de cette sous-variété sur un ouvert de Z contenant $\gamma(\frac{1}{2})$

(ii) il existe une carte $\phi_n : W_n \rightarrow \mathbb{R}^{mn}$ définie son voisinage ouvert W_n de γ dans $\Omega_{xy}^{n,h}$ telle que

$$\begin{aligned}\phi_n(\gamma) &= 0, \\ \phi_n(W_n \cap \Omega_{xy}^Z) &= \phi_n(W_n) \cap (\mathbb{R}^{m-1} \times \{0\})\end{aligned}$$

et

$$(E \circ \phi_n^{-1})(u_1, \dots, u_{mn}) = (E \circ \chi^{-1})(u_1, \dots, u_{m-1}) + u_m^2 + \dots + u_{mn}^2,$$

en notant χ la restriction de ϕ_n à $W_n \cap \Omega_{xy}^Z$.

Démonstration. En vertu de la proposition 7.12 le point $z = \gamma(\frac{1}{2})$ n'appartient ni au cut-locus de x ni à celui de y . En restreignant au besoin Z il en sera de même de tout point z' de Z . On démontre alors i) de la même façon que 8.3.i).

Pour n impair et à condition de restreindre encore Z on a l'inclusion $\Omega_{xy}^Z \subset \Omega_{xy}^{n,h}$. Les géodésiques brisées $c \in \Omega_{xy}^{n,h}$ qui rencontrent transversalement la sous-variété Z forment un ouvert de $\Omega_{xy}^{n,h}$. Nous allons appliquer le lemme 8.7 en prenant pour X cet ouvert, pour Y la sous-variété Ω_{xy}^Z , pour f la fonction énergie et pour $p : X \rightarrow Y$ l'application ainsi définie : si $c \in X$ on note $\xi(c)$ l'unique élément γ de Ω_{xy}^Z tel que $\gamma(\frac{1}{2}) = \xi(c)$. La vérification des hypothèses de 8.7 se fait comme dans 8.8 et le théorème est démontré.

Nous terminons ce chapitre par la démonstration de deux lemmes techniques qui seront utilisés au ch.9. La proposition 8.11 constitue la seule application dans cet ouvrage du th. 8.10 et n'interviendra que de façon accessoire (ch.9, 4.6 et 4.9). La formule (2.9) de 8.12 est, par contre, fondamentale.

8.11. Le théorème 8.10 affirme que la fonction énergie peut se mettre dans des coordonnées convenables au voisinage d'une géodésique minimisante sur $\Omega_{xy}^{n,h}$ sous la forme $f(u_1, \dots, u_{m-1}) + u_m^2 + \dots + u_{mn}^2$. Réciproquement on a la

PROPOSITION. Soit M une variété de classe C^∞ et soient x et y deux points distincts de M . Alors pour toute fonction f de classe C^∞ définie sur un voisinage de 0 dans

$\mathbb{R}^{m-1}$ et admettant au point 0 un minimum avec la valeur $f(0) = 0$ il existe une métrique riemannienne complète g sur M et pour cette métrique une géodésique minimale γ joignant x et y telle que pour $h > 0$ assez petit et n assez grand il existe une carte $\phi_n : W_n \rightarrow \mathbb{R}^{mn}$ définie au voisinage de γ dans $\Omega_{xy}^{n,h}$ telle que

$$\phi_n(\gamma) = (0, 0, \dots, 0)$$

et

$$(E \circ \phi_n^{-1})(u_1, \dots, u_{mn}) = E(\gamma) + f(u_1, \dots, u_{m-1}) + u_m^2 + \dots + u_{mn}^2.$$

Démonstration. Nous traiterons le cas de $M = \mathbb{R}^m$, en prenant $x = (0, 0, \dots, 0)$ et $y = (1, 0, \dots, 0)$ le cas général s'en déduisant par des arguments de partition de l'unité.

En vertu du théorème 8.9 il suffit de trouver une métrique riemannienne g sur $\mathbb{R}^m$ telles que

1) le chemin $\gamma : t \rightarrow (t, 0, \dots, 0)$ ($0 \leq t \leq 1$) soit une géodésique minimale ;

2) Il existe une surface Σ orthogonale en son milieu $z = (\frac{1}{2}, 0, \dots, 0)$ à Σ et des coordonnées $\theta_1, \dots, \theta_{m-1}$ sur Σ , nulles en z , telles que

$$\frac{1}{2}[d^2(x, \sigma) + d^2(\sigma, y)] = f(\theta_1, \dots, \theta_{m-1})$$

si σ désigne le point de coordonnées $\theta_1, \dots, \theta_{m-1}$.

Décrivons brièvement la construction de g : on prend pour surface Σ l'ensemble des points de coordonnées sphériques $\frac{1}{2} + f(\theta_1, \dots, \theta_{m-1})$, $\theta_1, \dots, \theta_{m-1}$. On considère au voisinage du segment xy un champ de vecteur X qui pour $1/3 < x_1 < 2/3$ coïncide avec le champ des normales à Σ , pour $x_1 > 5/6$ est le champ des vecteurs unitaires dirigé vers y et dont la composante suivant x_1 est toujours > 0 . On modifie X en le multipliant par une fonction à support compact dans $\frac{2}{3} < x_1 < \frac{5}{6}$ de manière à ce que les courbes intégrales $t \rightarrow \exp(tY)\sigma$ ($\sigma \in \Sigma$) du champ Y obtenu joignent toutes Σ à y dans le temps $\frac{1}{2}$. On modifie maintenant la métrique riemannienne canonique de $\mathbb{R}^m$ dans l'ensemble U_2 défini par

$$\frac{2}{3} < x_1 < \frac{5}{6}, \quad x_2^2 + \dots + x_m^2 < 4\varepsilon^2$$

(ε assez petit), de manière à ce que dans l'ensemble U_1 défini par

$$\frac{2}{3} < x_1 < \frac{5}{6}, \quad x_2^2 + \dots + x_m^2 < \varepsilon^2$$

elle s'écrit sous la forme

$$dt^2 + \sum_{i,j=1}^{m-1} g_{ij}(t, \theta_1, \dots, \theta_{m-1}) d\theta_i d\theta_j,$$

les coordonnées $t, \theta_1, \dots, \theta_{m-1}$ étant définies par l'application

$(t, \theta_1, \dots, \theta_{m-1}) \rightarrow \exp(tY)\sigma(\theta_1, \dots, \theta_{m-1})$, tout en veillant à ce que tout chemin joignant x à y et ne rencontrant pas U_1 soit de longueur >1 . Pour la métrique g obtenue au voisinage du segment xy , les géodésiques joignant $\bar{\gamma}$ à y sont les courbes intégrales de Y et les géodésiques joignant x à $\bar{\gamma}$ sont les rayons issus de x . Si on a de plus pris la précaution plus haut de garder égales à $\frac{\partial}{\partial x_1}$ les valeurs de $X(x_1, 0, \dots, 0)$ et $Y(x_1, 0, \dots, 0)$ pour $0 \leq x_1 \leq 1$ de sorte que $\exp(tY)z = z + t(\frac{1}{2}, 0, \dots, 0)$ la métrique g répond aux conditions 1) et 2).

8.12. PROPOSITION. (Notations de 8.3). On suppose que la variété riemannienne M est complète et que le couple (x, y) n'appartient pas au cut-locus de M . On note γ l'unique géodésique minimisante joignant x et y .

Soient h tel que $0 < h \leq H$ et n un entier $\geq N$. Posons $y_1 = \gamma(\frac{1}{n+1}), \dots, y_n = \gamma(\frac{n}{n+1})$. Alors on a

$$(29) \det \text{Hess} (E|_{\Omega_{xy}^{n,h}}) = (n+1)^{m(n+1)} \theta(x, y_1)^{-1} \theta(y_1, y_2)^{-1} \dots \theta(y_n, y)^{-1} \theta(x, y).$$

(voir 7.11 pour la définition de $\theta(x, y)$).

Démonstration. La démonstration de la formule (29) repose sur une nouvelle interprétation géométrique de la quantité $\theta(x, y)$.

Introduisons d'abord une forme volume ω sur le fibré tangent TM , liée à la métrique riemannienne g de M . Si $x_1, \dots, x_n, v_1, \dots, v_m$ désignent des coordonnées locales sur TM correspondant à un système de coordonnées locales $x_1, \dots, x_m$ sur M (cf. ch.3) on pose

$$\omega = \det(g_{ij}(x)) dx_1 \dots dx_m dv_1 \dots dv_m.$$

En notant π la forme volume canonique sur M et π_x la forme volume correspondant sur $T_x M$, on a $\int_{TM} f(x, v) \omega = \int_M \pi(dx) \int_{T_x M} f(x, v) \pi_x(dv)$. On écrira avec un certain abus

$$\omega = \omega(dx \, dv) = \pi(dx) \pi(dv).$$

On appelle flot géodésique de M le groupe à un paramètre (ϕ_t) de difféomorphismes

de TM ainsi défini : pour $(x,v) \in TM$ on pose

$$\phi_t(x,v) = (\gamma(t), \dot{\gamma}(t))$$

où γ désigne la géodésique déterminée pour $\gamma(0) = x$, $\dot{\gamma}(0) = v$. Le flot géodésique laisse invariante la forme volume ω . En effet ω est la valeur absolue de la $2m$ -forme $d\alpha \wedge \dots \wedge d\alpha$ où $\alpha = \langle v, dx \rangle = \sum_{i=1}^m g_{ij}(x) v_i dx_j$. Il suffit donc de montrer que $\phi_t^*(d) = d$ ou encore que $\phi_t^* \alpha = \alpha + \beta$ avec $d\beta = 0$, ce qui (pour $t=1$) est exactement le contenu de la proposition 7.5.

Considérons maintenant le graphe Γ_t de ϕ_t . En vertu de l'invariance de ω , la sous-variété Γ_t de $TM \times TM$ se trouve munie naturellement d'une forme ω_t que l'on peut définir comme l'image réciproque de ω par la projection de Γ_t aussi bien sur le premier que sur le second facteur. Nous utiliserons ces projections (cf. diagramme)



suivies éventuellement de cartes canoniques de TM comme des "cartes" de Γ_t . On a ainsi

$$\omega_t = \pi(dx)\pi(dv) = \pi(dy)\pi(dw).$$

Si (x,y) n'appartient pas au cut-locus et si $v = \dot{\gamma}(0)$, $\dot{w} = \dot{\gamma}(t)$ pour l'unique géodésique minimisante $\gamma : [0,t] \rightarrow M$ qui joint x et y (on suppose $t > 0$) la projection $p_t : \Gamma_t \rightarrow M \times M$ définie par $p_t(x,v,y,w) = (x,y)$ est régulière au voisinage de x_0, v_0, y_0, w_0 et détermine une nouvelle "carte" de Γ_t . L'expression de Γ_t dans cette carte est

$$\omega_t = t^m \theta(x,y) \pi(dx) \pi(dy).$$

En effet on a $\omega_t = \lambda(x,y) \pi(dx) \pi(dy)$ où $\lambda(x,y)$ est le jacobien de l'application $p_t \circ pr_1^{-1} : (x,v) \rightarrow (x,y)$ par rapport aux formes $\pi(dx)\pi(dv)$ et $\pi(dx)\pi(dy)$. Comme cette dernière application n'est autre que l'application $(x,v) \rightarrow (x, \exp_x tv)$ on voit que $\lambda(x,y)$ est t^m fois le déterminant au point tv de $\exp_x$, par rapport aux formes $\pi(dv)$ et $\pi(dy)$, soit $\lambda(x,y) = t^m \theta(x,y)$.

Prouvons maintenant la formule (29). On définit une immersion

$$j : \Gamma_1 \rightarrow \Gamma_1 \times \dots \times \Gamma_1 \quad ((n+1) \text{ fois}) \text{ de la façon suivante : pour } (x,v,y,w) \in \Gamma_1 \text{ soit}$$

$$\frac{1}{n+1} \quad \frac{1}{n+1}$$

γ la géodésique qui vérifie $\gamma(0) = x$, $\dot{\gamma}(0) = v$, $\gamma(1) = y$, $\dot{\gamma}(1) = w$; on pose $y_i = \gamma(\frac{i}{n+1})$, $w_i = \dot{\gamma}(\frac{i}{n+1})$ ($1 \leq i \leq n$) et

$$j(x, v, y, w) = ((x, v, y_1, w_1), (y_1, w_1, y_2, w_2), \dots, (y_n, w_n, y, w)).$$

L'image de j est une sous-variété de dimension $2m$ de $\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}$ que l'on peut définir localement de la façon suivante : soit $x^\circ, v^\circ, y^\circ, w^\circ$ un point de Γ_1 on note $\gamma^\circ : [0, 1] \rightarrow M$ la géodésique correspondante et on pose $y_i^\circ = \gamma^\circ(\frac{i}{n+1})$, $w_i^\circ = \dot{\gamma}^\circ(\frac{i}{n+1})$. Fixons les cartes de M au voisinage de $x^\circ, y_1^\circ, \dots, y_n^\circ$. Au voisinage des points (x°, v°) , $(y_1^\circ, w_1^\circ), \dots, (y_n^\circ, w_n^\circ)$ dans TM on considère les cartes correspondantes.

Le produit $\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}$ est alors dans $(TM \times TM) \times \dots \times (TM \times TM)$ l'ensemble des points

$$((x, v, y_1', w_1'), (y_1', w_1', y_2', w_2'), \dots, (y_n, w_n, y, w))$$

tels que $y_{i+1}', w_{i+1}' = \phi_{\frac{1}{n+1}}(y_i, w_i)$ pour $i=0, \dots, n$ et $j(\Gamma_1)$ est dans $\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}$ l'ensemble des points tels que

$$y_1 = y_1', w_1 = w_1', \dots, y_n = y_n, w_n = w_n'.$$

Il est facile de voir que l'application

$$H : ((x, v, y_1', w_1'), (y_1, w_1, y_2', w_2'), \dots, (y_n, w_n, y, w)) \rightarrow ((x, v, y, w),$$

définie au moyen des cartes ci-dessus est localement un difféomorphisme de $\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}$ sur $\Gamma_1 \times \mathbb{R}^{2m} \times \dots \times \mathbb{R}^{2m}$ et que H fait se correspondre $j(\Gamma_1)$ et

$\Gamma_1 \times \{0\} \times \dots \times \{0\}$. Seule en fait l'application tangente à H en un point

$j(x^\circ, v^\circ, y^\circ, w^\circ)$ de $j(\Gamma_1)$ a un sens intrinsèque, comme application

$$dH : T_{j(x^\circ, v^\circ, y^\circ, w^\circ)}(\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}) \rightarrow T_{(x^\circ, v^\circ, y^\circ, w^\circ)} \Gamma_1 \times T_{(y_1, w_1)}^{TM} \times \dots \times T_{(y_n, w_n)}^{TM}$$

Nous allons calculer de deux façons différentes le jacobien de dH , par rapport aux formes volumes $\mu_1 = \omega_{\frac{1}{n+1}} \times \dots \times \omega_{\frac{1}{n+1}}$ et $\mu_2 = \omega_1 \times \dots \times \omega$ et la formule (29) résultera de la comparaison des résultats.

Sur les facteurs successifs du produit $\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}$ considérons les coordonnées $(x, v), (y_1, v_1), \dots, (y_n, v_n)$ définies par la projection sur le premier facteur de $TM \times TM$, sur Λ_1 considérons de même les coordonnées x, v et sur chacun des facteurs de $TM \times \dots \times TM$ considérons les coordonnées habituelles, notées $(z_1, u_1), \dots, (z_n, u_n)$.

H s'exprime par des équations de la forme

$$\begin{aligned} x &= x, \\ v &= v, \\ z_1 &= y_1 + \phi_1(x, v), \\ u_1 &= w_1 + \psi_1(x, v), \\ &\dots \\ z_n &= y_n + \phi_n(y_{n-1}, v_{n-1}), \\ u_n &= v_n + \psi_n(y_{n-1}, v_{n-1}). \end{aligned}$$

Le jacobien de H dans les coordonnées utilisées est égal à 1. Comme

$$\begin{aligned} \mu_1 &= \pi(dx)\pi(dv)\pi(dy_1)\pi(dw_1) \dots \pi(dy_n)\pi(dw_n) \text{ et} \\ \mu_2 &= \pi(dx)\pi(dv)\pi(dz_1)\pi(du_1) \dots \pi(dz_n)\pi(du_n) \end{aligned}$$

on en déduit que, calculé par rapport aux formes volume μ_1 et μ_2 en un point de $j(\Gamma_1)$, le jacobien de H à la valeur 1.

Utilisons maintenant sur Γ_1 les coordonnées x, y et de même sur les facteurs successifs du produit $\Gamma_{\frac{1}{n+1}} \times \dots \times \Gamma_{\frac{1}{n+1}}$ les coordonnées $(x, y'_1), (y_1, y'_2), \dots, (y_n, y)$.

L'application H s'écrit $(x, y'_1, y_1, y'_2, \dots, y_n, y) \rightarrow (x, y, z_1, u_1, \dots, z_n, u_n)$ avec

$$\begin{aligned} z_1 &= y_1 - y'_1, \\ u_1 &= \frac{\partial E(x, y'_1)}{\partial y'_1} + \frac{\partial E(y_1, y'_2)}{\partial y_1}, \\ &\dots \\ z_n &= y_n - y'_n, \\ u_n &= \frac{\partial E(y_{n-1}, y'_n)}{\partial y'_n} + \frac{\partial E(y_n, y)}{\partial y_n}. \end{aligned}$$

(la notation $\frac{\partial}{\partial y}$ désigne ici le gradient riemannien). En un point où $y'_1 = y_1, \dots, y'_n = y_n$

$y'_n = y_n$ son jacobien, calculé par rapport aux formes volumes

$$v_1 = \pi(dx)\pi(dy'_1) \dots \pi(dy_n)\pi(dy) \text{ et } v_2 = \pi(dx)\pi(dy)\pi(dz_1)\pi(dy_1) \dots \pi(dz_n)\pi(dy_n)$$

est

$$J = \det_{\pi \times \dots \times \pi} \frac{\partial^2}{\partial y_i \partial y_j} [E(x, y_1) + \dots + E(y_n, y)]$$

Or on sait que sur Γ_1 on a

$$\omega_1 = \pi(dx)\pi(dv) = \theta(x, y)\pi(dx)\pi(dy).$$

De même sur les facteurs successifs de $\Gamma_1 \times \dots \times \Gamma_1$ on a

$$\omega_{\frac{1}{n+1}} = \pi(dx)\pi(dv) = \left(\frac{1}{n+1}\right)^m \theta(x, y_1)\pi(dx)\pi(dy_1),$$

...

$$\omega_{\frac{1}{n+1}} = \pi(dy_n)\pi(dw_n) = \left(\frac{1}{n+1}\right)^m \theta(y_n, y)\pi(dy_n)\pi(dy).$$

On a donc $\mu_1 = (n+1)^{-m(n+1)} \theta(x, y_1) \dots \theta(y_n, y) v_1$ et $\mu_2 = \theta(x, y) v_2$.

Calculé par rapport aux formes μ_1 et μ_2 le jacobien de H en un point de $j(\Gamma_1)$ est égal à

$$(n+1)^{m(n+1)} \theta(x, y_1)^{-1} \dots \theta(y_n, y)^{-1} \theta(x, y) J$$

on sait déjà qu'il vaut 1. La formule (29) en découle.

BIBLIOGRAPHIE

- [1] V.I. Arnold, *Normal forms for functions near degenerate critical points, the Weyl groups of A_k, D_k, E_k and lagrangian singularities*, Functional Analysis and its appl., 6, 1972 p.254-272.
- [2] R. Azencott, *Grandes déviations et applications* in R. Azencott, Y. Guivarc'h, R.F. Gundy, Ecole d'été de Probabilités de Saint-Flour, VII-1978, Lecture Notes in Mathematics n°774, Springer-Verlag, Berlin-Heidelberg-New York, 1980.
- [3] N. Bourbaki, *Variétés différentielles et analytiques*, Fascicule de résultats, §§1-7, Hermann, Paris, 1971.

- [4] M.A. Buchner, *Structure of the cut-locus in dimension less than or equal to 6*, Compositio Math., 37 (1978) p.103-119.
- [5] J. Cheeger, D. Ebin, *Comparison theorems in Riemannian geometry*, North Holland, Amsterdam, 1975.
- [6] S. Kobayashi, K. Nomizu, *Foundations of differential geometry*, Wiley, New-York, 1963.
- [7] S. Lang, *Introduction to differentiable manifolds*, Wiley, New-York, 1962, Introduction aux variétés différentiables, Dunod, Paris, 1967.
- [8] J. Milnor, *Morse theory*, Ann. of Math. Studies, vol 51, Princeton University Press, Princeton, 1963-1969.
- [9] A.D. Ventsel, M.I. Freidlin, *On small random perturbations of dynamical systems*, Russian Math. surveys, 25 (1970) p.1-55.
- [10] C.T.C. Wall, *Geometrical properties of generic differentiable manifolds*, p.707-774 in Geometry and Topology, Rio de Janeiro 1976, Lecture Notes in Mathematics n°597, Springer Verlag, Berlin-Heidelberg-New York, 1977.

CHAPITRE 8

UN PROBLÈME POSÉ PAR LE PASSAGE DES ESTIMÉES LOCALES AUX ESTIMÉES GLOBALES POUR LA DENSITÉ D'UNE DIFFUSION.

par

Robert AZENCOTT

1. Estimées locales et globales de $p(t,x,y)$.

Soit M une variété et p la densité d'une L -diffusion sur M , avec L elliptique, à coefficients suffisamment différentiables. On a vu que (Elie ch. 5) pour tout compact K de U , on a

$$(1) \quad \lim_{t \rightarrow 0} t \log p(t,x,y) = -\frac{1}{2} d^2(x,y)$$

où d est la distance riemannienne associée à D , et ceci pourvu que $d(x,y) \leq \varepsilon(K)$, avec $\varepsilon(K) > 0$.

L'extension de (1) au cas de x,y quelconques, et, a fortiori, celle de l'équivalent de $p(t,x,y)$ obtenu au ch. 5, sont non triviales.

Dans [4] p. 18, Molchanov affirme la possibilité de restreindre l'étude de $\lim_{t \rightarrow 0} p(t,x,y)$ au cas où M est compacte. Son argument, extrêmement condensé, suppose implicitement que M est complète. Dans ce chapitre et le suivant est esquissée une méthode très simple de passage du local au global, qui a l'avantage de s'appliquer au cas où M est non complète.

En effet, il est possible (cf. ci-dessous) de construire une L-diffusion dans un ouvert M de $\mathbb{R}^2$, avec L elliptique et à coefficients C^∞ , telle que la relation (1) soit fausse, c'est-à-dire que

$$\lim_{t \rightarrow 0} t \log p(t, x, y) > -\frac{1}{2} d^2(x, y)$$

pour x, y dans le complémentaire d'un certain compact de $M \times M$. A fortiori l'équivalent global de $p(t, x, y)$ obtenu par Molchanov est faux dans ce cas.

De plus cet exemple montrera que si U est un ouvert relativement compact de M , même "convexe", même à bord très lisse, les densités resp. $p(t, x, y)$ et $p_U(t, x, y)$ des L-diffusions sur M et U ne sont en général pas équivalentes (lorsque $t \rightarrow 0$) pour tout $x, y \in Y$. Il se peut d'ailleurs que même $\log p$ et $\log p_U$ ne soient pas équivalentes.

Comme on le verra au ch. 9 (C. Bellaïche), la relation (1) est sûrement valable si $d(x, y) < \max [d(x, \infty), d(y, \infty)]$, et en particulier lorsque M est complète. Elle reste vraie aussi lorsque M est un ouvert relativement compact d'une variété $\tilde{M}$, pourvu que L se prolonge en un opérateur elliptique à coefficients C^3 sur un voisinage de $\bar{M}$ dans $\tilde{M}$.

Dans les mêmes situations, l'équivalent de $p(t, x, y)$ obtenu au ch. 5 sera vrai pour x, y quelconques, joints par une unique géodésique minimisante.

Au paragraphe 3 de ce chapitre nous donnerons une méthode élémentaire permettant de propager les majorations locales de $p(t, x, y)$ obtenues au ch. 5, pour obtenir des majorations globales du type

$$p(t, x, y) \leq C t^{-N} \exp\left[-\frac{d^2(x, y)}{2t}\right]$$

avec x, y vérifiant $d(x, y) < \max[d(x, \infty), d(y, \infty)]$ et t petit. Ces majorations globales serviront essentiellement au ch. 9, pour étendre la validité de l'équivalent naturel de $p(t, x, y)$ au cas où x, y sont éloignés.

Un intérêt de la méthode proposée, est qu'elle s'appliquera tout aussi bien au cas où le générateur différentiel L est seulement hypo-elliptique, une fois obtenues les majorations locales ; ce dernier problème (local) est non résolu (et délicat) dans le cas hypo-elliptique (cf. 2ième partie du Séminaire).

2. Construction d'un contre-exemple.

2.1. Une diffusion gênante.

Dans le plan considérons une base orthonormée $\vec{OA}, \vec{OC}$, et le carré OABC. Soit M l'intérieur de ce carré. Notons (z_1, z_2) les coordonnées du point courant $z \in M$ dans cette base.

Soit b un champ de vecteurs C^∞ sur M de la forme $b(z) = f(z_2) \vec{OA}$, où la fonction numérique f est C^∞ sur $]0,1[$, décroissante, et tend vers l'infini au voisinage de 0. Plus précisément imposons

$$(2) \quad f(z_2) \equiv n^5 \text{ pour } \frac{1}{n} - \frac{1}{3n^2} \leq z_2 \leq \frac{1}{n} + \frac{1}{3n^2}$$

Soit Δ le laplacien usuel sur $\mathbb{R}^2$ et soit $L = \frac{1}{2}\Delta + b$. L'opérateur L est elliptique de classe C^∞ sur M . La métrique d associée à L sur M (cf. Elie ch. 5) est la métrique euclidienne ordinaire. En particulier M n'est pas complète pour d , et l'opérateur L ne se prolonge pas à un voisinage de $\bar{M}$ dans $\mathbb{R}^2$.

Soit (X_t) la L -diffusion sur M ; soit $p(t, x, y)$ la densité de sa fonction de transition.

2.2. PROPOSITION. *Considérons la diffusion particulière décrite en 2.1. Soit ∞ le point à l'infini du compactifié d'alexandroff de M . Alors pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que les conditions*

$$(3) \quad d(x, \infty) = d(x, OA) \leq \alpha d(x, y)$$

$$d(y, \infty) = d(y, OA) \leq \alpha d(x, y)$$

impliquent pour la densité $p(t, x, y)$

$$(4) \quad \lim_{t \rightarrow 0} t \log p(t, x, y) > -\varepsilon d^2(x, y).$$

En particulier pour $\varepsilon = \frac{1}{2}$, on peut prendre $\alpha = \frac{1}{25}$.

Pour prouver 2.2, nous aurons besoin d'un calcul "brownien" désagréable.

2.3. LEMME. *Soit (β_1, β_2) un brownien plan issu de 0. Pour $r > 0$, $\sigma > 0$ soit $S(r, \sigma)$ l'évènement défini par*

$$(5) \quad |\beta_1(s)| \leq \sqrt{\sigma} \text{ et } -r - \sqrt{\sigma} \leq \beta_2(s) \leq \sqrt{\sigma} \text{ pour } 0 \leq s \leq \sigma$$

$$|\beta_2(\sigma) + r| \leq \frac{1}{2} \sqrt{\sigma}$$

Alors il existe des constantes c_1, c_2 positives, et il existe une fonction $\varepsilon(u)$ tendant vers 0 avec u tels que pour $\frac{r^2}{\sigma} \geq c_2$, on ait

$$P_0[S(r, \sigma)] \geq c_1 \exp \left\{ -\frac{r^2}{2\sigma} \left[1 + \varepsilon\left(\frac{\sigma}{r^2}\right) \right] \right\}$$

Preuve : Remarquons que $P_0[S(r, \sigma)] = P_0[S(\frac{r}{\sqrt{\sigma}}, 1)]$.

Il suffit donc d'étudier

$$P_0[S(r, 1)] = P_0[|\beta_1(s)| \leq 1, 0 \leq s \leq 1] P_0[-r-1 \leq \beta_2(s) \leq 1, 0 \leq s \leq 1; |\beta_2(1)+r| \leq \frac{1}{2}]$$

Le premier facteur est une constante $c_3 > 0$. Le second s'écrit $\int_{-r-\frac{1}{2}}^{-r+\frac{1}{2}} g(w) dw$, où g est la densité au temps 1 du brownien linéaire issu de 0 et tué à la sortie de $[-r-1, 1] = [\lambda, \mu]$. On aura donc

$$(6) \quad P_0[S(r, 1)] \geq c_3 \inf_{-r-\frac{1}{2} \leq w \leq -r+\frac{1}{2}} g(w).$$

La méthode des images (cf. Dynkin [2] par exemple) donne

$$\sqrt{2\pi} g(w) = \exp\left(-\frac{w^2}{2}\right) - \exp\left[-\frac{(2\lambda-w)^2}{2}\right] - \left\{ \exp\left[-\frac{(2\mu-w)^2}{2}\right] - \exp\left[-\frac{(2\lambda-2\mu+w)^2}{2}\right] \right\}$$

La somme des deux premiers termes s'écrit

$$\exp\left(-\frac{w^2}{2}\right) [1 - \exp 2\lambda(w-\lambda)]$$

et est donc minorée par $\frac{1}{2} \exp\left(-\frac{w^2}{2}\right)$ pour r assez grand et $w \in [-r-\frac{1}{2}, -r+\frac{1}{2}]$. Sous les mêmes conditions le terme $\{...\}$ dans $\sqrt{2\pi} g(w)$ se majore de même par $\frac{1}{2} \exp\left[-\frac{(2\mu-w)^2}{2}\right]$, d'où $\sqrt{2\pi} g(w) \geq \frac{1}{2} \exp\left(-\frac{w^2}{2}\right) - \frac{1}{2} \exp\left[-\frac{(2\mu-w)^2}{2}\right]$.

Le même calcul montre que pour r grand, et $w \in [-r-\frac{1}{2}, -r+\frac{1}{2}]$ on a

$$(7) \quad \sqrt{2\pi} g(w) \geq \frac{1}{4} \exp\left(-\frac{w^2}{2}\right) \geq \frac{1}{8} \exp\left(-\frac{r^2}{2}(1+\varepsilon)\right)$$

où ε tend vers 0 avec $\frac{1}{r}$.

De (6), (7) on conclut que $P_0[S(r, l)]$ est minoré par $c_4 \exp[-\frac{r^2}{2}(1+\varepsilon)]$ pour r assez grand, ce qui prouve le lemme.

2.4. Première étape : définition d'un paquet de chemins de x à y :

Soient $x, y \in M$, vérifiant (3) avec α "petit" à fixer plus bas. Supposons que $d(x, \infty) = d(y, \infty) = r$, et posons $d(x, y) = h$.

Pour chaque t , choisissons $n =$ partie entière de $(\frac{h}{t})^{1/5}$ d'où $n \sim (\frac{h}{t})^{1/5}$, et considérons les points x^n, y^n définis par

$$(8) \quad x_1^n = x_1, \quad x_2^n = \frac{1}{n}, \quad y_1^n = y_1, \quad y_2^n = \frac{1}{n}.$$

Pour $u, v \in \mathbb{R}^2$ appelons "boite d'axe uv", l'ensemble des points de la forme

$$\xi u + (1-\xi)v + \eta\sqrt{t} \overrightarrow{OA} + \zeta \sqrt{t} \overrightarrow{OC}$$

où $0 \leq \xi \leq 1, |\eta| \leq 1, |\zeta| \leq 1$.

Fixons $q \in]0, \frac{1}{2}[$ et soit V_x le paquet de trajectoires issues de x , définies sur $[0, qt]$, qui ne quittent pas la "boite d'axe xx^n ", et qui au temps qt se trouvent à une distance de x^n inférieure à $\sqrt{t}$.

Pour $u \in M$, posons $\hat{u} = u + y - x$; soit H_u le paquet de trajectoires issu de u au temps qt , définies sur $[qt, (1-q)t]$, qui ne quittent pas la "boite d'axe $u\hat{u}$ ", et qui au temps $(1-q)t$ se trouvent à une distance de $\hat{u}$ inférieure à $\sqrt{t}$.

Enfin pour $v \in M$, posons $\tilde{v} = v + y - y^n$. Soit W_v le paquet de trajectoires issues de v au temps $(1-q)t$ définies sur $[(1-q)t, t]$, qui ne quittent pas la "boite d'axe $v\tilde{v}$ ", et qui au temps t se trouvent à une distance de $\tilde{v}$ inférieure à $\sqrt{t}$.

Il est clair que l'évènement

$$(9) \quad \mathfrak{X} = \{X_{0,qt} \in V_x ; X_{qt,t-qt} \in H_{X_{qt}} ; X_{t-qt,t} \in W_{X_{t-qt}}\}$$

implique $d(y, X_t) \leq 3\sqrt{t}$. Par suite, la loi Q_x des trajectoires de la L-diffusion issue de x vérifie

$$(10) \quad Q_x[d(y, X_t) \leq 3\sqrt{t}] \geq Q_x(\mathfrak{X}).$$

Pour minorer $Q_x(\mathfrak{X})$, nous allons appliquer la propriété de Markov aux temps qt et $(1-q)t$, et minorer successivement chaque "facteur".

2.5. Deuxième étape ; minoration de $Q_x[X_{0,qt} \in V_x]$:

Soit $\beta(t)$ le brownien dans $\mathbb{R}^2$; notons P_x la loi du Brownien issu de x , et E_x l'espérance par rapport à P_x . Soit τ_M le temps de sortie de M .

La formule de Cameron-Martin s'écrit

$$(11) \quad Q_x[X_{0,qt} \in V_x] = E_x[1_{\{t < \tau_M\}} 1_{\{\beta_{0,qt} \in V_x\}} \exp J_{qt}]$$

où

$$(12) \quad J_{qt} = \left\{ \int_0^{qt} b[\beta(s)] d\beta(s) - \frac{1}{2} \int_0^{qt} |b[\beta(s)]|^2 ds \right\} 1_{\{\beta_{0,qt} \in V_x\}}$$

Puisque $x_2^n = \frac{1}{n} \sim (\frac{t}{h})^{1/5} \gg \sqrt{t}$, la condition $\beta_{0,qt} \in V_x$ entraîne $t < \tau_M$. Par définition de b , cf. 2.1, on a

$$J_{qt} = \left\{ \int_0^{qt} f[\beta_2(s)] d\beta_1(s) - \frac{1}{2} \int_0^{qt} f^2(\beta_2(s)) ds \right\} 1_{\{\beta_{0,qt} \in V_x\}}.$$

Le processus croissant associé à $\int f(\beta_2) d\beta_1$ est $\int f^2(\beta_2) ds$ et pour $\beta_{0,qt} \in V_x$, sa valeur en qt est majorée par

$$\phi(t) = qt f^2(x_2^n - \sqrt{t}) = qt \left(\frac{1}{3} \right)^2 = q \frac{h^2}{t} [1 + \varepsilon_1(t)]$$

où $\varepsilon_1(t) \rightarrow 0$ avec t .

La majoration exponentielle classique des martingales (cf. Stroock-Varadhan [4]) donne

$$P_x \left[\int_0^{qt} f(\beta_2) d\beta_1 \leq -a \right] \leq \exp \left(- \frac{a^2}{2\phi(t)} \right)$$

Des deux remarques ci-dessus, on conclut que

$$(13) \quad P_x[J_{qt} \leq -a - \frac{1}{2} \phi(t)] \leq \exp \left[- \frac{a^2}{2\phi(t)} \right]$$

La formule (11) donne

$$Q_x[X_{0,qt} \in V_x] \geq \int_{\{\beta_{0,qt} \in V_x ; J_{qt} \geq -a - \frac{1}{2} \phi(t)\}} \exp J_{qt} dP_x$$

d'où par (13)

$$(14) \quad Q_x[X_{0,qt} \in V_x] \geq \exp \left[-a - \frac{1}{2} \phi(t) \right] \{ P_x[\beta_{0,qt} \in V_x] - \exp \left[- \frac{a^2}{2\phi(t)} \right] \}$$

La définition de $S(r, \sigma)$ (cf. (5)) donne immédiatement, en tenant compte du lemme 2.3

$$P_x[\beta_{0,qt} \in V_x] \geq P_o[S(r, \sqrt{qt})] \geq c_1 \exp - \frac{r^2}{2qt}(1+\varepsilon_2(t))$$

pour t assez petit, où $\varepsilon_2(t) \rightarrow 0$ avec t .

Choisissons

$$a = \sqrt{2} \frac{rh}{t} \sqrt{1+\varepsilon_1} = \sqrt{2} \frac{rh}{t}(1+\varepsilon_3) \leq \frac{2rh}{t}$$

d'où $\frac{a^2}{2\phi(t)} = \frac{r^2}{qt}$; dans (14) le terme $\{...\}$ se minore donc par $c_2 \exp(-\frac{r^2}{2qt}[1+\varepsilon_2])$ pour t petit. Par suite (14) entraîne

$$(15) \quad Q_x(X_{0,qt} \in V_x) \geq c_2 \exp \left[-\frac{2rh}{t} - \frac{qh^2}{2t}(1+\varepsilon_1) - \frac{r^2}{2qt}(1+\varepsilon_2) \right]$$

Un calcul identique au précédent, à translation près dans le temps, et la propriété de Markov, montrent que pour t assez petit, on a Q_x -p.s., sur l'ensemble

$$\{d(X_{t-qt}, y^n) \leq 2\sqrt{t}\}, \text{ la relation}$$

$$(16) \quad Q[X_{t-qt,t} \in W_{X_{t-qt}} \mid \mathcal{F}_{t-qt}] \geq c_2 \exp \left[-\frac{2rh}{t} - \frac{qh^2}{2t}(1+\varepsilon_1) - \frac{r^2}{2qt}(1+\varepsilon_2) \right]$$

2.6. Troisième étape ; minoration de $Q(X_{qt,t-qt} \in H_{X_{qt}} \mid \mathcal{F}_{qt})$:

Cette minoration nous intéresse sur l'ensemble $\{d(X_{qt}, x^n) \leq \sqrt{t}\}$. Par la propriété de Markov et translation temporelle, ceci revient à évaluer $Q_z(X_{0,t-2qt} \in K_z)$ avec $d(z, x^n) \leq \sqrt{t}$, où le paquet de trajectoires K_z est défini exactement comme H_z , au remplacement près de l'intervalle $[qt, t-qt]$ par $[0, t-2qt]$.

Comme $\frac{1}{3n^2} \sim \frac{1}{3}(\frac{t}{h})^{2/5} \gg 2\sqrt{t}$, la définition du champ b montre que l'évènement

$\{\beta_{0,t-2qt} \in K_z\}$ entraîne $b(\beta_s) \equiv n^5 \overrightarrow{OA} = \frac{h}{t}[1+\varepsilon_4(t)] \overrightarrow{OA}$ pour $0 \leq s \leq t-2qt$ où $\varepsilon_4(t) \rightarrow 0$ avec t .

Dans la formule de Cameron Martin analogue à (11) le $\log$ de $\frac{dQ_z}{dP_z}$ s'écrit alors immédiatement

$$J_{t-2qt} = \frac{h}{t}(1+\varepsilon_4)[\beta_1(t-2qt)-z_1] - \frac{1}{2} \frac{h^2}{t}(1-2q)(1+\varepsilon_4)^2.$$

Par définition de K_z , on a $d(\beta_{t-2qt}, \hat{z}) \leq \sqrt{t}$ où $\hat{z} = z + h \overrightarrow{OA}$, ce qui entraîne directement $J_{t-2qt} \geq \frac{h^2}{2t}(1+2q+\varepsilon_5)$, où $\varepsilon_5(t) \rightarrow 0$ avec t .

La formule de Cameron-Martin fournit alors la minoration

$$(17) \quad Q_z(X_{0,t-2qt} \in K_z) \geq \exp \left[\frac{h^2}{2t}(1+2q+\epsilon_5) \right] P_z[\beta_{0,t-2qt} \in K_z].$$

Les définitions de K_z et $S(h,\sigma)$ montrent que

$$(18) \quad P_z[\beta_{0,t-2qt} \in K_z] \geq P_o[S(h,t-2qt)].$$

De (17) (18) et du lemme 2.3, on conclut que sur l'ensemble $\{d(X_{qt}, x^n) \leq \sqrt{t}\}$, on a Q_x -p.s.

$$(19) \quad Q[X_{qt,t-qt} \in H_{X_{qt}} \mid \mathfrak{F}_{qt}] \geq c_1 \exp \left[\frac{h^2}{2t}(1+2q+\epsilon_5) - \frac{h^2}{2(1-2q)t}(1+\epsilon_6) \right]$$

où $\epsilon_6(t) \rightarrow 0$ avec t .

2.7. Quatrième étape ; minoration de $Q_x[d(X_t, y) \leq 3\sqrt{t}]$:

La propriété de markov appliquée deux fois et les minoration (19) (16) et (15) fournissent une minoration immédiate de $Q_x(\mathfrak{Z})$ par le produit des seconds membres de (19) (16) (15). Grâce à (10) on obtient donc

$$(20) \quad Q_x[d(X_{t,y}) \leq 3\sqrt{t}] \geq c_3 \exp \left\{ + \frac{1}{2t} [B - \epsilon_7(t)] \right\}$$

où $\epsilon_7(t) \rightarrow 0$ avec t , avec

$$B = -8rh + h^2 - \frac{2r^2}{q} - \frac{h^2}{1-2q}$$

Pour maximiser B on fixe $q = \frac{r}{h+2r}$, d'où $\max_q B = -4r^2 - 12rh$, ce qui par rapport dans (20) donne

$$(21) \quad Q_x[d(X_{t,y}) \leq 3\sqrt{t}] \geq c_3 \exp - \frac{1}{2t} [4r^2 + 12rh + \epsilon_7(t)]$$

dès que t est assez petit.

2.8. Cinquième étape ; minoration de $p(t,x,z)$:

Soient maintenant deux points x et z de M , tels que $d(x,z) = h$ situés à des distances respectives r et ρ de OA , avec $\frac{r}{h}, \frac{\rho}{h}$ petits, $r < \rho$ et $d(x,\infty) = r$, $d(y,\infty) = \rho$. Soit y le point situé à la verticale de z et à la distance r de OA . On a $d(x,y) = \lambda \leq h$.

Soit V une boule compacte de centre y et de rayon $\gamma > 0$. Pour $u \in V$ et γ petit, on a $d(u, z) < \rho = d(z, \infty)$. L'équivalent naturel de la densité $p(t, u, z)$ pour la L -diffusion est valide dans cette situation, uniformément en $u \in V$ (cf. C. Bellaïche) ch. 9). Par suite on aura pour t assez petit

$$(22) \quad p(\frac{t}{2}, u, z) \geq c_4 \exp - \frac{d^2(u, z)}{t} \text{ pour tout } u \in V.$$

Soit W une boule de centre y et de rayon $3\sqrt{\frac{t}{2}}$. Pour t assez petit, (22) montre que

$$(23) \quad \inf_{u \in W} p(\frac{t}{2}, u, z) \geq c_4 \exp \left[- \frac{(\rho-r)^2 + 8}{t} \right]$$

où $\varepsilon_8(t) \rightarrow 0$ avec t , et $c_4 > 0$.

La formule de Chapman-Kolmogorov fournit

$$(24) \quad p(t, x, z) \geq \int_W p(\frac{t}{2}, x, u) p(\frac{t}{2}, u, z) du \geq \left[\inf_{u \in W} p(\frac{t}{2}, u, z) \right] Q_x(X_{\frac{t}{2}} \in W).$$

Les minoration (24) (23) et (21) donnent alors

$$(25) \quad p(t, x, z) \geq c_5 \exp \left[- \frac{\dot{I} + \varepsilon_9(t)}{2t} \right] \text{ où } \varepsilon_9(t) \rightarrow 0 \text{ avec } t, \text{ avec}$$

$$I = 8r^2 + 24rh + 2(\rho-r)^2$$

ce qui entraîne

$$(26) \quad \lim_{t \rightarrow 0} t \log p(t, x, z) \geq - \frac{1}{2} [8r^2 + 24rh + 2(\rho-r)^2] = - \frac{I}{2}$$

Il est évident que $\frac{I}{h^2}$ tend vers 0 avec $(\frac{r}{h} + \frac{\rho}{h})$, ce qui prouve la prop. 2.2.

3. Deux contre-exemples dérivés du précédent :

3.1. PROPOSITION. *Considérons la L -diffusion construite en 2.1. Il existe une infinité d'ouverts U de M , relativement compacts dans M , à bord lisse, tels que si p_U est la densité de la L -diffusion sur U on ait, pour une infinité de couples $x, y \in U$*

$$\lim_{t \rightarrow 0} t \log p(t, x, y) > \lim_{t \rightarrow 0} t \log p_U(t, x, y)$$

En particulier, p et p_U ne sont pas équivalentes quand $t \rightarrow 0$

Preuve : Soit F le centre du carré M . Prenons l'image de M par une homothétie (F, λ) avec $\lambda < 1$, λ proche de 1. Lissons de façon quelconque les coins de ce nouveau carré et soit U l'intérieur du domaine ainsi défini. Si λ est assez proche de 1, U contient des points x, y tels que

$$d(x, \infty) = d(x, OA) \leq \frac{1}{25} d(x, y)$$

$$d(y, \infty) = d(y, OA) \leq \frac{1}{25} d(x, y)$$

et $d(x, y) = d_U(x, y)$.

On a donc $\lim_{t \rightarrow 0} t \log p(t, x, y) > -\frac{1}{2} d^2(x, y)$, par la prop. 2.2. D'autre part, comme l'opérateur D se prolonge en un opérateur elliptique de classe C^∞ sur un voisinage de $\bar{U}$, on a (cf. C. Bellaïche ch. 9)

$$\lim_{t \rightarrow 0} t \log p_U(t, x, y) = -\frac{1}{2} d_U^2(x, y)$$

ce qui prouve le résultat annoncé.

3.2. Commentaire : Sur une variété complète M (pour la distance d associée à D), on a en fait $p(t, x, y) \sim p_U(t, x, y)$ pour tout $x, y \in U$ et tout ouvert relativement compact U de M pourvu que U contienne une géodésique minimisante (dans M) allant de x à y . Ceci résulte du calcul effectif des équivalents de p et de p_U , à la Molchanov (cf. C. Bellaïche ch. 9).

Lorsque M n'est pas complète, le même résultat reste vrai (pour les mêmes raisons) pourvu que $d(x, y) < \sup[d(x, \infty), d(y, \infty)]$.

3.3. PROPOSITION. *Considérons la variété M , le laplacien Δ et le champ b sur M définis en 2.1. Soit Λ la fonctionnelle de Cramer associée au problème de petites perturbations $L_\varepsilon = \frac{1}{2} \varepsilon^2 \Delta + \varepsilon b$. (Renvoyons à Azencott ch 12, ainsi qu'à [1]). Il existe un ouvert A de l'ensemble $C(M)$ des chemins continus sur $[0, 1]$ à valeurs dans M , et un point $x \in M$ tels que*

$$-\Lambda(\bar{A}) < \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log R_x(Y_\varepsilon \in A)$$

où Y_ε est la trajectoire (sur $[0, 1]$) de la D_ε -diffusion et où R_x est la loi des trajectoires issues de x . Ici $\bar{A}$ est l'adhérence de A dans $C(M)$ muni de la topologie de la convergence uniforme.

3.4. Remarque : Tout le sel du "contre-exemple" 3.3, provient de ce que si on considère A comme sous ensemble de l'espace $\mathcal{E}(M)$ des chemins "explosifs" définis sur $[0,1]$ à valeurs dans $M \cup \delta$, on a en fait toujours

$$\overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon^2 \log R_x(Y_\varepsilon \in A) \leq -\Lambda(\tilde{A})$$

où $\tilde{A}$ est l'adhérence de A dans $\mathcal{E}(M)$ muni d'une topologie adéquate naturelle, (cf. Azencott [1]). La prop. 3.3 montre l'intérêt de l'introduction de cette topologie.

3.5. Preuve de 3.3.

Soient $x, y \in M$ comme en 2.4 avec $r = d(x, OA) = d(y, OA)$ et $h = d(x, y)$. Prenons

$$A = \{\gamma \in C(M) \mid \gamma_0 = x, \gamma_1 \in V\}$$

où V est une boule de centre $y \in M$, de rayon $\rho > 0$. Dans le cas simple qui nous intéresse $\Lambda(\bar{A}) = \inf_{\gamma \in \bar{A}} \frac{1}{2} \lambda(\gamma)$ où $\lambda(\gamma)$ est le carré de la longueur euclidienne de γ .

Ceci donne donc $\Lambda(\bar{A}) = \frac{1}{2} d^2(x, V) = \frac{1}{2} [d(x, y) - \varepsilon]^2$.

Un simple changement de temps (cf. Azencott ch. 12) montre que

$$R_x(Y_\varepsilon \in A) = Q_x(X_t \in V) \text{ si } \varepsilon = \sqrt{t}.$$

Mais la formule (21) donne

$$\lim_{t \rightarrow 0} t \log Q_x(X_t \in V) \geq -\frac{1}{2}(4r^2 + 12rh)$$

et pour r/h assez petit on aura

$$\frac{1}{2}(4r^2 + 12rh) < \frac{1}{2}(h-\varepsilon)^2 = \Lambda(\bar{A})$$

d'où le résultat annoncé en 3.3.

3.4. Des diffusions gênantes génériques :

Soit M une variété C^∞ quelconque. Soit U un ouvert relativement compact de M , tel que pour tout $x \in \partial U$ il existe une carte (V, ϕ) autour de x telle que $\phi(V \cap U) = H \cap \phi(V)$ où H est un $\frac{1}{2}$ -espace euclidien ouvert.

Soit $\frac{1}{2}\Delta$ un opérateur elliptique d'ordre 2, de classe C^∞ sur M . Soit d la distance riemannienne sur M associée à $\frac{1}{2}\Delta$ (cf. Elie ch. 5). Supposons l'existence d'un point x_0 de ∂U et d'un voisinage V de x_0 tel que pour tout $x, y \in V \cap U$, la géodésique minimisante joignant x à y soit unique et contenue dans $V \cap U$.

Conjecture : Il existe un champ b , C^∞ sur U , avec $|b(x)| \rightarrow +\infty$ quand $x \rightarrow x_0$, tel que la $(\frac{1}{2}\Delta + b)$ -diffusion sur U ait une densité $p(t, x, y)$ vérifiant

$$\lim_{t \rightarrow 0} t \log p(t, x, y) > -\frac{1}{2} d^2(x, y)$$

pour une infinité de points $x, y \in U \cap V$.

Nous n'avons pas mené cette construction jusqu'au bout, mais le paragraphe 2 en donne évidemment le principe : Pour $x, y \in U$ très proche de x_0 , avec $d(x, \partial U)$ et $d(y, \partial U)$ négligeables devant $d(x, y)$, on considère un paquet de chemin allant de x "à ∂U ", puis suivant "sur ∂U " une "trajectoire du champ fb ", où b est un champ tangent à ∂U en tout point de $V \cap \partial U$ et où $f : U \rightarrow \mathbb{R}$ tend vers l'infini au bord de U , puis enfin revenant "de U " à y .

La construction se fait pratiquement, sur une sous-variété de dimension 2 de U , celle des géodésiques $x^n y^n$ utilisées au paragraphe 2. La preuve présente quelques difficultés techniques supplémentaires (remplacement du brownien plan par la diffusion associée à $\frac{1}{2}\Delta$) tout à fait surmontables.

4. Propagation des majorations locales de $p(t, x, y)$ par itération :

4.1. Utilité des majorations de $p(t, x, y)$ pour x, y éloignés :

Soit $p(t, x, y)$ la densité d'une diffusion sur une variété. Pour propager l'équivalent local de $p(t, x, y)$ au cas de x, y éloignés, il est naturel (Molchanov [3]) d'utiliser la formule de Chapman-Kolmogorov. Mais si M n'est pas compacte, cas que Molchanov néglige (à tort) d'étudier, il est essentiel, dans une intégrale du type

$$\int_M p(\frac{t}{k}, x, z_1) \dots p(\frac{t}{k}, z_k, y) dz_1 \dots dz_k$$

de pouvoir négliger la contribution des termes correspondant à $d(z_i, z_j) > \varepsilon$, avec ε petit. Ceci exige d'avoir obtenu une majoration de $p(t, x, y)$ pour x, y éloignés.

Nous allons donc, par itération, à l'aide d'ailleurs de la formule de Chapman-Kolmogorov, commencer par propager du local au global une majoration de $p(t, x, y)$ du type $ct^{-N} \exp[-\frac{d^2(x, y)}{2t}]$, ce qui servira ensuite (C. Bellaïche, ch 9) à obtenir l'équivalent de $p(t, x, y)$ pour x, y éloignés.

4.2. Hypothèses et notations :

Soit M une variété de dimension n , de classe 5. Soit $\mathcal{D}$ une diffusion sur M de générateur différentiel L .

Nous supposons L elliptique, à coefficients C^4 , vérifiant $L1 = 0$. Sur M on

note d la distance riemannienne associée à L (cf. Elie ch. 5). Pour $x \in M$ on définit $d(x, \infty)$, où ∞ est le point à l'infini du compactifié d'Alexandroff de M , par

$$\begin{aligned} d(x, \infty) &= \lim_{y \rightarrow \infty} d(x, y) && \text{si } M \text{ est non compacte} \\ d(x, \infty) &= +\infty && \text{si } M \text{ est compacte.} \end{aligned}$$

Remarquons que la fonction $x \rightarrow d(x, \infty)$ est soit identique à $+\infty$ (cas d'une distance complète), soit partout finie et continue.

On fixe sur M une autre structure riemannienne arbitraire mais de classe 3, et on note π le volume riemannien correspondant. Enfin on note p la densité (par rapport à π) du semi-groupe de transition (R_t) de $\mathcal{P}$.

4.3. Remarque : En fait toutes les démonstrations du paragraphe 4 n'utiliserons que les faits suivants :

1) $\mathcal{P}$ est un processus de Markov fort à trajectoires dans $\mathcal{D}(M)$ (cf. Azencott ch. 12), dont le semi-groupe admet par rapport à π une densité $p(t, x, y)$ continue en (t, x, y) .

2) p vérifie l'équation de Chapman-Kolmogorov.

3) p vérifie les majorations locales en $\exp[-\frac{d^2(x, y)}{2t}]$, où la fonction d satisfait à l'inégalité triangulaire (cf. 4.5 pour une formulation précise des "majorations locales").

4) $d(x, \infty)$ est soit identique à $+\infty$, soit finie et continue.

Cette remarque est à verser au dossier du cas hypoelliptique.

4.4. PROPOSITION. *Soit $\mathcal{P}$ une diffusion sur une variété M de classe 5. Supposons que le générateur différentiel L de $\mathcal{P}$ soit elliptique, à coefficients C^4 . Soit d la distance associée à L . Alors pour tout compact K de M , il existe des nombre c, T, N positifs stricts tels que*

$$\begin{aligned} p(t, x, y) &\leq c t^{-N} \exp[-\frac{d^2(x, y)}{2t}] \text{ sur l'ensemble} \\ \{0 < t \leq T, x \in K, y \in K, d(x, y) < \sup[d(x, \infty), d(y, \infty)]\} \end{aligned}$$

De plus pour toute fonction continue v sur M , vérifiant $0 < v(y) < d(y, \infty)$, et tout compact K de M , il existe c, T, N positifs stricts tels que sur l'ensemble

$$\{0 < t \leq T, x \in M, y \in K, d(x, y) > v(y)\},$$

les nombres $p(t, x, y)$ et $p(t, y, x)$ soient majorés par $c t^{-N} \exp[-\frac{v^2(y)}{2t}]$

Preuve : Elle est fragmentée en plusieurs sections ci-dessous.

4.5. Deux majorations intermédiaires :

A chaque partie K de M , et chaque $r > 0$, $s > 0$ associons les propriétés suivantes

$H(K, r, s)$: Il existe des nombres c, N, T positifs stricts tels que pour

$0 < t \leq T$, $x \in M$, $y \in K$, $d(x, y) \leq r \wedge sd(y, \infty)$ on ait

$$p(t, x, y) \leq ct^{-N} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

$U(K, r, s)$: Il existe c, N, T positifs stricts tels que pour

$0 < t \leq T$, $x \in M$, $y \in K$, $d(x, y) \geq r \wedge sd(y, \infty)$ on ait

$$p(t, x, y) \leq ct^{-N} \exp \left[-\frac{r^2 \wedge s^2 d^2(y, \infty)}{2t} \right]$$

L'étude de $p(t, x, y)$ quand $t \rightarrow 0$, avec x, y proches fixés (cf. L. Elie ch. 5, Molchanov [3]) montre que pour tout compact K de M , tout $s > 0$, il existe $r > 0$ tel que $H(K, r, s)$ soit vraie. Il s'agit d'améliorer la taille de r par itération. Pour chaque $r > 0$ $s > 0$, nous noterons $K^{r, s}$ l'ensemble des $x \in M$ tels qu'il existe $y \in K$ vérifiant $d(x, y) \leq r \wedge sd(y, \infty)$.

4.6. LEMME. (Hypothèses et notations de 4.4, 4.5). Pour tout $r, s > 0$ avec $r > 0$, $1 \geq s > 0$, toute partie K de M , et tout ouvert V de M inclus dans K , la validité de $H(K, r, s)$ entraîne celle de $U(\bar{V}, r, s)$.

Preuve : Soient r, K, V comme ci-dessus. Posons $v = r \wedge sd(y, \infty)$. Soit $y \in V$. Soit $\varepsilon > 0$ quelconque, et soit B la boule de centre y et de rayon $(v - \varepsilon)$. Soit A un voisinage compact de y contenu dans K et dans la boule de centre y , de rayon $\frac{\varepsilon}{2(1+s)}$. Posons pour $z \in \partial B$, $0 \leq u \leq t$

$$(27) \quad \Phi_t(u, z) = E_z[1_A(X_{t-u})] = R_{t-u}^{-1} 1_A(z)$$

Soit $x \in M$ tel que $d(x, y) \geq v$, donc vérifiant $x \notin B$. Notons η le temps d'entrée de $\mathcal{P}$ dans B ; la propriété de Markov forte entraîne

$$\begin{aligned} (28) \quad R_t^{-1} 1_A(x) &= E_x[1_A(X_t)] = E_x[1_{\{\eta < t\}} E[1_A(X_t) | \mathcal{F}_\eta]] \\ &= E_x[1_{\{\eta < t\}} \Phi_t(\eta, X_\eta)] \end{aligned}$$

Pour $z \in \partial B$, $w \in A$, on a par construction

$$d(z, w) \leq [r \wedge sd(y, \infty)] - \varepsilon + \frac{\varepsilon}{2(1+s)} \leq -\varepsilon + \frac{1+2s}{2(1+s)} + r \wedge [\frac{s\varepsilon}{2(1+s)} + sd(w, \infty)]$$

d'où $d(z, w) \leq r \wedge sd(w, \infty)$.

Comme de plus $A \subset K$, la validité de $H(K, r, s)$ montre que pour $z \in \partial B$, $w \in A$, on a

$$(29) \quad p(t, z, w) \leq c t^{-N} \exp \left[-\frac{d^2(z, w)}{2t} \right] \text{ pour } 0 < t \leq T$$

Par construction les conditions $z \in \partial B$, $w \in A$ entraînent $d(z, w) \geq v - 2\varepsilon \geq \frac{1}{2}v$ pourvu que $\varepsilon \leq \frac{v}{4}$. Le second membre de (29) est alors fonction croissante de t sur $]0, T]$, si on impose $T \leq \inf_{y \in K} \frac{1}{4N} v^2(y)$, $\varepsilon \leq \inf_{y \in K} \frac{1}{4} v(y)$. On conclut ainsi

$$p(t, z, w) \leq c t^{-N} \exp \left(-\frac{(v-2\varepsilon)^2}{2t} \right) \text{ pour } 0 < t \leq T, z \in \partial B, w \in A$$

D'après (27) ceci donne

$$(30) \quad \Phi_t(u, z) \leq c t^{-N} \exp \left[-\frac{(v-2\varepsilon)^2}{2t} \right] \pi(A) \text{ pour } 0 < t \leq T, z \in \partial B$$

ce qui, avec (28) implique

$$\frac{1}{\pi(A)} \int_A p(t, x, w) d\pi(w) \leq c t^{-N} \exp \left[-\frac{(v-2\varepsilon)^2}{2t} \right]$$

Lorsque $\varepsilon \rightarrow 0$, le diamètre de A tend vers 0, et p étant continue on obtient

$$p(t, x, y) \leq c t^{-N} \exp \left(-\frac{v^2}{2t} \right)$$

pour $0 < t \leq T$, $x \in M$, $d(x, y) \geq v(y)$, $y \in V$, inégalité qui reste vraie pour $y \in \bar{V}$ par continuité. Ainsi $U(\bar{V}, r, s)$ est vraie.

4.7. LEMME. (Hypothèses et notations de 4.4, 4.5). Pour tout $r > 0$, $u > 0$, $1 > s > 0$, toute partie K de M et tout ouvert V de M inclus dans K , la validité simultanée de $H(K, r, s)$ et $H(K^{r,s}, u, s)$ entraîne celle de $H[\bar{V}, \sqrt{2}(r \wedge u), s]$.

Preuve : Supposons provisoirement $\pi(M)$ fini. La propriété de semi-groupe $R_{\frac{t}{2}} R_{\frac{t}{2}} = R_t$ donne

$$(31) \quad p(t, x, y) = \int_M p\left(\frac{t}{2}, x, z\right) p\left(\frac{t}{2}, z, y\right) d\pi(z)$$

pour tout $t > 0$, $x, y \in M$ (pour l'égalité partout en $y \in M$, cf. Azencott ch. 2 §.5). Pour $A \subset M$ notons I_A l'intégrale en z de $p(\frac{t}{2}, x, z) p(\frac{t}{2}, z, y)$ sur A .

Posons $v(y) = r \wedge sd(y, \infty)$, $w(z) = u \wedge sd(z, \infty)$, et

$$M_y = \{z \in M \mid d(z, y) > v(y)\}$$

$$M_{xy} = \{z \in M \mid d(z, y) \leq v(y), d(x, z) > w(z)\}$$

$$N_{x,y} = \{z \in M \mid d(z, y) \leq v(y), d(x, z) \leq w(z)\}$$

Imposons $y \in V$, avec V ouvert de K .

D'après le lemme 4.6, la validité de $H(K, r, s)$ entraîne celle de $U(\bar{V}, r, s)$, et donc pour $z \in M_y$, $0 < t \leq T$, la fonction $p(\frac{t}{2}, z, y)$ est majorée par $2^N c t^{-N} \exp [-\frac{2v^2(y)}{2t}]$ tandis que $\int_M p(\frac{t}{2}, x, z) d\pi(z) \leq 1$. Ceci entraîne

$$(32) \quad I_{M_y} \leq 2^N c t^{-N} \exp [-\frac{2v^2(y)}{2t}] \quad 0 < t \leq T, y \in K.$$

Pour $z \in M_{xy}$ on a $z \in V^{r,s}$. La validité de $H(K^{r,s}, u)$ entraîne celle de $U(V^{r,s}, u)$; il existe donc $c_1 N_1 T_1$ tels que

$$p(\frac{t}{2}, x, z) \leq 2^{N_1} c_1 t^{-N_1} \exp [-\frac{2w^2(z)}{2t}], \quad 0 < t \leq T_1, z \in M_{xy}.$$

La validité de $H(K, r, s)$ permet de majorer $p(\frac{t}{2}, z, y)$ par

$$2^{N_2} c_2 t^{-N_2} \exp [-\frac{2d^2(z, y)}{2t}] \quad \text{pour } 0 < t \leq T_2, z \in M_{xy}, y \in K.$$

Comme $\pi(M)$ est fini par hypothèse, on conclut que

$$(33) \quad I_{M_{xy}} \leq c_3 t^{-N_3} \sup_{z \in M_{xy}} \exp [-\frac{2w^2(z) + 2d^2(z, y)}{2t}] \quad \text{pour } y \in V, 0 < t \leq T_2.$$

Il est clair que

$$2w^2(z) + 2d^2(z, y) \geq [w(z) + d(z, y)]^2 = [u \wedge sd(z, \infty) + d(z, y)]^2$$

et que le dernier terme écrit est minoré par $[u \wedge sd(y, \infty)]^2$, puisque $0 < s < 1$.

Si $u > sd(z, \infty)$, on a

$w(z) = sd(z, \infty)$ et donc, puisque $s < 1$

$$\begin{aligned} 2w^2(z) + 2d^2(z, y) &\geq s^2 [2d^2(z, \infty) + 2d^2(z, y)] \geq s^2 [d(z, \infty) + d(z, y)]^2 \\ &\geq s^2 d^2(y, \infty). \end{aligned}$$

Si $u < sd(z, \infty)$, on a $w(z) = u$, donc $2w^2(z) + 2d^2(z, y) \geq 2u^2$. Finalement on voit que

$$\inf_{z \in M_{xy}} [2w^2(z) + 2d^2(z, y)] \geq (2u^2) \wedge s^2 d^2(y, \infty)$$

et (33) implique, pour $0 < t \leq T_3$, $y \in V$

$$(34) \quad I_{M_{xy}} \leq c_3 t^{-N_3} \exp \left[- \frac{(2u^2) \wedge s^2 d(y, \infty)}{2t} \right].$$

Enfin pour $z \in N_{x,y}$ la validité de $H(K, r, s)$ et de $H(K^{r,s}, u, s)$ donne pour $0 < t \leq T_4$, $y \in K$,

$$p(\frac{t}{2}, x, z) p(\frac{t}{2}, z, y) \leq c_4 t^{-N_4} \exp \left[- \frac{2d^2(x, z) + 2d^2(z, y)}{2t} \right].$$

La relation

$$2d^2(x, z) + 2d^2(z, y) \geq [d(x, z) + d(z, y)]^2 \geq d^2(x, y)$$

entraîne alors pour $0 < t \leq T_4$, $y \in K$

$$(35) \quad I_{N_{x,y}} \leq c_5 t^{-N_4} \exp \left[- \frac{d^2(x, y)}{2t} \right].$$

Par conséquent si $x \in M$, $y \in V$ et $d(x, y) \leq r\sqrt{2} \wedge u\sqrt{2} \wedge sd(y, \infty)$ les relations (35) (34) (32) entraînent

$$p(t, x, y) = I_{M_y} + I_{M_{xy}} + I_{N_{xy}} \leq c_6 t^{-N_6} \exp \left[- \frac{d^2(x, y)}{2t} \right]$$

pour $0 < t \leq T_6$, avec c_6 , T_6 , N_6 constantes positives.

Par continuité ceci reste vrai si $y \in \bar{V}$, ce qui prouve $H[\bar{V}, \sqrt{2}(r \wedge u), s]$, lorsque $\pi(M)$ est fini.

Si maintenant on revient au cas général où $\pi(M)$ peut être infini, on peut toujours trouver un volume riemannien $\tilde{\pi}$ tel que $\tilde{\pi}(M)$ soit fini (il suffit de multiplier la métrique associée à π par une fonction positive bien choisie). La continuité de $\frac{d\tilde{\pi}}{d\pi}$ et la formule évidente

$$\tilde{p}(t, x, y) = p(t, x, y) \frac{d\pi}{d\tilde{\pi}}(y)$$

où $\tilde{p}$ est la densité du semi-groupe (R_t) par rapport à $\tilde{\pi}$, montrent que p vérifie $H(K, r, s)$ si et seulement si p vérifie $H(K, r, s)$, ce qui achève la preuve.

4.8. Preuve de la proposition 4.4.

Supposons d'abord M non complète, ce qui équivaut à dire que la fonction $y \rightarrow d(y, \infty)$ est finie et continue. Soit K un compact de M et soit $s < 1$. L'ensemble $F = \{x \in M \mid \text{il existe } y \in K \text{ vérifiant } d(x, y) \leq sd(y, \infty)\}$ est évidemment compact. L'équivalent local de $p(t, x, y)$ - cf. L. Elie ch. 5 - prouve l'existence de $r > 0$ tel que $H(F, r, s)$ soit vérifiée, et donc à fortiori $H(K, r, s)$.

Supposons construit r_n tel que $H(K, r_n, s)$ soit vraie, avec $r_n \geq r$. Puisque $K^{r_n, s}$ est inclus dans F , on voit que $H(K^{r_n, s}, r, s)$ est vrai. D'après le lemme 4.7, il s'ensuit que $H(K, \sqrt{2}(r_n \wedge r), s)$ est vraie ; en posant $r_n = 2^{n/2}r$, on voit donc par récurrence que $H(K, 2^{n/2}r, s)$ est vraie pour tout n . Prenons n assez grand pour que $2^{n/2}r > \text{diamètre}(F)$; on voit ainsi que pour tout $s < 1$, tout compact K de M , il existe $c, T, N > 0$ tels que

$$(36) \quad p(t, x, y) \leq c t^{-N} \exp - \frac{1}{2t} d^2(x, y)$$

pour $y \in K, x \in M, d(x, y) \leq sd(y, \infty)$.

Supposons maintenant M complète. Etant donné K compact de M , soit h son diamètre. L'existence d'un équivalent local pour $p(t, x, y)$ fournit $r > 0$ tel que $H(K^{h, \frac{1}{2}}, r, \frac{1}{2})$ soit vraie. L'argument ci-dessus avec $r_n = h \wedge 2^{n/2}r$ prouve par induction que $H(K, h, \frac{1}{2})$ est vraie, et donc, puisque $d(y, \infty) \equiv \infty$, que

$$(37) \quad p(t, x, y) \leq c t^{-N} \exp \left[- \frac{d^2(x, y)}{2t} \right]$$

pour $x \in K, y \in K$.

Montrons que (36) reste vraie si on échange les rôles de x et y . Il suffit de réaliser que toute la preuve 4.5, 4.6, 4.7, 4.8 reste valable si on échange les rôles de x et y . Comme les majorations locales qui servent ensuite de point départ pour obtenir (36) sont elles symétriques en x, y , la validité de (36) sur $\{x \in K, y \in M, d(x, y) \leq sd(x, \infty)\}$ est prouvée. On notera que l'application directe de (36) à la diffusion adjointe n'est pas possible en général car le générateur différentiel est seulement supposé C^4 , de sorte que son adjoint peut fort bien n'être que C^2 .

La validité de (36) (37) et le lemme 4.6 montrent que $U(K, r, s)$ est vraie pour tout r et tout $s < 1$; de même la propriété symétrique de $U(K, r, s)$ est vraie.

Pour compléter la preuve de la prop. 4.4, donnons nous une fonction v sur M

et un compact K de M , avec v continue $0 < v(\cdot) < d(\cdot, \infty)$. Considérons l'ensemble F des x, y tels que $\{x \in M, y \in K, d(x, y) > v(y)\}$.

La validité de $H(K, r, s)$ pour $s < 1, r > 0$, permet d'appliquer la preuve du lemme 4.6 à la situation considérée ici, en remplaçant le v de la preuve 4.6 par $v(y)$. On montre ainsi l'existence de c, T, N tels que pour $0 < t \leq T, (x, y) \in F$, on ait $p(t, x, y) \leq c t^{-N} \exp \left[-\frac{v^2(y)}{2t} \right]$. Ceci achève la preuve de 4.4.

4.9. Remarque : La majoration naturelle de $p(t, x, y)$ par $c t^{-N} \exp \left[-\frac{d^2(x, y)}{2t} \right]$ est vraie pour x, y dans un compact et $d(x, y) < \max[d(x, \infty), d(y, \infty)]$.

La proposition 4.4 par conséquent ne s'applique plus dès que $\{d(x, \infty) \leq d(x, y) \text{ et } d(y, \infty) \leq d(x, y)\}$. Le contre-exemple que nous avons construit en 2.1, 2.2 montre d'ailleurs que cette majoration "naturelle" est certainement fausse, pour $\{d(x, \infty) \leq \frac{1}{25} d(x, y) \text{ et } d(y, \infty) \leq \frac{1}{25} d(x, y)\}$. Le nombre $\frac{1}{25}$ résulte d'un calcul peut efficace, et peut sans doute être remplacé par α plus proche de 1, sans changer la dernière assertion. Comme $\frac{1}{25} < 1$, ces deux résultats ne suffisent pas à élucider le domaine de validité de 4.4. Il nous semble que 1 est en fait la "bonne" valeur dans le résultat obtenu en 4.4, c'est-à-dire que l'on peut trouver des exemples de diffusions à générateur elliptique et C^∞ , et des couples x, y vérifiant $\max[d(x, \infty), d(y, \infty)] < d(x, y)$ tels que $\lim_{t \rightarrow 0} t \log p(t, x, y) > -\frac{1}{2} d^2(x, y)$.

Pour le prouver il suffit sans doute d'améliorer la précision des calculs dans l'étude du contre-exemple 2.1, 2.2.

Cependant lorsque le générateur différentiel L se prolonge agréablement "au delà du bord de M ", la majoration en $t^{-N} \exp \left[-\frac{d^2(x, y)}{2t} \right]$ est valable globalement que M soit complète ou non ; en effet on a

4.10. PROPOSITION. Soit M une variété et $\mathcal{P}$ une L -diffusion sur M avec L elliptique de classe C^4 . Alors pour tout ouvert relativement compact U de M , il existe c, T, N positifs stricts tels que la densité p_U de la L -diffusion sur U vérifie pour $0 < t \leq T$ et tout $x \in U, y \in U$,

$$(38) \quad p_U(t, x, y) \leq c t^{-N} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

Preuve : Soit $\mathfrak{m}$ la structure riemannienne associée à L . Soit K un compact de M contenant un voisinage de $\bar{U}$. Soit f une fonction C^3 , positive stricte sur M , tendant vers 0 à l'infini. Considérons la métrique riemannienne $\tilde{\mathfrak{m}}$ associée à $\tilde{L} = fL$, qui, par définition, vérifie si $x \in M$,

$$(39) \quad \|v\| = \frac{1}{\sqrt{f(x)}} \|v\| \quad \text{pour tout vecteur tangent } v \in T_x(M).$$

On peut toujours construire f comme ci-dessus, telle que $0 < f \leq 1$, $f(x) = 1$ pour $x \in K$, et vérifiant de plus $\tilde{d}(x, \infty) = +\infty$ pour tout $x \in K$, où $\tilde{d}$ est la distance associée à $\tilde{m}$. Fixons une telle fonction f .

Soit $\tilde{p}$ la π -densité de la L -diffusion. Il existe (prop. 4.4) c, N, T tels que

$$(40) \quad \tilde{p}(t, x, y) \leq c t^{-N} \exp \left[-\frac{1}{2t} \tilde{d}^2(x, y) \right] \text{ sur l'ensemble } \{0 < t \leq T, x \in K, y \in K\}.$$

Comme $0 < f \leq 1$, la relation (39) montre que

$$(41) \quad \tilde{d}(x, y) \geq d(x, y) \text{ pour tout } x, y \in M.$$

Soit V un ouvert tel que $\bar{U} \subset V \subset K$. Comme $L = \tilde{L}$ sur V , on a $p_V \equiv \tilde{p}_V$. La définition des diffusions induites et la continuité de $p_U, p_V, \tilde{p}_V, \tilde{p}$ entraîne $p_U \leq p_V$ et $\tilde{p}_V \leq \tilde{p}$, de sorte que finalement pour $t > 0$, $x \in U, y \in U$

$$(42) \quad p_U(t, x, y) \leq \tilde{p}(t, x, y)$$

Combinant (40) (41) (42) on obtient l'inégalité (38) annoncée.

BIBLIOGRAPHIE

- [1] R. Azencott, *Grandes déviations pour la loi des grands nombres, et petites perturbations de systèmes dynamiques*. Ecole d'été St Flour 1978, Lecture Notes, v. 774, Springer Verlag 1980.
- [2] E. Dynkin, *Markov processes*, Springer Verlag, Berlin, 1965.
- [3] S.A. Molchanov, *Diffusion processes and riemannian geometry*. Russ. Math. Surveys 30:1 (1975) p.1 à 63.
- [4] D. Stroock ; S. Varadhan, *Diffusion processes*. Comm. Pure. App. Math. 22, 1969, p.345-400 et 479-530.

CHAPITRE 9

LE COMPORTEMENT ASYMPTOTIQUE DE $p(t,x,y)$ QUAND $t \rightarrow 0$ (POINTS ELOIGNES)

par

Catherine BELLAÏCHE

1. Introduction

1.1.- Soit M une variété connexe de classe C^5 , de dimension m . On considère sur M une diffusion dont le générateur infinitésimal L est un opérateur différentiel elliptique vérifiant $L1 = 0$. En coordonnées locales, L s'exprime sous la forme

$$L = \frac{1}{2} \sum_{1 \leq i, j \leq m} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq m} b_i(x) \frac{\partial}{\partial x_i}.$$

On suppose que les coefficients a_{ij} sont de classe C^4 et les coefficients b_i de classe C^3 . On introduit sur M une structure riemannienne de classe C^4 associée à L par

$$ds^2 = \sum_{1 \leq i, j \leq m} g_{ij}(x) dx_i dx_j$$

où $g(x) = [a(x)]^{-1}$. L'opérateur L s'écrit alors $\frac{1}{2} \Delta + Y$, Δ étant l'opérateur de Laplace-Beltrami et Y un opérateur différentiel du premier ordre.

1.2.- Notre but est d'étudier le comportement asymptotique quand $t \rightarrow 0$ de la densité $p(t,x,y)$ par rapport au volume riemannien π du semi-groupe de transition de la diffusion. On a déjà vu (L. Elie, ch. 5 et Molchanov [10], § 2) que pour des points x et y suffisamment voisins, on a l'équivalent

$$(1) \quad p(t,x,y) \sim (2\pi t)^{-m/2} H(x,y) \exp \left[-\frac{d^2(x,y)}{2t} + A(x,y) \right]$$

quand $t \rightarrow 0$, où $H(x,y) = \det [d \exp_x(\dot{\gamma}(0))]^{-1/2}$, $\gamma : [0,1] \rightarrow M$ étant l'unique géodésique de longueur minimale joignant x à y et où $A(x,y) = \int_0^1 \langle Y, \dot{\gamma}(s) \rangle ds$. De plus, cet équivalent est uniforme pour x et y assez voisins dans un même compact.

Dans la partie 2, en utilisant des majorations globales de la densité (R. Azencott, ch. 8), on ramène le calcul de l'équivalent de $p(t,x,y)$ au calcul de l'équivalent d'une intégrale de la forme

$$(2) \int_{\Omega_{xy}^{n,h}} \exp \left[-\frac{E(\gamma)}{t} + A(\gamma) \right] H(\gamma) \pi_n(d\gamma).$$

On rappelle que l'énergie d'un chemin $\gamma : [0,1] \rightarrow M$ est défini par $E(\gamma) = \frac{1}{2} \int_0^1 |\dot{\gamma}|^2 dt$. Dans (2), $\Omega_{xy}^{n,h}$ est l'espace des géodésiques brisées en n points $y_1, \dots, y_n$ ($y_i = \gamma(\frac{i}{n+1})$) et d'énergie $\leq \frac{1}{2} [d^2(x,y) + h]$. Pour n assez grand, la restriction de γ à l'intervalle $[\frac{i}{n+1}, \frac{i+1}{n+1}]$ est pour tout $\gamma \in \Omega_{xy}^{n,h}$, l'unique géodésique minimisante joignant $\gamma(\frac{i}{n+1})$ à $\gamma(\frac{i+1}{n+1})$ et on peut identifier par l'application $\gamma \mapsto (\gamma(\frac{i}{n+1}))$, $\Omega_{xy}^{n,h}$ à une partie de M^n . On pose

$$A(\gamma) = \int_0^1 \langle Y, \dot{\gamma} \rangle dt, \quad H(\gamma) = \prod_{i=0}^n H(y_i, y_{i+1})$$

et on désigne par π_n la restriction à $\Omega_{xy}^{n,h}$ de la mesure $\pi^{\otimes n}$ sur M^n . D'autre part, on montre que l'estimation de Varadhan $\lim_{t \rightarrow 0} 2t \log p(t,x,y) = -d^2(x,y)$ est vraie en général sur une variété complète. Pour une variété non complète, on prouve seulement que $\liminf_{t \rightarrow 0} 2t \log p(t,x,y) \geq -d^2(x,y)$ (lemme 2.1).

Le théorème principal de ce chapitre (Molchanov[10] th. 3.1 pour les variétés compactes) que nous démontrons dans la partie 4 affirme que si M est complète pour la distance riemannienne d , la formule asymptotique (1) est encore valable pour presque tous les points x et y de M - précisément pour ceux qui sont joints par une unique géodésique de longueur minimale le long de laquelle ils sont non conjugués.

La formule (1) est valable uniformément sur tout compact de son domaine de validité.

Ce théorème, comme les suivants, se démontre en appliquant à l'intégrale (2) la méthode de Laplace (dont nous donnons un exposé dans la partie 3). La méthode de Laplace montre que, dans tous les cas, l'équivalent de $p(t,x,y)$ dépend du comportement de la fonction énergie $E(\gamma)$ au voisinage de ses minimums (les géodésiques minimisantes).

Dans le cas où x et y sont conjugués - qui est le cas où le minimum de $E(\gamma)$ est

dégénéré - il est impossible de faire une étude exhaustive. Suivant Molchanov, nous déterminons le comportement asymptotique de $p(t, x, y)$ lorsque la variété M est complète

1) si les géodésiques minimisantes joignant x à y forment une variété de dimension k et si la multiplicité de la conjugaison sur chacune d'elles est exactement égale à k (c'est le cas par exemple lorsque x et y sont des pôles opposés d'une sphère ou les sommets d'une surface de révolution) ; on trouve un équivalent de la forme

$$p(t, x, y) \sim \text{const. } t^{-\frac{m}{2} - \frac{k}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right] ;$$

2) pour une diffusion elliptique générique si la variété M est de dimension ≤ 1 et pour x et y quelconques dans M . On trouve des formules du type

$$(3) \quad p(t, x, y) \sim \text{const. } t^{-\frac{m}{2} - \frac{\alpha}{2}} |\log t|^\beta \exp \left[-\frac{d^2(x, y)}{2t} \right] .$$

L'interprétation de α et β est beaucoup plus difficile que dans le cas 1). Elle fait intervenir la nature de dégénérescence du minimum de E et on doit utiliser la classification d'Arnol'd des points critiques des fonctions différentiables.

On a d'ailleurs toujours un équivalent du type (3) (avec $\alpha \in \mathbb{Q}$, $0 \leq \alpha \leq m-1$, $\beta \in \mathbb{N}$) si on suppose la variété et sa métrique analytiques. Un contre-exemple montre que ce n'est pas le cas si on suppose seulement la métrique de classe C^∞ .

Toujours dans le cas où M est complète et la métrique de classe C^4 , on a un encadrement de $p(t, x, y)$ qui précise l'estimation de Varadhan

$$\text{const. } t^{-m/2} \exp \left[-\frac{d^2(x, y)}{2t} \right] \leq p(t, x, y) \leq \text{const. } t^{-m + \frac{1}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

pour t assez petit.

Tous les résultats précédents restent valables, lorsque M n'est plus supposée complète mais lorsque les points x et y vérifient la condition $d(x, y) < \sup [d(x, \infty), d(y, \infty)]$, ∞ désignant le point à l'infini du compactifié d'Alexandroff de M .

On termine en calculant à titre d'application quelques équivalents de probabilités de transition.

2. Approximation de $p(t, x, y)$ par une intégrale en dimension finie

On considère sur M une diffusion satisfaisant aux hypothèses données en 1.1. On ne suppose pas pour l'instant la variété M complète pour la distance riemannienne associée.

2.1. LEMME. Pour tout compact K de M et pour tout $\chi > 0$, il existe $T > 0$ tel que l'on ait

$$(4) \quad p(t, x, y) \geq \exp \left[-\frac{d^2(x, y)}{2t} - \frac{\chi}{t} \right]$$

pour $0 < t \leq T$ et $x, y \in K$.

Démonstration. Soient $x_0, y_0 \in K$. On part de la formule de Chapman-Kolmogoroff

$$p(t, x_0, y_0) = \int_{M^n} p\left(\frac{t}{n+1}, x_0, y_1\right) p\left(\frac{t}{n+1}, y_1, y_2\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n)$$

Il résulte du théorème 1 (L. Elie, ch. 5) que pour tout compact H de M , il existe $\varepsilon, C, T > 0$ tels que

$$p(t, x, y) \geq C t^{-m/2} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

pour $0 < t \leq T$, $x, y \in H$ et $d(x, y) \leq \varepsilon$.

La minoration (4) sera démontrée pour x_0 et y_0 si on peut trouver un compact H contenant x_0 et y_0 , un entier n et une partie compacte L , d'intérieur non vide, de H^n tels que

$$(5) \quad (y_1, \dots, y_n) \in L \Rightarrow$$

$$d(x_0, y_1) < \varepsilon(H), \quad d(y_n, y_0) < \varepsilon(H), \quad d(y_i, y_{i+1}) < \varepsilon(H) \quad (1 \leq i \leq n-1)$$

et

$$(n+1) [d^2(x_0, y_1) + \dots + d^2(y_n, y_0)] < d^2(x_0, y_0) + \chi.$$

En effet, on aura alors pour $t \leq T$,

$$\begin{aligned} p(t, x_0, y_0) &\geq \int_L p\left(\frac{t}{n+1}, x_0, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y_0\right) \pi(dy_1) \dots \pi(dy_n) \\ &\geq C^{n+1} t^{-m(n+1)/2} \int_L \exp \left[-\frac{n+1}{2t} (d^2(x_0, y_1) + \dots + d^2(y_n, y_0)) \right] \pi(dy_1) \dots \pi(dy_n) \\ &\geq C^{n+1} t^{-m(n+1)/2} \exp \left[-\frac{d^2(x_0, y_0)}{2t} - \frac{\chi}{2t} \right] \pi^n(L) \\ &\geq \exp \left[-\frac{d^2(x_0, y_0)}{2t} - \frac{\chi}{t} \right]. \end{aligned}$$

Si on suppose de plus que x_0 et y_0 sont dans l'intérieur de H , comme (5) reste vérifiée si on remplace x_0 et y_0 par des points x et y suffisamment voisins, le même raisonnement montre que la minoration (4) sera encore valable pour $x \in B$ et $y \in B'$ où B (resp. B') est un voisinage de x_0 (resp. y_0), donc par un argument de compacité pour tout $x \in K$ et tout $y \in K$.

Il reste à prouver l'existence de H , n , L . Soit $c : [0,1] \rightarrow M$ une courbe C^1 par morceaux parcourue à vitesse uniforme, joignant x_0 à y_0 et d'énergie $< \frac{1}{2} [d^2(x_0, y_0) + \frac{\chi}{2}]$. On prend pour H un compact dont l'intérieur contient l'image de c et pour L l'ensemble des $(y_1, \dots, y_n) \in H^n$ vérifiant

$$(n+1) [d^2(x_0, y_1) + \dots + d^2(y_n, y_0)] \leq d^2(x_0, y_0) + \frac{\chi}{2}.$$

Si l'on choisit l'entier n vérifiant $\frac{1}{n+1} [d^2(x_0, y_0) + \frac{\chi}{2}] < \varepsilon(H)^2$, (5) sera alors vérifiée puisque si $(y_1, \dots, y_n) \in L$, on a

$$d^2(y_i, y_{i+1}) \leq \frac{d^2(x_0, y_0) + \frac{\chi}{2}}{n+1} \leq \varepsilon^2 \quad \text{pour } i = 1, \dots, n-1.$$

De plus, L est d'intérieur non vide car si l'on pose $y_1^0 = c(\frac{1}{n+1})$ on a

$$(n+1) [d^2(x_0, y_1^0) + \dots + d^2(y_n^0, y_0)] \leq 2 E(c) < d^2(x_0, y_0) + \frac{\chi}{2},$$

inégalité valable encore pour des points $y_1, \dots, y_n$ suffisamment voisins de $y_1^0, \dots, y_n^0$, ce qui prouve que le point $(y_1^0, \dots, y_n^0)$ est dans l'intérieur de L .

2.2. Lorsque la variété M n'est pas complète pour la distance riemannienne d , on note ∞ le point à l'infini du compactifié d'Alexandroff de M et on dit que deux points x et y vérifient la condition (C) si

$$d(x, y) < \sup [d(x, \infty), d(y, \infty)].$$

COROLLAIRE. Pour tout compact K de $M \times M$, on a

$$(6) \quad \liminf_{t \rightarrow 0} 2 t \operatorname{Log} p(t, x, y) \geq -d^2(x, y)$$

uniformément pour $(x, y) \in K$. Si la variété M est complète pour d ou si pour tout $(x, y) \in K$, x et y vérifient la condition (C), on a

$$(7) \quad \lim_{t \rightarrow 0} 2 t \operatorname{Log} p(t, x, y) = -d^2(x, y)$$

uniformément sur K .

Démonstration. L'inégalité (6) résulte immédiatement du lemme 2.1. Si M est complète ou si K est contenu dans l'ouvert $\{(x,y) \mid d(x,y) < \sup [d(x,\infty), d(y,\infty)]\}$, il résulte de la proposition 4.4 (ch.8 R. Azencott) qu'il existe $C, T, N > 0$ tels que

$$p(t,x,y) \leq C t^{-N} \exp \left[- \frac{d^2(x,y)}{2t} \right]$$

pour $0 < t \leq T$ et $(x,y) \in K$. On en déduit que

$$\limsup_{t \rightarrow 0} 2t \operatorname{Log} p(t,x,y) \leq -d^2(x,y)$$

uniformément sur K .

2.3. LEMME. Soit U un ouvert relativement compact de M . Pour tout compact $K \subset M \times M$ tel que pour tout $(x,y) \in K$ on ait $d(x,y) < \sup d(x, \partial U), d(y, \partial U)$, on a

$$(8) \quad p(t,x,y) \sim p_U(t,x,y)$$

quand $t \rightarrow 0$, uniformément sur K .

Démonstration. Soit $(x_0, y_0) \in K$. Supposons par exemple que $d(x_0, y_0) < d(y_0, \partial U)$ et soit r tel que $d(x_0, y_0) < r < d(y_0, \partial U)$. On note B la boule ouverte de centre y_0 et de rayon r . On a $B \subset U$. Soit $\varepsilon > 0$ tel que $\varepsilon < \inf(r, \frac{1}{2}[d(y_0, \infty) - r])$ et soit V la boule ouverte de centre y_0 et de rayon ε . Pour $z \in \partial B$ et $y \in V$ on a

$$d(z,y) \leq d(z, y_0) + d(y_0, y) < r + \varepsilon \leq d(y_0, \infty) - \varepsilon \leq d(y, \infty)$$

donc il existe $C, N > 0$ tels que

$$p(s,z,y) \leq C s^{-N} \exp \left[- d^2 \frac{(z,y)}{2s} \right]$$

pour s assez petit (cf. R. Azencott ch. 8 prop. 4.4). Pour $x \in B$ et $y \in V$ on a

$$\begin{aligned} p(t,x,y) &\leq p_B(t,x,y) + \sup_{\substack{s \leq t \\ z \in \partial B}} p(s,z,y) \\ &\leq p_B(t,x,y) + C t^{-N} \exp \left[- \frac{(r - \varepsilon)^2}{2t} \right] \end{aligned}$$

pour t assez petit, puisque $d(z, y) \geq r - \varepsilon$ pour $z \in \partial B$.

D'autre part, pour x assez proche de x_0 on a

$$d(x, y) \leq d(x_0, y_0) + 2\varepsilon < r - \varepsilon$$

si l'on impose à ε de vérifier de plus la condition $\varepsilon < \frac{1}{3} [r - d(x_0, y_0)]$.

On en déduit que pour t assez petit, x (resp. y) assez voisin de x_0 (resp. y_0), on a

$$p_B(t, x, y) \leq p(t, x, y) \leq p_B(t, x, y) + \exp \left[-\frac{d^2(x, y)}{2t} \right],$$

et en utilisant le lemme 2.1, il en résulte que $p(t, x, y) \sim p_B(t, x, y)$ uniformément sur des voisinages de x_0 et y_0 . Puisque $p_B(t, x, y) \leq p_U(t, x, y) \leq p(t, x, y)$, on a aussi $p_U(t, x, y) \sim p(t, x, y)$ uniformément sur des voisinages de x_0 et y_0 donc aussi sur K par un argument de compacité.

2.4. PROPOSITION. *Supposons la variété M non complète pour d et soient $x_0, y_0 \in M$ vérifiant la condition (C) de 2.2. Il existe des ouverts U et N contenant x_0 et y_0 tels que $U \subset N$ et une métrique riemannienne g_N sur N coïncidant sur U avec la métrique initiale faisant de N une variété riemannienne complète, il existe de plus des voisinages B et B' de x_0 et y_0 contenus dans U tels que, si l'on note d_N la distance associée à g_N , l'on ait pour $(x, y) \in B \times B'$.*

$$1) \quad d(x, y) = d_N(x, y)$$

$$2) \quad p(t, x, y) \sim p_N(t, x, y) \quad \text{quand } t \rightarrow 0 \text{ uniformément sur } B \times B'$$

3) *Pour ε assez petit, les chemins paramétrés de 0 à 1 joignant x à y , d'énergie $\leq \frac{d^2(x, y)}{2} + \varepsilon$ et contenus dans N sont les mêmes que ceux contenus dans M et sont d'ailleurs dans U .*

Démonstration. Supposons par exemple que $d(x_0, \infty) \leq d(y_0, \infty)$ et soient r et R tels que $d(x_0, y_0) < r < R < d(y_0, \infty)$. On note U (resp. N) la boule ouverte de centre y_0 de rayon r (resp. R)

Puisque $d(x_0, y_0) < \sup(d(x_0, \partial U), d(y_0, \partial U))$, par continuité il existe des voisinages compacts B et B' de x_0 et y_0 contenus dans U tels que l'on ait encore

$$d(x, y) < \sup[d(x, \partial U), d(y, \partial U)]$$

pour $(x, y) \in B \times B'$. Le lemme 2.3 montre alors que

$$p(t, x, y) \sim p_U(t, x, y)$$

uniformément sur $B \times B'$.

Soit ϕ une fonction de classe $C^3 : N \rightarrow [1, +\infty[$ telle que $\phi(z) = 1$ si $z \in U$. On considère sur N la métrique $g^N(z) = \phi(z) g(z)$. Puisque $\phi \geq 1$, on a $d(x, y) \leq d_N(x, y)$ pour $x, y \in N$. Soit $\alpha > 0$ tel que

$$(9) \quad d(x, y) + \alpha < \sup [d(x, \partial U), d(y, \partial U)]$$

pour $x \in B, y \in B'$. Tout chemin de longueur $\leq d(x, y) + \alpha$ joignant x à y est alors contenu dans U , il a donc dans N la même longueur ce qui prouve que $d_N(x, y) \leq d(x, y) + \alpha$. On en déduit 1).

Pour $x \in B, y \in B'$ on a donc $d_N(x, y) < \sup [d_N(x, \partial U), d_N(y, \partial U)]$ et d'après le lemme 2.3, $p_N(t, x, y) \sim p_U(t, x, y)$ uniformément sur $B \times B'$. On en déduit 2).

Soit $\alpha > 0$ vérifiant (9) et soit γ un chemin paramétré de 0 à 1 joignant x à y . Si γ sort de U à l'instant s on a

$$\begin{aligned} E(\gamma) &\geq \frac{1}{2} \left[\frac{d^2(x, \gamma(s))}{s} + \frac{d^2(\gamma(s), y)}{1-s} \right] \\ &\geq \frac{1}{2} \left[\frac{d^2(x, \partial U)}{s} + \frac{d^2(y, \partial U)}{1-s} \right] \\ &\geq \frac{1}{2} \sup [d(x, \partial U), d(y, \partial U)]^2 \\ &\geq \frac{1}{2} d^2(x, y) + \frac{1}{2} \alpha^2. \end{aligned}$$

La partie 3) de 2.4 en résulte en prenant $\varepsilon < \frac{1}{2} \alpha^2$.

On vérifie d'autre part que si l'on impose à ϕ de vérifier la condition

$$\phi(z) \geq \frac{1}{R - d(y_0, z)} \quad \text{si} \quad \frac{r + R}{2} \leq d(y_0, z) < R,$$

pour la nouvelle métrique la longueur de toute courbe $\gamma : [a, b] \rightarrow N$ telle que $\gamma(t) \rightarrow \infty$ dans N lorsque $t \rightarrow b$ est infinie, ce qui montre que N est complète.

2.5.- La proposition 2.4 permettant de ramener le cas de points x et y vérifiant la condition (C) au cas d'une variété complète, nous supposons dans le lemme suivant la variété M complète.

LEMME. On suppose la variété M complète pour la distance riemannienne d . Pour tout compact K de M , il existe un compact L contenant K tel que l'on ait pour tout entier $n \geq 1$,

$$(10) \quad p(t, x, y) \sim \int_{L^n} p\left(\frac{t}{n+1}, x, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n)$$

lorsque $t \rightarrow 0$, uniformément pour $x \in K$, $y \in K$.

Démonstration. Soit r le diamètre de K et soit $L = K_r$ l'ensemble des points dont la distance à K est $\leq r$. On a pour $x \in K$, $y \in K$

$$\int_{(L^n)^c} p\left(\frac{t}{n+1}, x, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n) \leq \sum_{k=1}^n J_k$$

$$\text{où } J_k = \int_{M^{k-1} \times_L C \times_M M^{n-k}} p\left(\frac{t}{n+1}, x, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n).$$

L'intégration sur les variables $y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_n$ donne

$$J_k = \int_{L^C} p\left(\frac{kt}{n+1}, x, y_k\right) p\left(\frac{(n-k+1)t}{n+1}, y_k, y\right) \pi(dy_k).$$

Comme $y_k \in L^C$ et $y \in K$, on a $d(y_k, y) > r$ et il existe $C, T, N > 0$ tels que

$$p\left(\frac{(n-k+1)t}{n+1}, y_k, y\right) \leq C \left(\frac{(n-k+1)t}{n+1}\right)^{-N} \exp \left[-\frac{(n+1)r^2}{2(n-k+1)t}\right]$$

pour $t \leq T$ (cf. R. Azencott ch. 8 Prop. 4.4).

Puisque $\int_{L^C} p\left(\frac{kt}{n+1}, x, y_k\right) \pi(dy_k) \leq 1$, on en déduit que pour $0 < t \leq T$ et

$k = 1, \dots, n$

$$J_k \leq C \left(\frac{t}{n+1}\right)^{-N} \exp \left[-\frac{(n+1)r^2}{2nt}\right].$$

Posons $J_0 = \int_{L^n} p\left(\frac{t}{n+1}, x, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n)$. On a pour $t \leq T$,

$$0 \leq p(t, x, y) - J_0 \leq n C \left(\frac{t}{n+1} \right)^{-N} \exp \left[- \frac{(n+1)r^2}{2nt} \right]$$

D'après le lemme 2.1, la différence $p(t, x, y) - J_0$ est uniformément négligeable devant $p(t, x, y)$ si on peut trouver $\chi > 0$ tel que

$$d^2(x, y) + 2\chi < \frac{n+1}{n} r^2$$

pour tout $x \in K$, $y \in K$. Comme $d^2(x, y) \leq r$ pour $x \in K$, $y \in K$, l'existence de χ est assurée et on a $p(t, x, y) \sim J_0$ uniformément sur $K \times K$.

2.6.- Etant donné $x, y \in M$ vérifiant la condition (C), un entier $n \geq 0$ et $h > 0$ on note $M_{xy}^{n,h}$ l'ensemble des $(y_1, \dots, y_n) \in M^n$ tels que

$$(n+1) [d^2(x, y_1) + \dots + d^2(y_n, y)] < d^2(x, y) + h$$

Si $(y_1, \dots, y_n) \in M_{xy}^{n,h}$, un calcul simple montre que

$$d(x, y_i) \leq \sqrt{d^2(x, y) + h} \quad \text{et} \quad d(y, y_i) \leq \sqrt{d^2(x, y) + h}$$

pour $i = 1, \dots, n$. Pour h vérifiant

$$\sqrt{d^2(x, y) + h} < \sup [d(x, \infty), d(y, \infty)],$$

les points $x, y_1, \dots, y_n, y$ sont donc dans un même compact K ne dépendant que de x, y, h . Soit $\varepsilon(x, y, h)$ le rayon d'injectivité de K , c'est-à-dire la borne inférieure des rayons d'injectivité des points de K . Si n est choisi assez grand pour que

$$\frac{d^2(x, y) + h}{n+1} < \varepsilon^2(x, y, h)$$

on peut associer à tout point $(y_1, \dots, y_n)$ de $M_{xy}^{n,h}$ une géodésique brisée $\gamma(x, y_1, \dots, y_n, y)$ (cf. A. Bellaïche, ch. 7, 8.3) et $M_{xy}^{n,h}$ est en bijection avec l'espace $\Omega_{xy}^{n,h}$ des géodésiques brisées d'énergie $< \frac{1}{2} [d^2(x, y) + h]$. Pour $\gamma \in \Omega_{xy}^{n,h}$, on pose

$$A(\gamma) = \int_0^1 \langle Y, \dot{\gamma} \rangle dt = \sum_{i=0}^n A(y_i, y_{i+1})$$

$$H(\gamma) = \prod_{i=0}^n H(y_i, y_{i+1})$$

et on note $\pi_n(dy)$ la mesure image de $\pi^{\otimes n}$ par l'application $(y_1, \dots, y_n) \mapsto \gamma(x, y_1, \dots, y_n, y)$.

PROPOSITION.— Soient x_o et $y_o \in M$. On suppose que M est complète ou que x_o et y_o vérifient la condition (C) de 2.2. Pour $h > 0$ assez petit et n assez grand, il existe des voisinages B de x_o et B' de y_o tels que

- 1) si $(y_1, \dots, y_n) \in M_{x_o y_o}^{n, h}$, $x \in B$, $y \in B'$, il existe une unique géodésique brisée $\gamma = \gamma(x, y_1, \dots, y_n, y)$ joignant les points $x, y_1, \dots, y_n, y$
- 2) on a uniformément sur $B \times B'$

$$(11) \quad p(t, x, y) \sim \left(\frac{2\pi t}{n+1}\right)^{-m(n+1)/2} \int_{M_{x_o y_o}^{n, h}} \exp \left[-\frac{E(\gamma)}{t} + A(\gamma)\right] H(\gamma) \pi(dy_1) \dots \pi(dy_n)$$

lorsque $t \rightarrow 0$.

Démonstration.— D'après la proposition 2.4, il existe une variété complète N et des voisinages B et B' de x_o et y_o tels que $p(t, x, y) \sim p_N(t, x, y)$ sur $B \times B'$, et d'après 2.4. 3) on a $M_{x_o y_o}^{n, h} = N_{x_o y_o}^{n, h}$ pour h assez petit et les fonctions énergies associées à M et à N coïncident sur $M_{x_o y_o}^{n, h}$. Il suffit donc de démontrer (11) dans le cas d'une variété complète.

Soit K la boule fermée de centre x_o et de rayon $\sqrt{d^2(x_o, y_o)} + h$ et L le compact qui lui est associé par le lemme 2.5. On a pour tout entier $n \geq 1$,

$$\begin{aligned} p(t, x, y) &\sim \int_{M_{x_o y_o}^{n, h}} p\left(\frac{t}{n+1}, x, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n) \\ &+ \int_{L^n - M_{x_o y_o}^{n, h}} p\left(\frac{t}{n+1}, x, y_1\right) \dots p\left(\frac{t}{n+1}, y_n, y\right) \pi(dy_1) \dots \pi(dy_n) \\ &= I + J \end{aligned}$$

lorsque $t \rightarrow 0$ uniformément pour $x \in K$, $y \in K$.

Il existe ε (ne dépendant que de K donc de x_o, y_o, h) tel que

$$p(t, x, y) \sim (2\pi t)^{-m/2} H(x, y) \exp \left[-\frac{d^2(x, y)}{2t} + A(x, y)\right]$$

uniformément pour $x \in K$, $y \in K$ et $d(x,y) \leq \varepsilon$. De plus ε est inférieur au rayon d'injectivité de K (cf. ch. 4, th. 1).

Comme on a vu que si $(y_1, \dots, y_n) \in M_{x_0, y_0}^{n,h}$ les points $y_1, \dots, y_n$ sont tous dans le compact K , on peut dans I remplacer chaque terme $p(\frac{t}{n+1}, y_i, y_{i+1})$ par son équivalent pourvu que $d(y_i, y_{i+1}) \leq \varepsilon$. Si l'on choisit l'entier n assez grand pour que

$$\frac{d^2(x_0, y_0) + h}{n+1} < \varepsilon^2$$

on aura bien pour $(y_1, \dots, y_n) \in M_{x_0, y_0}^{n,h}$ $d(y_i, y_{i+1}) < \varepsilon$ et $d(x_0, y_1) < \varepsilon$, $d(y_n, y_0) < \varepsilon$. En réduisant éventuellement la taille de B et B' , on aura encore $d(x, y_1) < \varepsilon$ et $d(y_n, y) < \varepsilon$, pour $x \in B$, $y \in B'$, $(y_1, \dots, y_n) \in M_{x_0, y_0}^{n,h}$. On a alors

$$\begin{aligned} I &\sim \left(\frac{2\pi t}{n+1}\right)^{-m(n+1)/2} \int_{M_{x_0, y_0}^{n,h}} \exp \left[-\frac{n+1}{2t} (d^2(x, y_1) + \dots + d^2(y_n, y)) \right. \\ &\quad \left. + A(x, y_1) + \dots + A(y_n, y) \right] H(x, y_1) \dots H(y_n, y) \pi(dy_1) \dots \pi(dy_n) \\ &= \left(\frac{2\pi t}{n+1}\right)^{-m(n+1)/2} \int_{M_{x_0, y_0}^{n,h}} \exp \left[-\frac{E(\gamma)}{t} + A(\gamma) \right] H(\gamma) \pi(dy_1) \dots \pi(dy_n) \end{aligned}$$

uniformément sur $B \times B'$.

D'autre part, en utilisant les majorations globales de la densité (R. Azencott, ch. 8, prop. 4.4), on obtient

$$J \leq C^{n+1} \left(\frac{t}{n+1}\right)^{-N(n+1)} \int_{L^n - M_{x_0, y_0}^{n,h}} \exp \left[-\frac{n+1}{2t} (d^2(x, y_1) + \dots + d^2(y_n, y)) \right] \pi(dy_1) \dots \pi(dy_n)$$

Comme pour x et y suffisamment voisins de x_0 et y_0 , on a

$$(12) \quad M_{x,y}^{n,h/2} \subset M_{x_0,y_0}^{n,h} \subset M_{x,y}^{n,h}$$

en réduisant éventuellement la taille de B et B' on obtient pour $(y_1, \dots, y_n) \in L^n - M_{x_0, y_0}^{n,h}$, $x \in B$, $y \in B'$

$$\begin{aligned}
 (n+1) [d^2(x, y_1) + \dots + d^2(y_n, y)] &\geq (n+1) [d^2(x_0, y_1) + \dots + d^2(y_n, y_0)] \\
 &\geq d^2(x, y) + \frac{h}{2},
 \end{aligned}$$

donc

$$J \leq \text{const.} \exp \left[-\frac{d^2(x, y) + h/2}{2t} \right]$$

pour t assez petit, $x \in B$, $y \in B'$. Le lemme 2.1 montre alors que J est uniformément négligeable devant $p(t, x, y)$ lorsque $t \rightarrow 0$ et (11) en découle.

3. La méthode de Laplace

3.1.- Pour obtenir des équivalents explicites de $p(t, x, y)$ à partir de la formule (11) nous sommes amenés à considérer des intégrales de la forme

$$I(t) = \int_K \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx).$$

L'évaluation asymptotique de telles intégrales porte le nom de méthode de Laplace (Dieudonné [5], Erdelyi [6]).

Dans la suite on désignera par f et g des fonctions continues sur un compact K d'intérieur non vide d'une variété N de dimension n munie d'une forme volume π avec $g > 0$.

On observe d'abord que la principale contribution à l'intégrale $I(t)$ est apportée par un voisinage V arbitraire de l'ensemble M_f des minimums de la fonction f .

En effet, pour tout voisinage V de M_f , il existe $\varepsilon > 0$ tel que

$$f(x) < \varepsilon \implies x \in V$$

(on suppose pour simplifier que la valeur minimale de f est 0). On a alors

$$\begin{aligned}
 \int_{K-V} \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) &\leq \int_{\{f \geq \varepsilon\}} \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) \\
 &\leq \exp \left(-\frac{\varepsilon}{t} \right) \int_{\{f \geq \varepsilon\}} g(x) \pi(dx) \\
 &= O \left(\exp \left(-\frac{\varepsilon}{2t} \right) \right) \quad \text{quand } t \rightarrow 0
 \end{aligned}$$

tandis que

$$\begin{aligned} \int_V \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) &\geq \int_{\{f < \varepsilon\}} \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) \\ &\geq \int_{\{f < \varepsilon/2\}} \\ &\geq C \exp(-\varepsilon/2t) \end{aligned}$$

où C est une constante > 0 , ce qui prouve que

$$(13) \quad I(t) \sim \int_V \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) \quad \text{quand } t \rightarrow 0.$$

Remarquons aussi que si f admet en $a \in \mathring{K}$ un minimum isolé, on a pour tout voisinage compact V de a ne contenant pas d'autre minimum de f

$$(14) \quad I_V(t) = \int_V \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) \sim g(a) \int_V \exp \left[-\frac{f(x)}{t} \right] \pi(dx)$$

lorsque $t \rightarrow 0$.

En effet, on peut écrire

$$I_V(t) = g(a) \int_V \exp \left[-\frac{f(x)}{t} \right] \pi(dx) + \int_V \exp \left[-\frac{f(x)}{t} \right] [g(x) - g(a)] \pi(dx)$$

donc il suffit de montrer que si $h(a) = 0$, on a

$$\int_V \exp \left[-\frac{f(x)}{t} \right] h(x) \pi(dx) = 0 \quad \left(\int_V \exp \left[-\frac{f(x)}{t} \right] \pi(dx) \right) \quad \text{quand } t \rightarrow 0$$

On peut supposer $f(a) = 0$. Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que $x \in V$ et $f(x) < \eta$ implique $h(x) < \varepsilon$. Il en résulte que

$$\begin{aligned} \int_V \exp \left[-\frac{f(x)}{t} \right] h(x) \pi(dx) &= \int_{\{f < \eta\} \cap V} \exp \left[-\frac{f(x)}{t} \right] h(x) \pi(dx) + \\ &\quad + \int_{\{f > \eta\} \cap V} \exp \left[-\frac{f(x)}{t} \right] h(x) \pi(dx) \\ &< \varepsilon \int_V \exp \left[-\frac{f(x)}{t} \right] \pi(dx) + \text{const.} \int_{\{f > \eta\}} \exp \left[-\frac{f(x)}{t} \right] \pi(dx) \\ &< 2\varepsilon \int_V \exp \left[-\frac{f(x)}{t} \right] \pi(dx) \end{aligned}$$

pour t assez petit puisque d'après ce qui précède, $\int_{\{f>\eta\}} \exp \left[-\frac{f(x)}{t}\right] \pi(dx)$ est négligeable devant $\int_V \exp \left[-\frac{f(x)}{t}\right] \pi(dx)$.

3.2.- On ne traitera pas le cas où le minimum de f est atteint sur le bord de K , sauf en dimension 1 :

PROPOSITION.- Soit f une fonction de classe C^1 sur un intervalle $[a, b] \subset \mathbb{R}$. On suppose que f admet en a un minimum unique et que $f'(a) > 0$. Soit g une fonction continue sur $[a, b]$ telle que $g(a) > 0$. Alors

$$(15) \quad I(t) = \int_a^b \exp \left[-\frac{f(x)}{t}\right] g(x) dx \sim t \frac{g(a)}{f'(a)} \exp \left[-\frac{f(a)}{t}\right] \quad \text{quand } t \rightarrow 0.$$

Démonstration.- Sans changer l'équivalent, on peut diminuer b de telle sorte que $f'(x) > 0$ sur $[a, b]$. On a

$$\begin{aligned} I(t) &\sim t \exp \left[-\frac{f(a)}{t}\right] \int_0^{\frac{b-a}{t}} \exp \left[-\frac{f(a+ty) - f(a)}{t}\right] g(a+ty) dy \\ \text{Or } \lim_{t \rightarrow 0} \int_0^{\frac{b-a}{t}} \exp \left[-\frac{f(a+ty) - f(a)}{t}\right] g(a+ty) dy &= g(a) \int_0^{\infty} \exp \left[-f'(a)y\right] dy \\ &= \frac{g(a)}{f'(a)}. \end{aligned}$$

3.3.- Dans toute la suite, on supposera que f est de classe C^2 et que M_f est contenu dans l'intérieur de K . Les minima de f sont alors des points critiques ($df = 0$).

On peut déterminer complètement l'équivalent de $I(t)$ quand ce sont des points critiques non dégénérés (rappelons qu'un point critique a de f est dit non dégénéré si la forme quadratique $\text{Hess}_a f$ sur $T_a N$ est non dégénérée. S'il s'agit d'un minimum elle est alors définie positive). Puisqu'un point critique non dégénéré est isolé dans M_f , il suffit de faire l'étude pour un minimum unique.

PROPOSITION.- (Formule asymptotique de Laplace). Soit K un compact d'intérieur non vide de $\mathbb{R}^n$ et soit f une fonction définie sur K , de classe C^2 , prenant sa plus petite valeur en un seul point a , intérieur à K , et qui est pour f un point critique non dégénéré. Soit g une fonction continue sur K telle que $g(a) > 0$. Alors

$$(16) \quad I(t) = \int_K \exp \left[-\frac{f(x)}{t}\right] g(x) dx \sim (2\pi t)^{n/2} g(a) [\det f''(a)]^{-1/2} \exp \left[-\frac{f(a)}{t}\right]$$

quand $t \rightarrow 0$.

Si f et g dépendent d'un paramètre λ parcourant un espace compact Λ de telle sorte que les applications $\lambda \rightarrow f_\lambda$ (resp. $\lambda \rightarrow g_\lambda$) à valeurs dans l'espace de Banach $C^2(K, \mathbb{R})$ (resp. $C^0(K, \mathbb{R})$) soient continues et si les hypothèses ci-dessus sont vérifiées pour chaque $\lambda \in \Lambda$, on a

$$(17) \quad I_\lambda(t) = \int_K \exp \left[-\frac{f_\lambda(x)}{t} \right] g_\lambda(x) dx \\ \sim (2\pi t)^{n/2} g_\lambda(a_\lambda) [\det f''_\lambda(a_\lambda)]^{-1/2} \exp \left[-\frac{f_\lambda(a_\lambda)}{t} \right]$$

quand $t \rightarrow 0$, uniformément sur Λ .

Démonstration. - Écrivons la formule de Taylor

$$f(a+u) = f(a) + \frac{1}{2} f''(a).u^2 + ||u||^2 \phi(u)$$

où $\phi(u) \rightarrow 0$ quand $u \rightarrow 0$. Comme $f''(a)$ est une forme quadratique définie positive, il existe un voisinage V de 0 dans $\mathbb{R}^n$ tel que $a+V \subset K$ et

$$||u||^2 |\phi(u)| \leq \frac{1}{4} f''(a).u^2$$

pour $u \in V$. On a d'après (13) et (14)

$$I(t) \sim g(a) \int_{a+V} \exp \left[-\frac{f(x)}{t} \right] dx \\ = g(a) \exp \left[-\frac{f(a)}{t} \right] t^{n/2} \int_{\frac{V}{\sqrt{t}}} \exp \left[-\frac{1}{2} f''(a).y^2 - ||y||^2 \phi(y \sqrt{t}) \right] dy.$$

En appliquant le théorème de Lebesgue à l'intégrale ci-dessus, on obtient

$$I(t) \sim t^{n/2} g(a) \exp \left[-\frac{f(a)}{t} \right] \int_{\mathbb{R}^n} \exp \left[-\frac{1}{2} f''(a).y^2 \right] dy$$

d'où l'on déduit (16).

Si f et g dépendent d'un paramètre λ , la continuité uniforme de f''_λ sur $\Lambda \times K$ permet de déduire de la formule de Taylor avec reste intégral

$$f_\lambda(a_\lambda + u) = f_\lambda(a_\lambda) + \int_0^1 (1-z) f''_\lambda(a_\lambda + zu).u^2 dz$$

le développement

$$f_{\lambda}(a_{\lambda} + u) = f_{\lambda}(a_{\lambda}) + \frac{1}{2} f''_{\lambda}(a_{\lambda}) \cdot u^2 + \|u\|^2 \phi_{\lambda}(u)$$

où $\phi_{\lambda}(u) \rightarrow 0$ quand $u \rightarrow 0$ uniformément sur Λ .

La fin de la démonstration est alors semblable à la précédente : il existe un voisinage V de 0 dans $\mathbb{R}^n$ tel que $a_{\lambda} + V \subset K$ et

$$|\phi_{\lambda}(u)| \leq \frac{1}{4} m \quad \text{pour } u \in V$$

où $m > 0$ est tel que $f''_{\lambda}(a_{\lambda}) \cdot u^2 \geq m \|u\|^2$ pour tout u .

On vérifie facilement que les formules (13) et (14) sont valables uniformément en $\lambda \in \Lambda$ et on applique le théorème de Lebesgue à l'intégrale de la fonction $\sup_{\lambda \in \Lambda} [|\exp[-\frac{1}{2} f''_{\lambda}(a_{\lambda}) \cdot y^2 - \|y\|^2 \phi_{\lambda}(y/\lambda)] - \exp[-\frac{1}{2} f''_{\lambda}(a_{\lambda}) \cdot y^2]|]$.

Remarque.— Les formules asymptotiques (16) et (17) sont encore valables sur une variété N munie d'une forme volume π à condition de remplacer $f''(a)$ par $\text{Hess}_a f$ et d'interpréter ainsi le déterminant du Hessien : si π s'exprime dans des coordonnées $u_1, \dots, u_n$ au voisinage de a par

$$\pi(dx) = h(u) du_1 \dots du_n$$

on a

$$\det_{\pi} \text{Hess}_a f = \det \left[\frac{\partial^2 f}{\partial u_i \partial u_j} (a) \right] h(a)^{-2}$$

expression qui ne dépend pas des coordonnées choisies.

3.4.— Si f possède des points critiques dégénérés, on ne peut pas en général déterminer un équivalent simple de $I(t)$. On dispose néanmoins d'une méthode qui fournit le résultat dans un grand nombre de cas (cf. Arnol'd [2]). Elle consiste essentiellement à effectuer dans l'intégrale un changement de variable $x = \phi(y)$ de manière à remplacer f par une fonction $f \circ \phi$ plus simple (polynômiale par exemple) pour laquelle on peut effectuer les calculs.

Montrons d'abord pour donner un exemple simple comment on peut utiliser cette méthode pour retrouver les résultats de la proposition 3.3 (on reprend les hypo-

thèses de 3.3).

Le lemme de Morse (cf. ch. 7 Lemme 8.4 A. Bellaïche) affirme l'existence d'un voisinage U de a , d'un voisinage V de 0 et d'un difféomorphisme $\Phi : V \rightarrow U$ tel que $\Phi(0) = a$ et

$$f[\Phi(v)] = f(a) + \sum_{i=1}^n v_i^2 \quad \text{pour } v \in V.$$

On a alors

$$\begin{aligned} I(t) &\sim \exp \left[-\frac{f(a)}{t} \right] \int_V \exp \left[-\frac{\sum v_i^2}{t} \right] g[\Phi(v)] [\det \Phi'(v)]^{-1} dv_1 \dots dv_n \\ &\sim (\pi t)^{n/2} g(a) [\det \Phi'(0)]^{-1} \exp \left[-\frac{f(a)}{t} \right]. \end{aligned}$$

Remarquons que si Q est la matrice de la forme quadratique $\sum v_i^2$ on a au point a ,

$$\frac{1}{2} \left(\frac{\partial^2 f}{\partial u_i \partial u_j} \right) = {}^t \Phi'(0) Q \Phi'(0), \text{ de telle sorte que } \det \Phi'(0) = 2^{-n/2} [\det f''(a)]^{1/2}.$$

On peut également utiliser le lemme de Morse avec paramètre (cf. ch. 7, Lemme 8.5) pour démontrer (17).

Soit maintenant $a \in K$ un minimum dégénéré (non nécessairement isolé) de f et soit $k > 0$ le corang de $\text{Hess}_a f$ (i.e. par définition la codimension du noyau de cette forme quadratique). Le lemme de Gromoll-Meyer (voir Arnol'd [2] lemme 4.1) affirme l'existence d'un voisinage V de a , d'un voisinage $U = U' \times U''$ de 0 dans $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ et d'un difféomorphisme $\Phi : U \rightarrow V$ tel que $\Phi(0) = a$ et

$$f[\Phi(u)] = f(a) + \phi(u_1, \dots, u_k) + u_{k+1}^2 + \dots + u_n^2$$

$$\text{avec } \phi(u_1, \dots, u_k) = o(u_1^2 + \dots + u_k^2).$$

On a alors

$$\begin{aligned} I_U(t) &= \int_U \exp \left[-\frac{f(x)}{t} \right] g(x) \pi(dx) \\ &= \exp \left[-\frac{f(a)}{t} \right] \int_{U'} \exp \left[-\frac{\phi(u')}{t} \right] du' \int_{U''} \exp \left[-\frac{|u''|^2}{t} \right] g(u', u'') p(u', u'') du'' \end{aligned}$$

lorsque dans les coordonnées $u_1, \dots, u_n$ la forme volume $\pi(dx)$ s'exprime sous la forme $p(u) du$.

D'après la formule asymptotique de Laplace (17), on a

$$(18) \quad I_U(t) \sim \exp \left[-\frac{f(a)}{t} \right] (\pi t)^{\frac{n-k}{2}} \int_{U'} \exp \left[-\frac{\phi(u')}{t} \right] g(u', 0) p(u', 0) du'.$$

Le calcul de l'équivalent de $I_U(t)$ est ainsi ramené à celui d'une intégrale analogue où le développement de Taylor de f ne possède plus de termes d'ordre 2 : les termes d'ordre 2 ont été éliminés par intégration dans les variables correspondantes et fournissent à l'équivalent une contribution en $t^{\frac{n-k}{2}}$.

Remarque.— Puisque pour $|u'|$ assez petit, on a $0 \leq \phi(u') \leq |u'|^2$, on voit que

$$\text{const. } t^{k/2} \leq \int_{U'} \exp \left[-\frac{\phi(u')}{t} \right] g(u', 0) p(u', 0) du' \leq \text{Const.}$$

pour t assez petit et que par conséquent

$$(19) \quad \text{Const. } t^{n/2} \exp \left[-\frac{f(a)}{t} \right] \leq I_U(t) \leq \text{Const. } t^{\frac{n-k}{2}} \exp \left[-\frac{f(a)}{t} \right].$$

Si l'on veut obtenir des résultats plus précis que (19), il faut faire des hypothèses sur les termes suivants du développement de Taylor de f . On supposera dans la suite que f est une fonction de classe C^∞ . On peut alors choisir le difféomorphisme ϕ ci-dessus de classe C^∞ et la fonction ϕ est elle aussi C^∞ .

3.5.— Supposons que $\text{corang Hess}_a f = 1$. La fonction ϕ est alors fonction d'une variable. Si ϕ n'a pas toutes ses dérivées nulles à l'origine, soit $2p$ l'ordre de la première dérivée non nulle de ϕ en 0. On peut alors par un nouveau changement de variable mettre ϕ sous la forme $\phi(u) = u^{2p}$ et f s'écrit

$$f(x) = x_1^{2p} + x_2^2 + \dots + x_n^2.$$

On dit alors que f possède en a un point critique dégénéré de type A_{2p-1} (terminologie d'Arnol'd). Remarquons qu'un point critique de type A_{2p-1} est nécessairement isolé. On a

$$\begin{aligned} \int_{U'} \exp \left[-\frac{\phi(u')}{t} \right] g(u', 0) p(u', 0) du' &\sim \text{Const. } t^{1/2p} \int_{t^{-1/2p} U'} e^{-y^{2p}} dy \\ &\sim \text{Const. } t^{1/2p}. \end{aligned}$$

Donc en utilisant (18), on a

$$(20) \quad I_U(t) \sim \text{Const.} \cdot t^{\frac{n-1}{2} + \frac{1}{2p}} \exp \left[-\frac{f(a)}{t} \right] \quad \text{quand } t \rightarrow 0.$$

La formule (20) montre a posteriori que l'entier p ne dépend que de la fonction p .

Si toutes les dérivées de ϕ sont nulles à l'origine, il n'y a aucune classification possible. On a cependant pour tout entier p ,

$$(21) \quad I_U(t) \gg t^{\frac{n-1}{2} + \frac{1}{2p}} \exp \left[-\frac{f(a)}{t} \right] \quad \text{quand } t \rightarrow 0.$$

En effet pour $|u'|$ assez petit, on a $\phi(u') \leq u'^{2p}$ pour tout entier p . Donc en diminuant éventuellement la taille de U' , on obtient

$$\begin{aligned} \int_{U'} \exp \left[-\frac{\phi(u')}{t} \right] g(u', 0) p(u', 0) du' &\geq \text{Const.} \int_{U'} \exp \left[-\frac{u'^{2p}}{t} \right] du' \\ &\geq \text{Const.} \cdot t^{1/2p} \end{aligned}$$

pour t assez petit.

3.6.- Si $\text{corang } f''(a) = 2$, la classification est beaucoup plus compliquée. Donnons seulement un exemple où un phénomène intéressant se produit : l'apparition d'un terme en $\text{Log } \frac{1}{t}$ dans l'équivalent.

PROPOSITION. - Si dans des coordonnées u_i au voisinage de a telles que $u_i(a) = 0$, on peut écrire

$$f(u) = f(a) + u_1^4 + (u_1 u_2)^2 + u_2^{2p} + \sum_{k=3}^n u_k^2 \quad *$$

alors si $p = 2$

$$(22) \quad I(t) \sim \text{Const.} \cdot t^{\frac{n-1}{2}} e^{-f(a)/t}$$

et si $p > 2$

(*) C'est la classe X_{2p+5} dans la terminologie d'Arnol'd [1].

$$(23) \quad I(t) \sim \text{Const. } t^{\frac{n-1}{2}} \text{Log } \frac{1}{t} e^{-f(a)/t}$$

quand $t \rightarrow 0$.

Démonstration.— On a d'après (18) et (14)

$$I(t) \sim \text{Const. } e^{-f(a)/t} t^{\frac{n-2}{2}} \int_W \exp \left[-\frac{x^4 + x^2 y^2 + y^{2p}}{t} \right] dx dy$$

où W est un voisinage compact de 0 dans $\mathbb{R}^2$.

Si $p = 2$ le changement de variable $x = t^{1/4}u$, $y = t^{1/4}v$ montre que

$$\int_W \exp \left[-\frac{x^4 + x^2 y^2 + y^4}{t} \right] dx dy \sim t^{1/2} \int_{\mathbb{R}^2} \exp \left[-u^4 - u^2 v^2 - v^4 \right] du dv$$

Si $p > 2$, posons

$$J(t) = \int_{\mathbb{R}^2} \exp \left[-\frac{x^4 + x^2 y^2 + y^{2p}}{t} \right] dx dy.$$

En faisant le changement de variable $x = t^{1/4}u$, $y = t^{1/4}v$, on obtient

$$J(t) = 4t^{1/2} \int_0^\infty \exp \left[-v^{2p} t^{p/2-1} \right] \phi(v) dv$$

$$\text{où } \phi(v) = \int_0^\infty \exp \left[-u^4 - u^2 v^2 \right] du.$$

D'après la formule asymptotique de Laplace, on a quand $v \rightarrow \infty$

$$\int_0^A \exp \left(-u^4 - u^2 v^2 \right) du \sim \frac{1}{2} \frac{\sqrt{2\pi}}{v}$$

et

$$\int_A^\infty \exp \left(-u^4 - u^2 v^2 \right) du \leq C \exp \left(-A^2 v^2 \right)$$

donc $\phi(v) \sim \text{Const. } \frac{1}{v}$ quand $v \rightarrow \infty$. Soit $\varepsilon > 0$. Pour A assez grand on a

$$\begin{aligned} (1 - \varepsilon) \int_A^\infty \exp \left[-v^{2p} t^{p/2-1} \right] \frac{dv}{v} &\leq \int_A^\infty \exp \left[-v^{2p} t^{p/2-1} \right] \phi(v) dv \\ &\leq (1 + \varepsilon) \int_A^\infty \exp \left[-v^{2p} t^{p/2-1} \right] \frac{dv}{v} \end{aligned}$$

$$\text{Or } \int_A^\infty \exp[-v^{2p} t^{p/2-1}] \frac{dv}{v} = \int_{t^{p/2-1} A^{2p}}^\infty e^{-u} \frac{du}{u} \sim \text{Const.} \cdot \text{Log } \frac{1}{t}$$

tandis que

$$\int_0^A \exp[-v^{2p} t^{p/2-1}] \phi(v) dv \leq \int_0^A \phi(v) dv = O(1)$$

donc $J(t) \sim \text{Const.} \cdot t^{1/2} \text{Log } \frac{1}{t}$ quand $t \rightarrow 0$.

La formule (23) en résulte si on peut montrer que

$$\int_W \exp\left[-\frac{x^4 + x^2 y^2 + y^{2p}}{t}\right] dx dy \sim \int_{\mathbb{R}^2} \exp\left[-\frac{x^4 + x^2 y^2 + y^{2p}}{t}\right] dx dy$$

Or pour $y \geq 1$, on a $x^4 + x^2 y^2 + y^{2p} \geq \frac{1}{2} (x^2 + y^2)^2$ et pour $x \geq 1$, on a $x^4 + x^2 y^2 + y^{2p} \geq x^2 + y^2$ donc $x^4 + x^2 y^2 + y^{2p} \geq r^2$ pour $r \geq \sqrt{2}$ ($r^2 = x^2 + y^2$).

On peut supposer que W contient le disque $r \leq \sqrt{2}$. Donc

$$\begin{aligned} \int_{\mathbb{R}^2 - W} \exp\left[-\frac{x^4 + x^2 y^2 + y^{2p}}{t}\right] dx dy &\leq \int_2^\infty \exp\left(-\frac{r^2}{t}\right) 2\pi r dr = \pi t e^{-2/t} \\ &= o\left(\int_{\mathbb{R}^2} \exp\left[-\frac{x^4 + x^2 y^2 + y^{2p}}{t}\right] dx dy\right). \end{aligned}$$

3.7.- Signalons enfin que si f est analytique et g est C^∞ , on peut montrer que

$$(24) \quad I(t) \sim \text{Const.} \cdot t^{\alpha(\text{Log } \frac{1}{t})^q} \cdot e^{-\frac{m}{t}}$$

où m est le minimum de f (on rappelle que ce minimum est atteint dans l'intérieur de K) et où α est un rationnel compris entre 0 et $\frac{n}{2}$ et où $q \in \mathbb{N}$. La démonstration de Malgrange ([8] th. 1.1, pp. 406 et 425-426) concerne des intégrales complexes de la forme

$$J(\tau) = \int e^{i\tau f(x)} g(x) dx$$

mais elle se transpose facilement.

Ce résultat est en défaut si on ne suppose plus f analytique :

3.8. Exemple.— Soit $f(x) = \exp(-\frac{1}{x^2})$ et $I(t) = \int_{-1}^1 \exp[-\frac{f(x)}{t}] dx$. Alors $I(t)$ n'est équivalent à aucun $Ct^\alpha (\log \frac{1}{t})^\beta$, $\beta \in \mathbb{N}$ quand $t \rightarrow 0$. En effet, d'une part $\lim_{t \rightarrow 0} I(t) = 0$. D'autre part $I(t)$ est infiniment grand devant t^α pour tout $\alpha > 0$ quand $t \rightarrow 0$. En effet, puisque f est plate à l'origine, pour tout entier n , il existe $\varepsilon > 0$ tel que $f(x) \leq |x|^n$ pour $|x| < \varepsilon$. On a alors

$$I(t) \geq \int_{-\varepsilon}^{\varepsilon} \exp[-\frac{|x|^n}{t}] dx \geq \text{Const. } t^{1/n}$$

Donc si on avait $I(t) \sim Ct^\alpha (\log \frac{1}{t})^\beta$ avec $\alpha \geq 0$, $\beta \in \mathbb{N}$, on aurait $\alpha = 0$ ce qui est impossible puisque $\lim_{t \rightarrow 0} I(t) = 0$.

4.- Equivalent de $p(t, x, y)$ quand $t \rightarrow 0$

On note $C(M)$ le cut-locus de M . Lorsque la variété M est complète, $C(M)$ est l'ensemble des couples $(x, y) \in M \times M$ tels que x et y sont joints par plusieurs géodésiques minimisantes ou conjugués le long d'une géodésique minimisante les joignant. Si M n'est pas complète, la définition de $C(M)$ est plus délicate (cf. ch. 7, §7, A. Bellaïche).

On note $V(M)$ l'ensemble des couples $(x, y) \in M \times M$ de points vérifiant la condition (c) : $d(x, y) < \sup [d(x, \infty), d(y, \infty)]$, et on pose $V(M) = M \times M$ si M est complète.

4.1.- THÉOREME.— (cf. [10], th. 3.1). Soit M une variété riemannienne de classe C^4 et $\mathcal{P}$ une diffusion sur M de générateur $\frac{1}{2}\Delta + Y$ où Δ est l'équation de Laplace-Beltrami et Y un champ de vecteurs. On note $p(t, x, y)$ la densité du semi-groupe de transition de $\mathcal{P}$ par rapport au volume riemannien π . Si $(x, y) \in V(M)$ et $(x, y) \notin C(M)$, on a

$$(25) \quad p(t, x, y) \sim (2\pi t)^{-m/2} H(x, y) \exp[-\frac{d^2(x, y)}{2t} + A(x, y)]$$

quand $t \rightarrow 0$, où $H(x, y) = \det [d \exp_x(\dot{\gamma}(0))]^{-1/2}$, $A(x, y) = \int_0^1 \langle \ddot{\gamma}, \dot{\gamma} \rangle ds$, γ étant l'unique géodésique de longueur minimale joignant x à y .

De plus, l'équivalent (25) est uniforme sur tout compact de $V(M)$ ne rencontrant pas $C(M)$.

Démonstration.— Supposons d'abord M complète pour d et soit $(x_0, y_0) \notin C(M)$. Il existe d'après la proposition 2.6, un entier n , $h > c$ et des voisinages B de x_0 , B' de y_0

$$p(t, x, y) \sim \left(\frac{2\pi t}{n+1} \right)^{-\frac{m(n+1)}{2}} \int_{M_{x_0 y_0}^{n, h}} \exp \left[-\frac{E(\gamma)}{t} + A(\gamma) \right] H(\gamma) \pi(dy_1) \dots \pi(dy_n)$$

Comme $C(M)$ est fermé, on peut supposer que $B \times B'$ ne rencontre pas $C(M)$. Soient $x \in B$ et $y \in B'$. L'hypothèse $(x, y) \notin C(M)$ signifie que la fonction $\gamma \mapsto E(\gamma)$ a un minimum unique non dégénéré sur la variété Ω_{xy} des chemins paramétrés de 0 à 1 d'origine x et d'extrémité y , ainsi que sur l'espace $\Omega_{xy}^{n, h}$ des géodésiques brisées (cf. ch. 7, § 7 et 8, A. Bellaïche) et que ce minimum est l'unique géodésique γ_{xy} de longueur minimale joignant x à y . On en déduit au moyen de l'identification entre $M_{xy}^{n, h}$ et $\Omega_{xy}^{n, h}$ et de l'inclusion (12) $M_{xy}^{n, h/2} \subset M_{x_0 y_0}^{n, h} \subset M_{xy}^{n, 2h}$ que le point $\gamma_{xy}(\frac{i}{n+1})_{i=1, \dots, n}$ est le minimum unique non dégénéré de E sur $M_{x_0 y_0}^{n, h}$. Comme la fonction $d^2(x, y)$ est de classe C^2 dans le complémentaire de $C(M)$, la fonction $E[\gamma(x, y_1, \dots, y_n, y)]$ est de classe C^2 dans $B \times M_{x_0 y_0}^{n, h} \times B'$ et on peut appliquer la formule asymptotique de Laplace (17), en posant $D_{xy} = (\det \text{Hess } E(\gamma_{xy}))^{-1/2}$,

$$\int_{M_{x_0 y_0}^{n, h}} \exp \left[-\frac{E(\gamma)}{t} + A(\gamma) \right] H(\gamma) \pi(dy_1) \dots \pi(dy_n) \sim$$

$$(2\pi t)^{mn/2} D_{xy} H(x, y_1) \dots H(y_n, y) \exp \left[-\frac{d^2(x, y)}{2t} + A(x, y_1) + \dots + A(y_n, y) \right]$$

lorsque $t \rightarrow 0$, uniformément sur $B \times B'$. L'équivalent (25) en résulte puisque

$$(n+1)^{m(n+1)/2} H(x, y_1) \dots H(y_n, y) [\det \text{Hess } E(\gamma_{xy})]^{-1/2} = H(x, y)$$

(voir ch. 7, prop. 8.12. On rappelle que $H(x, y) = \theta(x, y)^{-1/2}$).

Si M n'est pas complète, on a d'après la proposition 2.4, $p(t, x, y) \sim p_N(t, x, y)$ uniformément sur $B \times B'$ où B et B' sont des voisinages de x_0 et y_0 et N une variété complète. On en déduit (25) puisque les fonctions $H(x, y)$, $d(x, y)$, $A(x, y)$ associées à M et à N sont égales, d'après 2.4.

Remarques. 1) Puisque pour tout x , $C(x)$ est de mesure nulle, l'hypothèse $(x, y) \notin C(M)$ est vérifiée par presque tout couple $(x, y) \in M \times M$.

2) Si x et y sont joints par un nombre fini de géodésiques minimisantes $\gamma_1, \dots, \gamma_\ell$ et non conjugués le long de chacune d'elles, il suffit d'après (13), pour

obtenir l'équivalent de $p(t, x, y)$ d'ajouter les contributions de voisinages disjoints U_i de chaque γ_i . On en déduit que

$$p(t, x, y) \sim (2\pi t)^{-m/2} \sum_i H(\gamma_i) \exp[A(\gamma_i)] \exp\left[-\frac{d^2(x, y)}{2t}\right]$$

$$\text{où } H(\gamma_i) = \det [d \exp_x \dot{\gamma}_i(0)]^{-1/2}, \quad A(\gamma_i) = \int_0^1 \langle Y, \dot{\gamma}_i \rangle dt.$$

4.2. On suppose maintenant que les coefficients $a_{ij}(x)$ de L sont de classe C^∞ . La métrique riemannienne est alors de classe C^∞ . Soient x et y deux points de M vérifiant la condition (C) et soient n assez grand et h assez petit pour que (11) soit vérifiée en (x, y) .

4.3. THÉORÈME.— ([10], subcase 3.2.a). Avec les notations et hypothèses 4.2, supposons que x et y soient joints par une seule géodésique minimisante γ_0 et conjugués le long de γ_0 avec multiplicité 1. Si la singularité en γ_0 de la restriction de E à $\Omega_{xy}^{n, h}$ est de type A_{2p-1} c'est-à-dire si dans des coordonnées $u_1, \dots, u_{mn}$ au voisinage de γ_0 , on peut écrire

$$E(\gamma) = E(\gamma_0) + u_1^{2p} + \sum_{k=2}^{mn} u_k^2$$

on a

$$(26) \quad p(t, x, y) \sim \text{Const.} \cdot t^{\frac{1}{2p} - \frac{m+1}{2}} \exp\left[-\frac{d^2(x, y)}{2t}\right]$$

quand $t \rightarrow 0$

Démonstration.— Il suffit d'appliquer (11) et (20).

4.4. THÉORÈME.— ([10], subcase 3.2.b). Avec les notations et hypothèses 4.2, supposons que x et y soient joints par une géodésique minimale γ_0 et conjugués le long de γ_0 avec multiplicité 2. Si la singularité en γ_0 de la restriction de E à $\Omega_{xy}^{n, h}$ est de type X_{2p+5} c'est-à-dire si dans des coordonnées $u_1, \dots, u_{mn}$ au voisinage de γ_0 on a

$$E(\gamma) = E(\gamma_0) + u_1^4 + u_1^2 u_2^2 + u_2^{2p} + \sum_{k=3}^{mn} u_k^2$$

alors si $p = 2$

$$(27) \quad p(t, x, y) \sim \text{Const.} \cdot t^{-\frac{m+1}{2}} \exp\left[-\frac{d^2(x, y)}{2t}\right]$$

et si $p > 2$

$$(28) \quad p(t, x, y) \sim \text{Const. } t^{-\frac{m+1}{2}} \text{Log}\left(\frac{1}{t}\right) \exp\left[-\frac{d^2(x, y)}{2t}\right]$$

Démonstration.— Il suffit d'appliquer (22) et (23).

4.5. Remarque.— Pour chacun des équivalents (26) (27) (28) on peut effectivement construire, étant donné une variété M de dimension ≥ 2 pour (26), ≥ 3 pour (27) et (28) et deux points distincts x et y de M , une diffusion sur M dont la densité de transition $p(t, x, y)$ satisfasse (26) ou (27) ou (28). Pour le prouver, on construit une métrique complète sur M telle que x et y soient joints par une seule géodésique minimisante et que la singularité de E au voisinage de cette géodésique soit du type voulu A_{2p-1} , X_9 ou X_{2p+5} ($p > 2$) (cf. Prop. 8.10, ch. 7).

4.6. Remarque.— On peut montrer (au moins dans le cas où M est compacte) que si la dimension de M est ≤ 10 , il existe dans l'espace des métriques riemanniennes C^∞ sur M un ouvert dense U tel que si la métrique g appartient à U on ait la propriété suivante :

pour tout $x \in M$ et tout $y \in M$, les géodésiques minimisantes joignant x à y sont en nombre fini $\gamma_1, \dots, \gamma_r$, et pour $i = 1, \dots, r$, la géodésique γ_i est soit un minimum non dégénéré de E , soit un minimum dégénéré avec l'une des singularités A_{2p-1} ($2 \leq p \leq 5$), X_{2p+5} ($p = 2, 3$) (cf. Arnol'd [1] cor. 3.3.9 et Buchner [4]).

L'équivalent de $p(t, x, y)$ est alors somme de contributions du type (25), (26), (27) ou (28) correspondant à chacune des géodésiques γ_i . On peut donc dire que l'équivalent de $p(t, x, y)$ est déterminé génériquement en dimension ≤ 10 .

4.7. THÉORÈME.— Avec les hypothèses de 4.2, si les géodésiques minimisantes joignant x à y forment une sous-variété Γ de dimension k de la variété Ω_{xy} des chemins joignant x à y (cf. ch. 7, 2.6), ou de $\Omega_{xy}^{n, h}$ et si x et y sont conjugués avec multiplicité k le long de chaque $\gamma \in \Gamma$, on a

$$(29) \quad p(t, x, y) \sim \text{Const. } t^{-\frac{m}{2} - \frac{k}{2}} \exp\left[-\frac{d^2(x, y)}{2t}\right]$$

quand $t \rightarrow 0$.

Démonstration.— On se ramène au cas où M est complète. Par hypothèse Γ est une sous-variété de Ω_{xy} , et on a $\Gamma \subset \Omega_{xy}^{n, h}$. Montrons d'abord que Γ est une sous-

variété compacte de $\Omega_{xy}^{n,h}$. On sait que l'application qui à une géodésique γ associe $(\gamma(0), \dot{\gamma}(0))$ est un difféomorphisme de la variété des géodésiques sur TM . Il suffit donc de montrer que l'ensemble

$$\Gamma_x = \{\dot{\gamma}(0) \mid \gamma \in \Gamma\}$$

est compact. Or il est borné puisque $\|\dot{\gamma}(0)\| = d(x, y)$ et fermé puisque si $\dot{\gamma}_n(0) \rightarrow X \in T_x M$, on a $\lim \exp_x(\dot{\gamma}_n(0)) = y$ et la géodésique $\gamma(t) = \exp_x(tX)$ ($0 \leq t \leq 1$) appartient à Γ .

Soit $\gamma_0 \in \Gamma$. Puisque Γ est une sous-variété compacte de $\Omega_{xy}^{n,h}$, on peut trouver un voisinage U de γ_0 dans $\Omega_{xy}^{n,h}$ et un voisinage de 0 dans $\mathbb{R}^{mn}$ de la forme $U' \times U''$ avec $U' \subset \mathbb{R}^k$ et $U'' \subset \mathbb{R}^{mn-k}$, et un difféomorphisme $\phi : U' \times U'' \rightarrow U$ tels que

$$\phi(U' \times \{0\}) = \Gamma \cap U \quad \text{et} \quad \phi(0, 0) = \gamma_0.$$

La fonction $E \circ \phi$ au voisinage d'un point $(u, 0)$ ($u \in U'$) peut être considérée comme fonction de $u'' = (u_{k+1}, \dots, u_{mn}) \in U''$ paramétrée par $u' = (u_1, \dots, u_k) \in U'$. On a

$$\frac{\partial^2 E}{\partial u^2}(u', 0) = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & \frac{\partial^2 E}{\partial u''^2}(u', 0) \end{pmatrix}$$

L'hypothèse que x et y sont conjugués d'ordre k le long de toute géodésique $\gamma \in \Gamma$ assure que pour tout $u \in U'$, la matrice $\frac{\partial^2 E}{\partial u^2}(u', 0)$ est de rang $mn-k$. Par conséquent la matrice $\frac{\partial^2 E}{\partial u''^2}(u', 0)$ est non dégénérée pour tout $u' \in U'$. On peut alors appliquer le lemme de Morse avec paramètres (cf. ch. 7, lemme 8.5, A. Bellaïche). Celui-ci affirme qu'il existe un changement de coordonnées $(u', u'') \rightarrow (v', v'')$ défini sur un voisinage de $(0, 0)$ dans $U' \times U''$, avec $v' = u'$, $v'' = \phi(u', u'')$ tel que dans les nouvelles coordonnées ainsi obtenues, E prenne la forme

$$E(v', v'') = \frac{1}{2} d^2(x, y) + \sum_{k+1}^{mn} v_i^2$$

On peut supposer en réduisant éventuellement la taille de ce voisinage de $(0, 0)$ dans $U' \times U''$ que v' décrit un voisinage compact V' de 0 dans $\mathbb{R}^k$ et v'' un voisinage compact V'' de 0 dans $\mathbb{R}^{mn-k}$. On a alors

$$\int_{V' \times V''} \exp \left[-\frac{E(v', v'')}{t} + A(v', v'') \right] H(v', v'') \, dv' \, dv'' \\ \sim \text{const.} \, t^{\frac{mn-k}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

Puisque Γ est compact, en faisant un recouvrement fini de Γ , on voit qu'il existe un voisinage ouvert Ω de Γ dans $\Omega_{xy}^{n,h}$ tel que

$$\int_{\Omega} \exp \left[-\frac{E(\gamma)}{t} + A(\gamma) \right] H(\gamma) \, \pi_n(d\gamma) \sim \text{const.} \, t^{\frac{mn-k}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

D'après la proposition 2.6 et (13) on aura donc

$$p(t, x, y) \sim \text{const.} \, t^{-\frac{m}{2} - \frac{h}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

quand $t \rightarrow 0$.

4.8. THÉORÈME.— Soient x et y vérifiant la condition (C). Pour t assez petit on a

$$(30) \quad \text{const.} \, t^{-\frac{m}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right] \leq p(t, x, y) \leq \text{const.} \, t^{-m + \frac{1}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

Démonstration.— D'après 2.4 on peut supposer M complète. Soit γ_0 une géodésique minimale joignant x à y et supposons que x et y sont conjugués avec multiplicité k le long de γ_0 (k entier ≥ 0). D'après le lemme de Gromoll-Meyer, il existe un voisinage U de γ_0 dans $\Omega_{xy}^{n,h}$ et des coordonnées $u_1, \dots, u_{mn}$ sur U , nulles en γ_0 , dans lesquelles l'énergie s'exprime sous la forme

$$E(\gamma) = \frac{1}{2} d^2(x, y) + \phi(u_1, \dots, u_k) + \sum_{k+1}^{mn} u_i^2$$

avec $\phi(u_1, \dots, u_k) = 0(u_1^2 + \dots + u_k^2)$. Il est clair que $\phi(u_1, \dots, u_k) \geq 0$ pour $|u_1^2 + \dots + u_k^2|$ assez petit. On a alors d'après (19)

$$\text{const.} \, t^{\frac{mn}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right] \leq \int_U \exp \left[-\frac{E(\gamma)}{t} + A(\gamma) \right] H(\gamma) \, \pi_n(d\gamma) \\ \leq \text{const.} \, t^{\frac{mn-k}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right]$$

Soit Γ l'ensemble des géodésiques minimales joignant x à y (Γ est un compact de $\Omega_{xy}^{n,h}$) et soit k l'ordre maximal de conjugaison de x et y le long de γ lorsque γ par-

court Γ . On sait que $k \leq m-1$. En recouvrant Γ par un nombre fini de voisinages U_i de γ_i et en appliquant la proposition 2.6, et (13) on obtient (30).

Remarque.— Lorsque x et y sont deux pôles opposés d'une sphère de dimension m , il existe une variété de dimension $m-1$ de géodésiques minimales joignant x à y et l'ordre de conjugaison de x et y le long de chacune d'elles est égal à $m-1$.

On a d'après le théorème 4.7.

$$p(t, x, y) \sim \text{const. } t^{-m + \frac{1}{2}} \exp \left[- \frac{d^2(x, y)}{2t} \right]$$

et on voit qu'une des bornes extrêmes permises par le théorème 4.8 est atteinte (l'autre borne est atteinte dans le cas du théorème 4.1.).

4.9.— Si la variété M est analytique et si sa métrique riemannienne est analytique, d'après (24) on a toujours un équivalent du type

$$p(t, x, y) \sim \text{const. } t^{-\frac{m}{2} - \frac{\alpha}{2}} |\log t|^\beta \exp \left[- \frac{d^2(x, y)}{2t} \right]$$

($0 \leq \alpha \leq m-1$, et $\beta \in \mathbb{N}$).

Ce n'est pas le cas si M et sa métrique sont supposés seulement de classe C^∞ . En effet si la dimension de M est ≥ 2 , on peut trouver d'après la proposition 8.10 du chapitre 7 une métrique riemannienne complète g sur M et pour cette métrique une géodésique minimale γ_0 joignant x à y telle que pour n assez grand et h assez petit on ait dans des coordonnées $u_1, \dots, u_{mn}$ sur $\Omega_{xy}^{n, h}$ au voisinage de γ_0

$$E(\gamma) = \exp \left(- \frac{1}{u_1^2} \right) + \frac{1}{u_2^2} + \dots + \frac{1}{u_{mn}^2} + E(\gamma_0)$$

et on en déduit d'après 3.8 que $p(t, x, y)$ n'est équivalent à aucun

$$Ct^\alpha (\log \frac{1}{t})^\beta \exp \left[- \frac{d^2(x, y)}{2t} \right].$$

5.- Applications

Dans cette partie, la variété M est supposée complète pour la distance riemannienne d . Soit x un point de M et A un domaine fermé à bord C^2 tel que $x \notin A$. Nous allons chercher un équivalent de $P_x(X_t \in A)$ quand $t \rightarrow 0$. Pour simplifier, nous supposons que le générateur infinitésimal de la diffusion est $L = \frac{1}{2} \Delta$, ce qui dans l'équivalent

de $P_x(X_t \in A)$ ne modifie que la constante.

5.1. LEMME.— Soit A un domaine fermé de M à bord C^2 et soit $x \notin A$. Alors si $y_0 \in \partial A$ vérifie $d(x, y_0) = d(x, A)$, y_0 n'appartient pas au cut-locus de x .

Démonstration.— Soit γ une géodésique minimisante joignant x à y_0 . Alors γ est orthogonal à ∂A en y_0 (c'est une conséquence de la formule de la variation première). Donc γ est unique (deux géodésiques de même longueur, de même extrémité et ayant même vecteur tangent en leur extrémité commune coïncident).

Si y_0 était conjugué de x le long de γ , tout point $y \in A$ situé sur le prolongement de γ serait joint à x par une géodésique γ_1 plus courte que γ car γ n'est plus minimisante après le premier point conjugué. (cf. Milnor [9], théorème de l'indice). Si y est assez proche de y_0 , la géodésique γ entre y_0 et y réalise le minimum de la distance de y à ∂A . On a alors en notant y_1 un point d'intersection de γ_1 et ∂A ,

$$d(x, y_0) + d(y_0, y) > d(x, y_1) + d(y_1, y)$$

tandis que $d(y_0, y) < d(y_1, y)$. Donc $d(x, y_0) > d(x, y_1)$ ce qui contredit l'hypothèse.

5.2. LEMME.— (corollaire de la proposition 4.4, ch. 8, R. Azencott). Soit K un compact de M et soit A une partie mesurable de M telle que $d(K, A) > 0$. Il existe $C, T, N > 0$ tels que

$$P_x(X_t \in A) \leq C t^{-N} \exp \left[-\frac{d^2(x, A)}{2t} \right]$$

pour $t \leq T$ et $x \in K$.

Démonstration.— On choisit un volume riemannien $\tilde{\pi}$ tel que $\tilde{\pi}(M) < +\infty$ et on note $\tilde{p}$ la densité par rapport à $\tilde{\pi}$ du semi-groupe de transition de la diffusion $\mathcal{P}$. On a alors d'après la proposition 4.4 (ch. 8, R. Azencott)

$$P_x(X_t \in A) = \int_A \tilde{p}(t, x, y) \tilde{\pi}(dy) \leq C t^{-N} \tilde{\pi}(A) \exp \left[-\frac{d^2(x, A)}{2t} \right]$$

5.3. LEMME.— (Hypothèses de 5.1). Soit $\varepsilon > 0$ et soit B l'ensemble des $y \in A$ tels que $d(x, y) \leq d(x, A) + \varepsilon$. On a

$$P_x(X_t \in A) \sim P_x(X_t \in B)$$

quand $t \rightarrow 0$. De plus pour tout $h > 0$ on a pour t assez petit

$$P_x(X_t \in A) \geq \exp \left[-\frac{d^2(x, A)}{2t} - \frac{h}{t} \right]$$

Démonstration.— D'après le lemme 5.2, on a

$$P_x(X_t \in A - B) \leq C t^{-N} \exp \left[-\frac{(d(x, A) + \varepsilon)^2}{2t} \right]$$

pour t assez petit. Soit B' l'ensemble des points y de A tels que $d(x, y) \leq d(x, A) + \frac{\varepsilon}{2}$. D'après le lemme 2.1, on a pour tout $\chi > 0$

$$\begin{aligned} P_x(X_t \in B) &\geq P_x(X_t \in B') = \int_{B'} p(t, x, y) \pi(dy) \\ &\geq \exp \left[-\frac{(d(x, A) + \frac{\varepsilon}{2})^2}{2t} - \frac{\chi}{t} \right] \end{aligned}$$

pour t assez petit car $\pi(B') > 0$. Si on choisit χ assez petit, $P_x(X_t \in A - B)$ est négligeable devant $P_x(X_t \in B)$ ce qui démontre le lemme 5.3.

5.4 LEMME.— (Hypothèses de 5.1.) Si ε est assez petit, l'ensemble B des points $y \in A$ tels que $d(x, y) < d(x, A) + \varepsilon$ ne rencontre pas le cut-locus de x .

Démonstration.— Soit A_0 l'ensemble des points $y \in A$ tels que $d(x, y) = d(x, A)$. D'après le lemme 5.1, A_0 ne rencontre pas le cut-locus $C_x(M)$ de x . Comme A_0 est compact et $C_x(M)$ fermé, il existe un voisinage de A_0 ne rencontrant pas $C_x(M)$. Pour ε assez petit, B est contenu dans ce voisinage.

5.5. PROPOSITION.— Soit A un domaine fermé de M à bord C^2 et soit $x \notin A$. S'il existe un unique point $y_0 \in \partial A$ tel que $d(x, y_0) = d(x, A)$ et si la fonction F définie sur ∂A par $F(y) = \frac{1}{2} d^2(x, y)$ admet en y_0 un minimum non dégénéré, on a

$$(31) \quad P_x(X_t \in A) \sim \sqrt{\frac{t}{2\pi}} \frac{H(x, y_0)}{d(x, A)} [\det \text{Hess } F(y_0)]^{-1/2} \exp \left[-\frac{d^2(x, A)}{2t} \right]$$

quand $t \rightarrow 0$.

Remarque.— Si $M = \mathbb{R}^m$ dire que F admet en y_0 un minimum non dégénéré revient à dire que x n'est pas point focal de l'hypersurface ∂A en y_0 (cf. Milnor [9], p.36).

Démonstration.— Soit B l'ensemble des points $y \in A$ tels que $d(x, y) \leq d(x, A) + \varepsilon$ (on choisit ε assez petit pour que le lemme 5.4 soit vérifié) et soit C l'ensemble des vecteurs $X \in \exp_x^{-1}(B)$ tels que $t \mapsto \exp_x(tX)$ soit une géodésique minimisante joignant x à $\exp_x(X)$. Alors $\exp_x$ est un difféomorphisme de C sur B . D'après le théorème 4.1 et le lemme 5.3, on a en posant $\theta(x, y) = \det[d \exp_x(X)] = H(x, y)^{-2}$,

$$\begin{aligned} P_x(X_t \in A) &\sim (2\pi t)^{-m/2} \int_B \theta(x, y)^{-1/2} \exp\left[-\frac{d^2(x, y)}{2t}\right] \pi(dy) \\ &= (2\pi t)^{-m/2} \int_C \theta(x, \exp_x X)^{-1/2} \exp\left[-\frac{\|X\|^2}{2t}\right] \theta(x, \exp_x X) \pi_x(dX) \\ &= (2\pi t)^{-m/2} \int_C \theta(x, \exp_x X)^{1/2} \exp\left[-\frac{\|X\|^2}{2t}\right] \pi_x(dX). \end{aligned}$$

En coordonnées sphériques sur $T_x M$, on repère un point X par $r > 0$ et $\omega \in S^{m-1}$. Le point $\exp_x^{-1}(y_0)$ est alors repéré par $d(x, A)$ et $\omega_0 \in S^{m-1}$ et les points $\exp_x^{-1}(y)$ où $y \in \partial A \cap B$ par $f(\omega)$ et ω . Par définition de y_0 , on a $f(\omega) \geq d(x, A)$ pour tout ω .

En diminuant éventuellement ε , l'ensemble C est alors défini par les conditions $\omega \in \Omega$, $f(\omega) \leq r \leq g(\omega)$, où Ω est un voisinage compact de ω_0 dans S^{m-1} et où g est une fonction continue. On a alors

$$(32) \quad P_x(X_t \in A) \sim (2\pi t)^{-m/2} \int_{\Omega} d\omega \int_{f(\omega)}^{g(\omega)} \theta(x, r, \omega)^{1/2} e^{-\frac{r^2}{2t}} r^{m-1} dr$$

quand $t \rightarrow 0$. D'après la méthode de Laplace (cf (15)), on a

$$\int_{f(\omega)}^{g(\omega)} \theta(x, r, \omega)^{1/2} \exp\left[-\frac{r^2}{2t}\right] r^{m-1} dr \sim t \theta(x, f(\omega), \omega)^{1/2} f(\omega)^{m-2} \exp\left[-\frac{f(\omega)^2}{2t}\right]$$

quand $t \rightarrow 0$. De plus cet équivalent est uniforme en ω à condition que $g(\omega) - f(\omega)$ reste minoré par un nombre > 0 . Au lieu d'intégrer sur Ω dans (32), on intègre donc sur l'ensemble Ω_0 des $\omega \in \Omega$ tels que $g(\omega) - f(\omega) \geq \frac{1}{2}[(g(\omega_0) - f(\omega_0))]$ ce qui ne modifie pas l'équivalent puisque Ω_0 est un voisinage de ω_0 . On en déduit que

$$P_x(X_t \in A) \sim (2\pi t)^{-m/2} t \int_{\Omega_0} \theta(x, f(\omega), \omega)^{1/2} f(\omega)^{m-2} \exp\left[-\frac{f(\omega)^2}{2t}\right] d\omega$$

Or la fonction f admet un minimum en ω_0 dans Ω_0 et est de classe C^2 puisque ∂A est de classe C^2 . De plus si x n'est pas point focal de ∂A , le minimum ω_0 de f est non dé-

généralisé. On a alors d'après (16)

$$P_x(X_t \in A) \sim (2\pi t)^{-m/2} t^{(m-1)/2} \theta(x, f(\omega_0) \cdot \omega_0)^{1/2} d(x, A)^{m-2} \\ \left[\det \text{Hess} \frac{f^2}{2}(\omega_0) \right]^{-1/2} \exp \left[-\frac{d^2(x, A)}{2t} \right]$$

Or $\frac{f^2}{2}(\omega) = F(\exp_x f(\omega) \cdot \omega)$. On en déduit que

$$\det \text{Hess} \left(\frac{f^2}{2} \right)(\omega_0) = \det \text{Hess}[F(y_0)] f(\omega_0)^{2m-2} \theta(x, y_0)^2$$

et (31) en résulte.

Remarque.— S'il existe un nombre fini de points $y_1, \dots, y_n \in \partial A$ réalisant le minimum de la distance de x à A , les autres hypothèses de 5.5 restant vérifiées, on a encore

$$P_x(X_t \in A) \sim \sqrt{\frac{t}{2\pi}} \frac{1}{d(x, A)} \exp \left[-\frac{d^2(x, A)}{2t} \right] \sum_{i=1}^n H(x, y_i) \left[\det \text{Hess} F(y_i) \right]^{-1/2}$$

En effet si ε est assez petit, B est formé de voisinages disjoints de chaque y_i et il suffit d'ajouter les contributions de ces voisinages.

5.6.— Donnons un exemple où le minimum de la distance de x à A est atteint en une infinité de points, le calcul de l'équivalent restant possible.

PROPOSITION.— On suppose la variété M complète. Soit $x \in M$ et soit A le complémentaire de la boule de centre x et de rayon ρ où ρ est supposé strictement inférieur au rayon d'injectivité $i(x)$ de x . On a

$$P_x(X_t \in A) \sim \text{Const.} \cdot t^{-m/2} \exp \left[-\frac{\rho^2}{2t} \right]$$

quand $t \rightarrow 0$.

Démonstration.— La condition $\rho < i(x)$ assure que ∂A est une hypersurface de classe C^2 . Soit $\varepsilon > 0$ tel que $\rho + \varepsilon < i(x)$ et soit B l'ensemble compact des points y tels que $\rho \leq d(x, y) \leq \rho + \varepsilon$. Le lemme 5.3 montre que

$$P_x(X_t \in A) \sim \int_B p(t, x, y) \pi(dy) .$$

Comme B ne rencontre pas $C_x(M)$ on a d'après le théorème 4.1,

$$\begin{aligned} P_x(X_t \in A) &\sim (2\pi t)^{-m/2} \int_B H(x, y) \exp \left[-\frac{d^2(x, y)}{2t} \right] \pi(dy) \\ &= (2\pi t)^{-m/2} \int_C \theta(x, \exp_x X)^{1/2} \exp \left[-\frac{\|X\|^2}{2t} \right] \pi_x(dX) \end{aligned}$$

où C est l'ensemble des $X \in T_x M$ tels que $\rho \leq d(x, \exp_x X) \leq \rho + \varepsilon$ et $\theta(x, \exp_x X) = \det [d \exp_x(X)]$.

En passant en coordonnées sphériques dans $T_x M$, on obtient

$$P_x(X_t \in A) \sim (2\pi t)^{-m/2} \int_{S^{m-1}} d\omega \int_{\rho}^{\rho+\varepsilon} \theta(x, r, \omega)^{1/2} \exp\left(-\frac{r^2}{2t}\right) r^{m-1} dr.$$

On peut appliquer (15) et on a

$$\begin{aligned} P_x(X_t \in A) &\sim (2\pi t)^{-m/2} \int_{S^{m-1}} \theta(x, \rho\omega)^{1/2} \rho^{m-2} \exp\left(-\frac{\rho^2}{2t}\right) d\omega \\ &= \text{const. } t^{-m/2} \exp\left(-\frac{\rho^2}{2t}\right). \end{aligned}$$

5.7.- PROPOSITION. - Soit A un domaine fermé de M à bord C^2 . Alors pour $z \in \partial A$, on a

$$P_z(X_t \in A) \rightarrow \frac{1}{2}$$

quand $t \rightarrow 0$ uniformément sur tout compact de ∂A .

Démonstration. - Soit $z_0 \in \partial A$. Il existe un voisinage U de z_0 et des coordonnées $x_1, \dots, x_m$ sur U centrées en z_0 tels que $A \cap U$ soit défini dans U par la condition $x_m \geq 0$. le bord de A est alors défini dans U par l'équation $x_m = 0$.

Soit $\varepsilon > 0$ et soit K (resp. L) l'ensemble des points de U tels que $x_1^2 + \dots + x_m^2 \leq \varepsilon^2$ (resp. $x_1^2 + \dots + x_{m-1}^2 \leq \frac{\varepsilon^2}{4}$ et $x_m = 0$). Pour ε assez petit, K (resp. L) est un voisinage compact de z_0 contenu dans U (resp. $\partial A \cap \partial U$).

On a

$$P_z(X_t \in K) = \int_K p(t, z, y) \pi(dy) \sim (2\pi t)^{-m/2} \int_K \exp \left[-\frac{d^2(z, y)}{2t} \right] H(z, y) \pi(dy)$$

quand $t \rightarrow 0$. La fonction E_z définie par $E_z(y) = \frac{1}{2} d^2(z, y)$ admet en z un minimum unique non dégénéré et la démonstration de la méthode de Laplace (Prop. 3.3) prouve

que

$$(33) \quad P_z(X_t \in K) \rightarrow (2\pi)^{-m/2} \frac{d\pi}{dx}(z) \int_{\mathbb{R}^m} \exp \left[-\frac{1}{2} E_z''(z) \cdot u^2 \right] du$$

quand $t \rightarrow 0$ uniformément pour $z \in L$. La même démonstration où l'on remplace le domaine d'intégration K par l'intersection de K et du demi-espace $x_m \geq 0$ montrerait aussi que

$$(34) \quad P_x(X_t \in A \cap K) \rightarrow (2\pi)^{-m/2} \frac{d\pi}{dx}(z) \int_{\mathbb{R}^{m-1} \times \mathbb{R}^+} \exp \left[-\frac{1}{2} E_z''(z) \cdot u^2 \right] du$$

quand $t \rightarrow 0$, uniformément pour $z \in L$. la dernière intégrale est la moitié de l'intégrale dans (33). Puisque $P_z(X_t \in K) \rightarrow 1$ quand $t \rightarrow 0$, on en déduit que

$$P_z(X_t \in A \cap K) \rightarrow \frac{1}{2}$$

quand $t \rightarrow 0$, uniformément pour $z \in L$.

Comme d'autre part

$$P_z(X_t \in A \cap K^0) \rightarrow 0$$

quand $t \rightarrow 0$ uniformément sur L d'après le lemme 5.2, la proposition 5.7 en résulte.

5.8. PROPOSITION.— Soit A un domaine fermé de M à bord C^2 et soit $x \notin A$. On note τ_A le temps d'entrée dans A pour la diffusion (X_t) . On a alors

$$(35) \quad P_x(X_t \in A) \sim \frac{1}{2} P_x(\tau_A \leq t)$$

quand $t \rightarrow 0$.

Démonstration.—La propriété de Markov forte entraîne

$$\begin{aligned} P_x(X_t \in A) &= E_x [1_{(\tau_A \leq t)} E[1_A(X_t) | \mathcal{F}_{\tau_A}]] \\ &= E_x [1_{(\tau_A \leq t)} \phi_t(\tau_A, X_{\tau_A})] \end{aligned}$$

où $\phi_t(u, z) = E_z [1_A(X_{t-u})]$ pour $z \in \partial A$ et $0 \leq u \leq t$.

Si ∂A est compact, on a d'après la proposition 5.7, $\phi_t(u, z) \rightarrow \frac{1}{2}$ quand $t \rightarrow 0$, uniformément en $z \in \partial A$ et on en déduit (35).

Si ∂A n'est pas compact, soit $\varepsilon > 0$ et posons

$$Z = \partial A \cap \{z ; d(x, y) \leq d(x, A) + \varepsilon\}$$

Comme Z est un compact de ∂A , la démonstration précédente montre que

$$(36) \quad P_x(X_t \in A \text{ et } X_{\tau_A} \in Z) \sim \frac{1}{2} P_x(\tau_A \leq t \text{ et } X_{\tau_A} \in Z).$$

Notons $\mathcal{B}$ l'ensemble des trajectoires continues $\omega : [0, 1] \rightarrow M \cup \{\delta\}$ telles que $\tau_A(\omega) \leq 1$ et $X_{\tau_A}(\omega) \in \partial A - Z$. On a avec les notations de R. Azencott [3] ;

$$\Lambda(\mathcal{B}) = \inf_{\omega \in \mathcal{B}} E(\omega) = \frac{1}{2} [d(x, A) + \varepsilon]^2.$$

L'ensemble $\mathcal{B}$ est fermé dans l'espace des trajectoires (éventuellement explosives) $\omega : [0, 1] \mapsto M \cup \{\delta\}$. On obtient donc en appliquant le théorème 6.4, (R. Azencott [3])

$$\overline{\lim} t \operatorname{Log} P_x(X_{\tau_A} \notin Z \text{ et } \tau_A \leq t) \leq -\frac{1}{2} [d(x, A) + \varepsilon]^2.$$

On a donc pour t assez petit

$$P_x(X_{\tau_A} \notin Z \text{ et } X_t \in A) \leq P_x(X_{\tau_A} \notin Z \text{ et } \tau_A \leq t) \leq \exp \left[-\frac{(d(x, A) + \varepsilon)^2}{2t} + \frac{\varepsilon^2}{2t} \right]$$

D'autre part, d'après le lemme 5.3 on a pour tout $h > 0$

$$P_x(\tau_A \leq t) \geq P_x(X_t \in A) \geq \exp \left[-\frac{d^2(x, A)}{2t} - \frac{h}{t} \right]$$

pour t assez petit.

On en déduit facilement que

$$P_x(\tau_A \leq t) \sim P_x(\tau_A \leq t \text{ et } X_{\tau_A} \in Z)$$

et que

$$P_x(X_t \in A) \sim P_x(X_t \in A \text{ et } X_{\tau_A} \in Z).$$

La formule asymptotique (35) est alors une conséquence de (36).

Remarque.— (35) généralise le principe de réflexion de D. André (cf. [7], p.26).

BIBLIOGRAPHIE

- [1] V.I. Arnol'd, On the stationary phase method and coxeter numbers, *Uspekhi Mat. Nauk*, 28 (1973) = *Russian Math. Surveys*, 28 (1973), p. 19-48.
- [2] V.I. Arnol'd, Normal forms of functions with simple critical points, the Weyl groups A_k , D_k , E_k , and Lagrange manifolds, *Funct. Anal. and its appl.*, 6, 1972, p. 3-25.
- [3] R. Azencott, Grandes déviations et applications, R. Azencott et al., *Ecole d'été de Probabilités de Saint-Flour.VII*, 1978, *Lecture Notes In Mathematics*, 774, Springer-Verlag 1980.
- [4] M.A. Buchner, Stability of the cut locus in dimension less or equal to six, *Inventiones Math.*, 43 (1977) p. 199-231.
- [5] J. Dieudonné, *Calcul infinitésimal*, Hermann, Paris, 1968.
- [6] A. Erdelyi, *Asymptotic expansions*, Dover, New-York 1956.
- [7] K. Ito, H.P. Mc Kean, *Diffusion processes and their sample paths*, Springer Verlag 1965.
- [8] B. Malgrange, Intégrales asymptotiques et monodromie, *Ann. Scient. Ec. Norm. Sup.*, 4^{ème} série, 7, 1974, p. 405-430.
- [9] J. Milnor, *Morse theory*, Ann. of Math. Studies N° 51, Princeton University Press, 1969.
- [10] S.A. Molchanov, Diffusion processes and Riemannian geometry, *Russ. Math. Surveys*, 30, 1975, p. 1-63.

CHAPITRE 10

LA CONDITION D'HYPONELLIPTICITÉ D'HÖRMANDER

par

Mireille CHALEYAT-MAUREL

Après l'étude des diffusions sur une variété associées à un opérateur non dégénéré, il est naturel de s'intéresser à une classe un peu plus générale d'opérateurs auxquels on peut associer une diffusion ; ce sont les opérateurs hypoelliptiques de la forme $\sum_{i=1}^r X_i^2 + Y$ sur $\mathbb{R}^m$ satisfaisant à la condition suffisante d'Hörmander, c'est-à-dire que, en tout point, l'algèbre de Lie engendrée par $X_1, \dots, X_r, Y$ est de rang m . Les diffusions associées à ces opérateurs possèdent alors de nombreuses propriétés analogues à celles des diffusions dont le générateur différentiel est non dégénéré.

Dans une première partie de ce travail nous allons donner les grandes lignes de la démonstration de J.J. Kohn [7] du théorème d'Hörmander et dans la seconde partie nous verrons à quelles conditions un opérateur différentiel du second ordre se met sous la forme d'une somme de carrés ; nous donnerons également des indications sur quelques résultats de théorie du potentiel concernant les opérateurs hypoelliptiques : ouverts très réguliers et problème de Dirichlet.

1ère partie : Condition d'Hörmander.

1. Introduction.

Nous allons donner les idées et les grandes lignes de la démonstration de J.J. Kohn [7] qui utilise des résultats sur les opérateurs pseudo-différentiels

pour étudier l'hypoellipticité de l'opérateur du second ordre :

$$Lu = \sum_{j=1}^r X_j^2 u + X_0 u + cu,$$

$$\text{où } X_j = \sum_{i=1}^m \sigma_{ij} \frac{\partial}{\partial x_i}, \quad j=0,1,\dots,r,$$

avec σ_{ij} appartenant à $C^\infty(U)$, U ouvert relativement compact de $\mathbb{R}^m$.

1.1. Définition.

L est dit hypoelliptique si, lorsque l'équation $Lu = f$ est satisfaite avec u et f distributions sur U , alors on a la condition suivante :

Si V est un ouvert de U tel que $f|_V$ appartienne à $C^\infty(V)$, alors $u|_V$ appartient à $C^\infty(V)$.

J.J. Kohn montre la même condition suffisante d'hypoellipticité qu'Hörmander mais en employant des opérateurs pseudo-différentiels alors que l'article original d'Hörmander [6] utilisait l'étude des groupes à un paramètre engendrés par les champs de vecteurs X_j .

1.2. THÉORÈME. *Si tout champ de vecteurs sur U peut s'exprimer comme combinaison linéaire à coefficients C^∞ de $X_0, X_1, \dots, X_r, \dots, [X_1, X_j], \dots, [X_j, [X_i, X_k]] \dots \dots [X_{i_1}, [X_{i_2}, \dots, X_{i_p}]]$, ... alors L est hypoelliptique.*

Nous donnerons dans la deuxième partie une forme explicite de cette condition en coordonnées locales.

Esquisons le plan de la démonstration du théorème 1.2 :

. On calcule d'abord une estimation à priori de "l'énergie" :

Il existe c tel que pour tout u de $C_0^\infty(U)$,

$$(1) \quad \sum_{j=1}^r \|X_j u\|^2 \leq c(|\langle Lu, u \rangle| + \|u\|^2)$$

(où $C_0^\infty(U)$ désigne les fonctions à support compact dans U).

. Pour u appartenant à $C_0^\infty(U)$ et s réel, on introduit les normes de Sobolev par :

$$\|u\|_s = \left[\int |\hat{u}(\xi)|^2 (1+|\xi|^2)^s d\xi \right]^{1/2}$$

où $\hat{u}$ est la transformée de Fourier de u .

On montre alors que les hypothèses sur l'algèbre de Lie engendrée par les champs de vecteurs entraînent :

Il existe ε et c tels que, pour tout u de $C_0^\infty(U)$,

$$(2) \quad \|u\|_\varepsilon \leq c(\|Lu\| + \|u\|).$$

. Ensuite, on localise (2) ; c'est-à-dire qu'on montre que l'inégalité (2) entraîne :

Il existe ε tel que pour tous ξ et ξ_1 de $C_0^\infty(U)$ avec $\xi_1 = 1$ sur le support de ξ , il existe c tel que pour tout u de $C^\infty(U)$,

$$(3) \quad \|\xi u\|_\varepsilon \leq c(\|\xi_1 Lu\| + \|\xi_1 u\|).$$

. On prouve enfin que (3) entraîne l'hypoellipticité de L .

2. Démonstration de la formule de l'énergie.

2.1. PROPOSITION. Il existe c tel que pour tout u de $C_0^\infty(U)$,

$$(1) \quad \sum_{j=1}^r \|X_j u\|^2 \leq c(|\langle Lu, u \rangle| + \|u\|^2).$$

Preuve : On a $\langle Lu, u \rangle = \langle \sum_{j=1}^r X_j^2 u, u \rangle + \langle X_0 u, u \rangle + \langle cu, u \rangle$.

Remarquons que pour tous $j = 0, \dots, r$,

$$X_j^* u = - \sum_{n=1}^m \frac{\partial}{\partial x_n} (\sigma_{nj} u) = - X_j u + f_j u.$$

Donc, pour tous $j = 1, \dots, r$,

$$\langle X_j^2 u, u \rangle = \langle X_j u, X_j^* u \rangle = - \|X_j u\|^2 + O(\|X_j u\| \|u\|)$$

et

$$\langle X_0 u, u \rangle = -\langle u, X_0 u \rangle + \langle u, f_0 u \rangle, \text{ d'où :}$$

$$|\langle X_0 u, u \rangle| = O(\|u\|^2).$$

En regroupant, on trouve :

$$\sum_{j=1}^r \|X_j u\|^2 \leq |\langle Lu, u \rangle| + \sum_{j=1}^r O(\|X_j u\| \|u\|) + O(\|u\|^2)$$

et en utilisant la formule simple $ab \leq \frac{1}{2n} a^2 + 2n b^2$, on obtient (1) en faisant tendre n vers l'infini.

3. Quelques rappels sur les opérateurs pseudo-différentiels.

3.1. On va définir une classe particulière d'opérateurs pseudo-différentiels qui s'introduira dans la démonstration de la formule (2) et dans la preuve finale de l'hypoellipticité. Pour plus de précisions on pourra consulter l'exposé de L. Nirenberg [8].

L'idée des opérateurs pseudo-différentiels vient du fait que l'on voudrait une classe plus large que celle des opérateurs différentiels et qui contienne, si possible, leurs inverses.

Les opérateurs différentiels sont caractérisés par le fait qu'ils sont locaux; c'est-à-dire, si $L : C_0^\infty(U) \rightarrow C^\infty(U)$ est différentiel, on a :

$$\text{supp } Lu \subset \text{supp } u.$$

Tandis que les opérateurs pseudo-différentiels sont pseudo-locaux; c'est-à-dire, si $L' : \mathcal{E}'(U) \rightarrow \mathcal{D}'(U)$ est pseudo-différentiel, on a $\text{sing supp } L'u \subset \text{sing supp } u$, avec $\text{sing supp } u = \text{complémentaire de l'ouvert où } u \text{ est } C^\infty$; et où $\mathcal{E}'(U)$ désigne les distributions à support compact et $\mathcal{D}'(U)$ les distributions.

On remarquera que l'hypoellipticité d'un opérateur se traduit par l'égalité des deux supports singuliers.

3.2. Considérons les trois familles d'opérateurs suivantes :

(i) multiplication par a appartenant à $C^\infty(U)$.

$$(ii) \quad D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_m^{\alpha_m}} \quad |\alpha| = \sum \alpha_i.$$

(iii) pour tout s réel, Λ^s est défini par :

$$\widehat{\Lambda^s u}(\xi) = (1 + |\xi|^2)^{s/2} \hat{u}(\xi) \text{ pour tout } u \text{ de } C_0^\infty(U).$$

Ces trois familles d'opérateurs n'engendrent pas une algèbre car Λ^s envoie $C_0^\infty(U)$ dans $C^\infty(U)$ mais on peut remplacer les Λ^s par des opérateurs Λ'^s qui envoient $C_0^\infty(U)$ dans $C_0^\infty(U)$ de la façon suivante qui est classique (voir L. Hörmander [5] et J. Dieudonné [3]).

Pour tout s , l'opérateur Λ^s est à noyau C^∞ en dehors de la diagonale, il est donc somme d'un opérateur à noyau C^∞ , Γ^s et d'un opérateur à noyau C^∞ en dehors de la diagonale proprement supporté : Λ'^s .

On rappelle qu'un opérateur à noyau est proprement supporté si le support Σ de son noyau est propre, c'est-à-dire si les deux projections p_x et p_y de $\Sigma \rightarrow U$ sont propres.

$\wedge^s$ envoie alors continûment $C_0^\infty(U)$ dans $C_0^\infty(U)$ et on fait toutes les majorations dans les espaces H_s pour ces opérateurs, elles seront alors valables pour les $\wedge^s$.

3.3. Définissons également l'ordre d'un opérateur.

Soit π un opérateur linéaire de $C_0^\infty(U)$ dans $C_0^\infty(U)$, on dira qu'il est d'ordre n si : pour tout s réel, pour tout compact K , il existe $c(s,K)$ tel que

$$\text{supp } u \subset K \quad \|\pi u\|_s \leq c(s,K) \|u\|_{s+n}.$$

Soient π et π' d'ordre n et n' respectivement on a alors :

$\pi\pi'$ est d'ordre $n+n'$

$[\pi, \pi']$ est d'ordre $n+n' - 1$

π^* est d'ordre n .

Indiquons alors les ordres de la famille d'opérateurs considérés :

- (i) la multiplication par a est d'ordre 0.
- (ii) D^α est d'ordre $|\alpha|$.
- (iii) $\wedge^s$ est d'ordre s .

4. Démonstration de l'estimation (2).

4.1. PROPOSITION. Les hypothèses du théorème 1.2. impliquent :

Il existe ε et c tels que, pour tout u de $C_0^\infty(U)$,

$$(2) \quad \|u\|_\varepsilon \leq c (\|Lu\| + \|u\|).$$

Preuve : Remarquons tout d'abord que si $\varepsilon \leq 1$, on a pour tout u de $C_0^\infty(U)$,

$$(4) \quad \|u\|_\varepsilon \leq \sum_{j=1}^m \|D_j u\|_{\varepsilon-1} + \|u\|.$$

En effet,

$$\begin{aligned} \|u\|_s^2 &= \int |\hat{u}(\xi)|^2 (1+|\xi|^2)^s d\xi = \|u\|_{s-1}^2 + \|\nabla u\|_{s-1}^2 \\ &\leq \|u\|^2 + \|\nabla u\|_{s-1}^2 \quad \text{si } s \leq 1. \end{aligned}$$

Ensuite l'hypothèse du théorème 1.2. entraîne que pour tout j , D_j s'exprime en fonction des crochets des X_i :

$$D_j = \sum a_j^{i_1 \dots i_p} F_{i_1 \dots i_p} \quad 0 \leq i_n \leq r$$

$$\begin{aligned} \text{où } F_{i_1 \dots i_p} &= X_{i_p} \quad \text{si } p = 1 \\ &= [X_{i_p}, F_{i_1 \dots i_{p-1}}] \quad \text{si } p > 1. \end{aligned}$$

Donc d'après (4), il suffit de borner les $\|F_{i_1 \dots i_p} u\|_{\varepsilon-1}$ par $c(\|Lu\| + \|u\|)$ pour obtenir la relation (2).

Ecrivons pour simplifier $F^p = [X, F^{p-1}]$.

On a : $\|F^p u\|_{\varepsilon-1}^2 = \langle F^p u, F^p u \rangle_{\varepsilon-1} = \langle \widehat{F^p u}, (1+|\cdot|^2)^{\varepsilon-1} \widehat{F^p u} \rangle_0$

On définit alors l'opérateur $T^{2\varepsilon-1}$ d'ordre $2\varepsilon-1$ par :

$$T^{2\varepsilon-1} = \wedge^{2\varepsilon-1} \circ F^p \text{ et on obtient :}$$

$$(5) \quad \|F^p u\|_{\varepsilon-1}^2 = \langle F^p u, T^{2\varepsilon-1} u \rangle = \langle XF^{p-1} u, T^{2\varepsilon-1} u \rangle - \langle F^{p-1} Xu, T^{2\varepsilon-1} u \rangle$$

L'idée de la démonstration est une récurrence sur les ε ; on va "descendre" les normes H_ε d'un cran. On estime chaque terme séparément ;

a) si $X = X_j$ $j = 1, \dots, r$, on a d'une part :

$$\langle XF^{p-1} u, T^{2\varepsilon-1} u \rangle = \langle F^{p-1} u, X^* T^{2\varepsilon-1} u \rangle = - \langle F^{p-1} u, XT^{2\varepsilon-1} u \rangle + \langle F^{p-1} u, f_j T^{2\varepsilon-1} u \rangle$$

le deuxième terme étant d'ordre $O(\|u\| \|F^{p-1} u\|_{2\varepsilon-1})$ car $(f_j T^{2\varepsilon-1})^*$ est d'ordre $2\varepsilon-1$.

D'autre part,

$$\begin{aligned} - \langle F^{p-1} u, XT^{2\varepsilon-1} u \rangle &= - \langle F^{p-1} u, [X, T^{2\varepsilon-1}] u \rangle - \langle F^{p-1} u, f_j T^{2\varepsilon-1} u \rangle \\ &= - \langle F^{p-1} u, T^{2\varepsilon-1} Xu \rangle + O(\|u\| \|F^{p-1} u\|_{2\varepsilon-1}) \\ &= O(\|u\| + \|Xu\|) (\|F^{p-1} u\|_{2\varepsilon-1}) \end{aligned}$$

On utilise alors la formule de l'énergie (i) pour obtenir :

$$\langle XF^{p-1} u, T^{2\varepsilon-1} u \rangle \leq c(\|F^{p-1} u\|_{2\varepsilon-1}^2 + \|Lu\|^2 + \|u\|^2)$$

et pour l'autre terme de (5) on ferait les mêmes évaluations.

b) si $X = X_0$, on n'a pas la formule de l'énergie mais on écrit :

$$2X_0 = -L^* + \sum_{j=1}^r X_j^2 + \sum_{j=1}^r b_j X_j + f_0,$$

et on obtient finalement, en faisant les mêmes types de majorations, que dans tous les cas :

$$\|F^p u\|_{\varepsilon-1} \leq c(\|F^{p-1} u\|_{4\varepsilon-1} + \|Lu\| + \|u\|).$$

On itère cette inégalité qui donne alors :

$$\|F^p u\|_{\varepsilon-1} \leq c \sum_{j=1}^r (\|X_j u\|_{4^{p-1}\varepsilon-1} + \|Lu\| + \|u\|)$$

Mais pour $j = 1, \dots, r$, $\|X_j u\|_{-1/2} \leq \|X_j u\| \leq c(\|Lu\| + \|u\|)$ ainsi que pour X_0

car on écrit : $X_0 u = Lu - \sum_{j=1}^r X_j^2 u - cu$

Donc, en prenant $\varepsilon \leq \frac{2}{4^p}$, on obtient :

$\|F^P u\|_{\varepsilon-1} \leq c(\|Lu\| + \|u\|)$, ce qui prouve (2) à l'aide de (4).

5. Localisation de l'inégalité (2).

5.1. LEMME. Soient ξ et ξ_1 appartenant à $C_0^\infty(U)$ tels que $\xi_1 = 1$ sur $\text{supp } \xi$. Supposons (2) vérifiée, alors, il existe c et ε tels que pour tout u de $C^\infty(U)$,

$$(6) \quad \|\xi u\|_\varepsilon \leq c(\|\xi_1 Lu\| + \|\xi_1 u\|)$$

Ce lemme technique est nécessaire car dans la démonstration de l'hypoellipticité, on considère des fonctions qui ne sont que C^∞ .

Preuve : On applique (2) à ξu qui est dans $C_0^\infty(U)$:

$$\|\xi u\|_\varepsilon \leq c(\|L\xi u\| + \|\xi u\|) ,$$

les hypothèses sur les coefficients entraînent qu'il existe c tel que :

$$\|\xi u\| \leq c \|\xi_1 u\|$$

$$\|X_j \xi u\| \leq c \|\xi_1 u\| \quad \text{pour tout } j = 0, \dots, r$$

On a donc :

$$\begin{aligned} \|L \xi u\| &\leq \|[L, \xi] u\| + \|\xi Lu\| \\ &\leq \|[L, \xi] u\| + c \|\xi_1 Lu\| \end{aligned}$$

Donc il suffit de majorer $\|[L, \xi] u\|$ par le deuxième membre de (6); on obtient facilement le résultat en écrivant :

$$[L, \xi] u = 2 \sum_{j=1}^r [X_j, \xi] X_j u + \sum_{j=1}^r [X_j, [X_j, \xi]] u + [X_0, \xi] u$$

6. Démonstration de l'hypoellipticité.

Avant de montrer que (6) entraîne l'hypoellipticité, établissons l'inégalité suivante :

6.1. LEMME. Etant donné un réel s et $N > 0$, il existe $c(s, N)$ tel que

$$(7) \quad \|u\|_{s+\varepsilon} \leq c(s, N)(\|Lu\|_s + \|u\|_{-N})$$

pour tout u de $C_0^\infty(U)$.

Preuve : On a $\|\wedge^s u\|_\varepsilon = \|u\|_{\varepsilon+s} \leq c(\|L \wedge^s u\| + \|\wedge^s u\|)$.

et on utilise le résultat suivant :

Pour tout δ , il existe $c(s, N, \delta)$ tel que

$$(8) \quad \|u\|_s \leq \delta \|u\|_{s+\varepsilon} + c(s, N, \delta) \|u\|_{-N},$$

ce qui se démontre en étudiant le sens de variation de la fonction :
 $t \rightarrow t^{1-\gamma} a + t^{-\gamma} (1-\gamma)b$.

(8) entraîne alors que :

$$\|L \wedge^s u\| + \|\wedge^s u\| \leq \|L \wedge^s u\| + \delta \|u\|_{s+\varepsilon} + c \|u\|_{-N},$$

donc, lorsque δ tend vers zéro, il suffit de majorer $\|L \wedge^s u\|$ pour obtenir (7).

$$\|L \wedge^s u\| \leq \|[L, \wedge^s] u\| + \|\wedge^s Lu\| \leq \|[L, \wedge^s] u\| + c \|Lu\|_s$$

car $\wedge^s$ est d'ordre s ; il suffit donc de majorer $\|[L, \wedge^s] u\|$ que l'on met sous la forme suivante :

$$\begin{aligned} [L, \wedge^s] &= \left[\sum_{j=1}^r X_j^2 + X_0, \wedge^s \right] = \sum_{j=1}^r [X_j^2, \wedge^s] + [X_0, \wedge^s] \\ (9) \quad [L, \wedge^s] &= \sum_{j=1}^r T_{s,j} X_j + T^s \end{aligned}$$

avec $T_{s,j}$ et T_s opérateurs d'ordre s car

$$[X_j^2, \wedge^s] = [X_j, \wedge^s] X_j + X_j [X_j, \wedge^s] = 2 [X_j, \wedge^s] X_j - [[X_j, \wedge^s], X_j]$$

où $[X_j, \wedge^s]$ et $[[X_j, \wedge^s], X_j]$ sont d'ordre s .

Donc

$$\|[L, \wedge^s] u\| \leq c \left(\sum_{j=1}^r \|X_j u\|_s + \|u\|_s \right),$$

mais

$$\|X_j u\|_s = \|\wedge^s X_j u\| \leq \|[X_j, \wedge^s] u\| + \|\wedge^s X_j u\| \leq c(\|u\|_s) + \|X_j \wedge^s u\|.$$

la formule (1) appliquée à $\wedge^s u$ donne alors :

$$\begin{aligned} \sum_{j=1}^r \|X_j \wedge^s u\|^2 &\leq c(|\langle L \wedge^s u, \wedge^s u \rangle| + \|\wedge^s u\|^2) \\ &\leq c(|\langle [L, \wedge^s] u, \wedge^s u \rangle| + \|Lu\|_s^2 + \|u\|_s^2) \end{aligned}$$

On utilise la formule (9) et on trouve :

$$\sum_{j=1}^r \|X_j \wedge^s u\|^2 \leq c \left(\sum_{j=1}^r \|X_j u\|_s \|u\|_s + \|Lu\|_s^2 + \|u\|_s^2 \right)$$

ce qui donne en employant le même argument que dans le paragraphe 2,

$$\sum_{j=1}^r \|X_j u\|_s^2 \leq c(\|Lu\|_s^2 + \|u\|_s^2),$$

d'où la formule (7).

6.2. Ensuite on localise (7) par la même méthode que dans le paragraphe 5 et on obtient :

Pour tous ξ et ξ_1 de $C_0^\infty(U)$ avec $\xi_1 = 1$ sur $\text{supp } \xi$, et pour tout u de $C^\infty(U)$,

$$(10) \quad \|\xi u\|_{s+\varepsilon} \leq c(\|\xi_1 Lu\|_s + \|\xi_1 u\|_{-N})$$

6.3. Nous sommes maintenant en mesure de prouver l'hypoellipticité de L .

Soit u appartenant à $\mathcal{D}'(U)$ tel que $Lu = f$ et supposons qu'il existe V ouvert de U tel que $f|_V$ appartienne à $C^\infty(V)$; on veut montrer que $u|_V$ appartient à $C^\infty(V)$.

On pose $S_\delta u(x) = \delta^{-m} \int u(y) \varphi\left(\frac{y-x}{\delta}\right) dy = u * \varphi_\delta(x)$,
avec $\varphi_\delta(x) = \delta^{-m} \varphi\left(\frac{x}{\delta}\right)$ où φ appartient à $C_0^\infty(U)$ et $\int \varphi(y) dy = 1$.

On va montrer que pour tout ξ de $C_0^\infty(U)$, ξu appartient à H_s quelque soit s , ce qui entraînera que ξu appartient à $C^\infty(V)$, d'où $u|_V$ appartient à $C^\infty(V)$.

Pour cela on montre d'abord un lemme :

6.4. LEMME. Les deux assertions suivantes sont équivalentes :

(i) il existe c indépendante de δ tel que

$$\|S_\delta \xi u\|_s \leq c.$$

(ii) ξu appartient à H_s .

Preuve : (i) $\Rightarrow$ (ii) : en effet, $S_\delta \xi u$ appartient à une boule de H_s qui est hilbertien donc ξu tend faiblement vers une limite $\ell = \xi u$ appartenant à H_s lorsque δ tend vers zéro.

$$\begin{aligned} \text{(ii)} \Rightarrow \text{(i)} \quad & \text{car} \\ \|S_\delta \xi u\|_s &= \int |S_\delta \xi u(x)|^2 (1+|x|^2)^s dx \\ &= \int |\hat{\varphi}(\delta x)|^2 |\hat{\xi u}(x)|^2 (1+|x|^2)^s dx \\ &\leq \|\xi u\|_s \quad \text{car} \quad |\hat{\varphi}(\delta x)| \leq |\hat{\varphi}(0)| = 1. \end{aligned}$$

D'après le lemme 6.4., il suffit de majorer $\|S_\delta \xi u\|_s$ indépendamment de δ et ceci pour tout s , on aura alors ξu dans H_s pour tout s .

6.5. Remarquons que si ξu appartient à H_s ,

$$(11) \quad \|S_\delta \xi u\|_{s+\varepsilon} \leq c(\|S_\delta \xi_1 f\|_s + \|S_\delta \xi_1 u\|_{-N} + \|\xi_1 u\|_s)$$

avec c indépendant de δ (pour démontrer (11) on fait le même raisonnement que pour (7) en remarquant que si T^1 est un opérateur pseudo-différentiel d'ordre 1,

$$\| [S_\delta \quad \xi, T^1] u \|_s \leq c \|u\|_s$$

avec c indépendante de δ .

6.6. Donc, soit u appartenant à $\mathcal{D}'(U)$, il existe N tel que $\xi_1 u$ soit dans H_{-N} si le support de ξ_1 est assez petit; en appliquant l'inégalité (11) avec $s = -N$, on trouve :

$$\| S_\delta \xi u \|_{-N+\epsilon} \leq c, \quad c \text{ indépendante de } \delta, \text{ ce qui entraîne : } \xi u \text{ appartient à } H_{-N+\epsilon}.$$

On recommence alors avec $s = -N+\epsilon$ et on obtient ξu appartient à H_s pour tout s , d'où ξu dans C^∞ donc u dans C^∞ pour un voisinage ouvert éventuellement plus petit que V .

2ème partie : Quelques propriétés des opérateurs hypoelliptiques.

1. Diffusion associée.

Les opérateurs hypoelliptiques s'introduisent naturellement dans l'étude des diffusions associées à des opérateurs différentiels du second ordre. En effet, lorsque les coefficients sont C^∞ , il y a existence et unicité de la diffusion associée à

l'opérateur $L = \sum_{i=1}^r X_i^2 + Y$; le résultat s'obtient très facilement en considérant l'équation différentielle associée.

Et si L et $L - \frac{\partial}{\partial t}$ satisfont à la condition suffisante d'hypoellipticité d'Hörmander, on montre que le semi-groupe de transition possède une densité $p_t(x, y)$ C^∞ sur $\mathbb{R}^x \times U \times U$ où U est l'ouvert de définition. Pour toutes ces questions voir l'exposé de R. Azencott (chap. 2) et l'article de A. Bonami et Altri [1].

2. Forme locale des opérateurs.

2.1. On peut se demander également à quelles conditions un opérateur différentiel du second ordre général, à coefficients C^∞ , peut se décomposer en somme de carrés d'opérateurs du premier ordre à coefficients C^∞ (ce qui suit est tiré du cours de J.M. Bony [2]).

Il est clair en considérant l'exemple simple sur $\mathbb{R}^2$, $L = (x^2 + y^2) \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$

qu'on ne peut pas avoir le résultat sur tout l'espace.

Quand on se donne un opérateur à coefficients C^∞ :

$$L = \sum_{i,j=1,\dots,m} a_{ij}(\cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1,\dots,m} b_i(\cdot) \frac{\partial}{\partial x_i},$$

défini sur un ouvert U de $\mathbb{R}^m$, la forme quadratique définie par la matrice $(a_{ij}(x))$ se décompose en chaque point x en somme de carrés de formes linéaires indépendantes; le nombre de ces formes linéaires est une fonction semi-continue inférieurement donc localement constante sur un ouvert dense.

On pourra donc écrire dans cet ouvert

$$L = \sum_{j=1, \dots, r} X_j^2 + Y$$

où les X_j et Y sont des opérateurs du premier ordre à coefficients C^∞ .

2.2. Explicitons localement l'hypothèse d'hypoellipticité d'Hörmander.

Soit M une variété C^∞ connexe de dimension m . Soit L un opérateur différentiel du second ordre sur M annulant les constantes et semi-elliptique. Ceci équivaut à dire que pour toute carte locale (U, φ) , l'image L_φ de L par φ peut s'écrire :

$$L_\varphi = \sum_{1 \leq i, j \leq m} a_{ij}(\cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq m} b_i(\cdot) \frac{\partial}{\partial x_i}$$

où la matrice symétrique $a_\varphi = (a_{ij})$ est positive au sens large; on note b_φ le vecteur de coordonnées b_i

(H) : Dans le cas hypoelliptique classique on suppose qu'il existe une base d'ouverts de M tels que sur chaque ouvert V de cette base on puisse écrire

$L = \sum_{i=1}^r X_i^2 + Y$ où $X_1, \dots, X_r, Y$ sont des champs de vecteurs de classe C^∞ sur V et où l'algèbre de Lie engendrée par $X_1, \dots, X_r, Y$ $\mathcal{L}(X_1, \dots, X_r, Y)$ est de dimension $m = \dim M$ en tout point de V . (On rappelle que $\mathcal{L}(X_1, \dots, X_r, Y)$ est le plus petit C^∞ -module engendré par $X_1, \dots, X_r, Y$ et leurs crochets).

Si on fait une hypothèse supplémentaire on peut avoir une forme équivalente de (H) en coordonnées locales :

LEMME. Si dans chaque ouvert de la base définie dans (H) la dimension de l'algèbre de Lie engendrée par $X_1, \dots, X_r$ est constante en tout point, il existe un atlas (U, φ) tel que (H) soit équivalente à la propriété suivante :

Les vecteurs colonnes de a_φ et b_φ engendrent par crochet de Lie une algèbre de Lie de dimension m en tout point.

Preuve. Soit V un ouvert de la base et $x \in V$; si la dimension de $\mathcal{L}(X_1, \dots, X_r)$ est égale à n_1 sur V , il existe d'après le théorème de Frobenius une carte (U, φ) telle que les X_i s'expriment en fonction des n_1 premiers $\frac{\partial}{\partial x_j}$:

$$X_i = \sum_{j=1}^{n_1} \sigma_{ji}(x) \frac{\partial}{\partial x_j}, \quad 1 \leq i \leq r$$

Dans cette base le champ Y s'écrit

$$Y = \sum_{i=1}^m (b_i(x) - \sum_{1 \leq j, k \leq n_1} \sigma_{jk}(x) \frac{\partial \sigma_{ik}(x)}{\partial x_j}) \frac{\partial}{\partial x_i}$$

Donc $\mathcal{L}(X_1, \dots, X_r, Y)(x) = \mathcal{L}(X_1, \dots, X_r, b_\varphi)(x)$ car $\sum_{i=1}^m (\sum_{j,k} \sigma_{jk}(x) \frac{\partial \sigma_{ik}(x)}{\partial x_j}) \frac{\partial}{\partial x_i}$ appartient à $\mathcal{L}(X_1 \dots X_r)(x)$.

Il suffit ensuite de remarquer que si on note $\tilde{X}_1, \dots, \tilde{X}_m$ les vecteurs colonnes de a_φ , les espaces vectoriels engendrés d'une part par $\tilde{X}_1(x), \dots, \tilde{X}_m(x)$ et $X_1(x), \dots, X_r(x)$ coïncident puisque $\text{Im } a_\varphi(x) = \text{Im } \sigma_\varphi(x)$ si $a_\varphi = \sigma_\varphi \sigma_\varphi^*$.

Les $\tilde{X}_i$ s'expriment comme combinaisons linéaires à coefficients C^∞ des X_j et réciproquement on peut exprimer les X_i en fonction de $n_1 \tilde{X}_j$ car le rang est constant.

Alors la formule

$$[fX, gX'] = fg [X, X'] + f(Xg)X' - g(X'f)X$$

avec X et X' champs de vecteurs et f, g fonctions, montre que les C^∞ -modules engendrés par $X_1, \dots, X_r, b$ et $\tilde{X}_1 \dots \tilde{X}_n, b$ sont les mêmes.

3. Problème de Dirichlet.

Un autre problème lié classiquement à l'opérateur est la résolution du problème de Dirichlet. On se place dans un ouvert U de $\mathbb{R}^m$, lorsque L est sous la forme $L = \sum_{i=1}^r X_i^2 + Y$ avec $\mathcal{L}(X_1 \dots X_r, Y)$ de rang m en tout point, J.M. Bony [2] a montré l'existence, l'unicité et la positivité de la solution du problème de Dirichlet dans les ouverts très réguliers.

3.1. Définition

Un ouvert ω est très régulier s'il possède la propriété suivante :

Pour tout x_1 de $\partial\omega$, il existe une sphère centrée en un point x_0 , ne rencontrant ω qu'en x_1 et telle que :

$$\sum a_{ij}(x_1)(x_1^i - x_0^i)(x_1^j - x_0^j) > 0.$$

J.M. Bony a alors démontré le théorème suivant :

3.2. THÉOREME (J.M. Bony). *On suppose que L vérifie la condition suffisante d'hy-poellipticité d'Hörmander et que ω est très régulier. Alors, pour toute fonction φ continue sur $\partial\omega$, il existe une et une seule fonction u , continue sur $\bar{\omega}$, C^∞ sur ω telle que :*

$$\begin{cases} Lu = 0 & \text{dans } \omega \\ u = \varphi & \text{sur } \partial\omega \end{cases}$$

et si $\varphi \geq 0$, cela entraîne $u \geq 0$

On en déduit l'existence d'une base d'ouverts très réguliers.

3.3. COROLLAIRE. Pour les solutions de $Lu = 0$ les ouverts très réguliers sont réguliers et forment une base de la topologie de U .

Preuve. Montrons que tout point z possède un système fondamental de voisinages très réguliers. L'opérateur L étant semi-elliptique non totalement dégénéré, il existe un vecteur θ unitaire tel que

$$\sum_{i,j} a_{ij}(z) \theta_i \theta_j > 0,$$

avec $a = \sigma\sigma^*$, σ étant la matrice des X_i . L'intersection des deux boules de rayon $M+\varepsilon$ centrées respectivement en $z + M\theta$ et $z - M\theta$ est un ouvert très régulier pourvu que M soit assez grand et ε assez petit.

Rappelons qu'un ouvert ω relativement compact dans U est régulier si pour toute fonction f continue sur $\partial\omega$, il existe une et une seule fonction $H^\omega f$ continue sur ω , harmonique dans ω et égale à f sur $\partial\omega$, et si de plus, $H^\omega f$ est positive dès que f l'est, le théorème 3.2. montre l'existence l'unicité et la positivité pour la résolution du problème de Dirichlet, il montre donc que les ouverts très réguliers sont réguliers et de plus forment une base de la topologie de U .

3.4. Rappelons les notions probabilistes de la régularité (on pourra consulter A. Bonami et Altri [1]).

On définit tout d'abord ce qu'est un point régulier.

Soit ω un ouvert de $\mathbb{R}^m$, x_0 appartenant à $\partial\omega$ est dit régulier si $P_{x_0}[\tau > 0] = 0$ où τ est le premier temps de sortie de ω pour X_t , diffusion associée à L .

L'ouvert ω sera dit régulier si tous les points de $\partial\omega$ sont réguliers.

De même, on dit qu'un point x_1 appartenant à $\partial\omega$ est très régulier si, quelque soit u positif, quelque soit ε , il existe un voisinage $V(x_1)$ de x_1 tel que si x appartient à $V(x_1)$,

$$P_x(\tau > u) < \varepsilon.$$

ω est très régulier si tous les points de $\partial\omega$ sont très réguliers.

Les ouverts très réguliers au sens de Bony sont très réguliers au sens probabiliste et on remarque que lorsque X_t est fortement fellerienne (ce qui est le cas pour les diffusions associées à un opérateur hypoelliptique à coefficients C^∞) les ouverts réguliers sont très réguliers au sens probabiliste (voir E.B. Dynkin [4]).

BIBLIOGRAPHIE

- [1] A. Bonami, N. El.Karoui , M. Reinhard B.Roynette. *Processus de diffusion associé à un opérateur elliptique dégénéré*. Ann. Inst. Henri Poincaré vol. VII n°1 (1971) p. 31-80.
- [2] J.M. Bony, *Opérateurs elliptiques dégénérés associés aux axiomatiques de la théorie du potentiel*. Cours du C.I.M.E. (1969).
- [3] J. Dieudonné, *Eléments d'Analyse. Tome 7*. Gauthier-Villars(1978).
- [4] E.B. Dynkin, *Markov processes*. Vol. II. Springer (1965).
- [5] L. Hörmander, *Pseudo-differential operators and hypoelliptic equations*. Proc. symp.pure math. Vol. 10. AMS providence RI (1966) p. 138-183.
- [6] L. Hörmander, *Hypoelliptic second order differential equation* . Acta. Math. Vol. 119 (1967) p.147-171.
- [7] J.J. Kohn, *Pseudo-differential operators and hypoellipticity*. Proc. symp. pure math. Vol. 23. AMS providence RI (1969) p. 61-69.
- [8] L. Nirenberg, *Pseudo-differential operators*. Proc. sym. math. Vol. 16 AMS providence RI (1970) p. 149-167.

CHAPITRE 11

DENSITÉS DES DIFFUSIONS INVARIANTES SUR CERTAINS GROUPES NILPOTENTS CALCUL D'APRÈS B. GAVEAU

par

Mireille CHALEYAT-MAUREL

Les opérateurs hypoelliptiques considérés ici, du type $L = \frac{1}{2} \sum_{i=1}^m X_i^2$, généralisent de façon naturelle pour les groupes de Lie nilpotents d'ordre quelconque le modèle qu'est le laplacien pour les opérateurs non dégénérés ; ce sont, en quelque sorte, ceux de rang minimum vérifiant la condition d'hypoellipticité d'Hörmander.

Nous allons montrer comment B. Gaveau [2] calcule explicitement les solutions fondamentales de l'équation de la chaleur pour des groupes de Lie simplement connexes, nilpotents d'ordre 2.

La méthode employée repose sur un calcul d'espérance conditionnelle où interviennent des intégrales stochastiques de la forme :

$$\int_0^t \tilde{\beta}_s d\beta_s - \beta_s d\tilde{\beta}_s, \beta \text{ et } \tilde{\beta} \text{ étant deux mouvements browniens indépendants.}$$

Dans un premier paragraphe, nous montrerons comment, suivant une idée de P. Levy [4], on développe chaque mouvement brownien en série de Fourier et on obtient une formule pour l'intégrale stochastique $\int_0^t \tilde{\beta}_s d\beta_s - \beta_s d\tilde{\beta}_s$ ne faisant intervenir que des variables aléatoires gaussiennes indépendantes ; ensuite, dans le paragraphe deux, nous donnerons les grandes lignes du calcul de la solution fondamentale.

1. Calcul de $\int_0^{2\pi} (\tilde{\beta}_s d\beta_s - \beta_s d\tilde{\beta}_s).$

1.1. Développement en série de Fourier du mouvement brownien.

Soit β un mouvement brownien sur $[0, 2\pi]$; si la dérivée $\dot{\beta}$ existait et était dans $L^2([0, 2\pi])$, on pourrait développer $\dot{\beta}$ par rapport à une base orthogonale de $L^2([0, 2\pi])$: $(f_n)_{n \in \mathbb{Z}}$ où $f_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}$.

Un calcul formel donnerait alors :

$$\dot{\beta} = \sum_{n \in \mathbb{Z}} c_n f_n,$$

$$\text{avec } c_n = \langle \dot{\beta}, f_n \rangle = \int_0^{2\pi} \dot{\beta}_s \overline{f_n(s)} ds = \int_0^{2\pi} \overline{f_n(s)} d\beta_s,$$

mais la dernière intégrale stochastique a maintenant un sens. On peut donc définir $c_n = \int_0^{2\pi} \overline{f_n(s)} d\beta_s$; ce sont des variables aléatoires gaussiennes, centrées, réduites, indépendantes et on montre alors que $\int_0^{2\pi} g(s) d\beta_s = \sum_{n \in \mathbb{Z}} \overline{c_n} \langle g, f_n \rangle$ pour g dans $L^2([0, 2\pi])$ car l'application $g \rightarrow \phi(g) = \sum_{n \in \mathbb{Z}} \overline{c_n} \langle g, f_n \rangle$ est une isométrie linéaire de $L^2([0, 2\pi])$ dans $L^2(\Omega)$.

De même, si g est un processus aléatoire indépendant de β , progressivement mesurable tel que : $E \int_0^{2\pi} g^2(s) ds < \infty$, on a :

$$\int_0^{2\pi} g(s, \omega) d\beta_s = \sum_{n \in \mathbb{Z}} \overline{c_n} \langle g, f_n \rangle.$$

D'où un développement de β_t en prenant $g = 1_{[0, t]}$,

$$\beta_t = \sum_{n \in \mathbb{Z}} \overline{c_n} \int_0^t \overline{f_n(s)} ds$$

$$\text{et en calculant } \int_0^t \overline{f_n(s)} ds = \begin{cases} \frac{i}{n} f_{-n}(t) - \frac{i}{n} f_0(t) & \text{si } n \neq 0 \\ \frac{t}{\sqrt{2\pi}} & \text{si } n = 0 \end{cases}$$

on trouve :

$$\beta_t = \sum_{n \neq 0} c_{-n} \frac{i}{n} f_{-n}(t) - \left(\sum_{n \neq 0} c_{-n} \frac{i}{n} \right) f_0(t) + \frac{c_0 t}{\sqrt{2\pi}}.$$

1.2. Calcul de l'intégrale stochastique.

Considérons maintenant un nouveau mouvement brownien $\tilde{\beta}$ avec les coefficients de Fourier $\tilde{c}_n$ par rapport à la base $(f_n)_{n \in \mathbb{Z}}$. On a :

$$\int_0^{2\pi} \tilde{\beta}_s d\beta_s = \sum_{\mathbb{Z}} \overline{c_n} \langle \tilde{\beta}, f_n \rangle$$

Calculons $\langle \tilde{\beta}, f_n \rangle$:

$$\langle \tilde{\beta}, f_n \rangle = \left\langle \sum_{n \neq 0} \tilde{c}_{-n} \frac{i}{n} f_{-n} - \left(\sum_{n \neq 0} \tilde{c}_{-n} \frac{i}{n} \right) f_0 + \frac{\tilde{c}_0 h}{\sqrt{2\pi}}, f_n \right\rangle$$

avec $h(t) = t$.

$$\begin{aligned} \text{Soit } \langle \tilde{\beta}, f_n \rangle &= -\frac{i}{n} \tilde{c}_n + \frac{\tilde{c}_0}{in} \quad \text{si } n \neq 0 \\ &= -\left(\sum_{n \neq 0} \tilde{c}_{-n} \frac{i}{n} \right) + \pi \sqrt{2\pi} \tilde{c}_0 \quad \text{si } n = 0 \end{aligned}$$

D'où finalement :

$$\int_0^{2\pi} (\tilde{\beta}_s d\beta_s - \beta_s d\tilde{\beta}_s) = \sum_{n \neq 0} \frac{1}{in} [\tilde{c}_n c_{-n} - c_n \tilde{c}_{-n} + 2(\tilde{c}_0 c_n - c_0 \tilde{c}_n).]$$

En revenant aux variables réelles, on pose

$$c_n = \frac{A_n + iB_n}{\sqrt{2}}, \quad \tilde{c}_n = \frac{\tilde{A}_n + i\tilde{B}_n}{\sqrt{2}}$$

où les $A_n, B_n, \tilde{A}_n, \tilde{B}_n$ sont des variables aléatoires gaussiennes, centrées, réduites indépendantes, alors :

$$\int_0^{2\pi} \tilde{\beta}_s d\beta_s - \beta_s d\tilde{\beta}_s = 2 \sum_{n \geq 1} \frac{1}{n} [B_n (A_n - c\sqrt{2}) - B_n (\tilde{A}_n - \tilde{c}\sqrt{2})]$$

$$\text{avec } c = \frac{\beta_{2\pi}}{\sqrt{2\pi}}, \quad \tilde{c} = \frac{\tilde{\beta}_{2\pi}}{\sqrt{2\pi}}.$$

2. Calcul de la densité.2.1. Position du problème.

On considère une algèbre de Lie nilpotente stratifiée, libre d'ordre 2, c'est-à-dire :

$$\mathcal{A} = V_1 \oplus V_2 \quad \text{avec} \quad [V_1, V_1] = V_2, \quad [V_1, V_2] = 0, \\ [V_2, V_2] = 0.$$

V_1 est de dimension m , engendrée par $X_1, X_2, \dots, X_m$;

V_2 est de dimension $\frac{m(m-1)}{2}$, engendrée par les $[X_i, X_k]_{i < k}$ que nous noterons X_{ik} .

Soit N le groupe nilpotent simplement connexe d'algèbre de Lie $\mathfrak{d}$ et L son laplacien hypoelliptique associé :

$$L = \frac{1}{2} \sum_{i=1}^m X_i^2.$$

L'objet du travail est le calcul de la solution fondamentale de

$$\frac{\partial}{\partial t} - L = 0.$$

Les opérateurs $\frac{\partial}{\partial t} - L$ et L satisfaisant évidemment à la condition suffisante d'hypoellipticité d'Hörmander, on connaît l'existence d'une solution régulière : $p_t(0, g)$ qui est la densité de la probabilité de transition de la diffusion de générateur différentiel L . (Voir l'exposé de R. Azencott (chap 2)).

On considère la courbe exponentielle de N

$$N \longrightarrow R^{m + \frac{m(m-1)}{2}} : g \longrightarrow (u_i, v_{ik})_{\substack{i < k \\ 1 \leq i \leq m}}$$

$$\text{Si } g = \exp\left(\sum_{i=1}^m u_i x_i + \sum_{i < k} v_{ik} x_{ik}\right)$$

si on note $\partial_i = \frac{\partial}{\partial u_i}$, $\partial_{ik} = \frac{\partial}{\partial v_{ik}}$, les champs X_i et X_{ik} se lisent par

$$\partial_i + \frac{1}{2} \left(\sum_{j < i} u_j \partial_{ji} - \sum_{j > i} u_j \partial_{ij} \right) \text{ et } \partial_{ij}$$

ce qui donne pour L :

$$\begin{aligned} 2L = & \sum_i \partial_i^2 + \sum_i \left(\sum_{k < i} u_k \partial_{ki} \partial_i - \sum_{k > i} u_k \partial_{ik} \partial_i \right) \\ & + \frac{1}{4} \sum_i \left(\sum_{k, l < i} u_k u_l \partial_{ki} \partial_{li} + \sum_{k, l > i} u_k u_l \partial_{ik} \partial_{il} - 2 \sum_{\substack{k < i \\ l > i}} u_k u_l \partial_{ki} \partial_{il} \right) \end{aligned}$$

Nous ne donnons pas les détails des calculs qui peuvent être trouvés dans B. Gaveau [2].

Soit maintenant un mouvement brownien standard de $\mathbb{R}^m$: $\beta_1(t), \dots, \beta_m(t)$; si on pose $\beta_{k1}(t) = \frac{1}{2} \int_0^t (\beta_k(s) d\beta_1(s) - \beta_1(s) d\beta_k(s))$,

$$G(t) = (\beta_1(t), \dots, \beta_m(t), \beta_{12}(t), \dots, \beta_{k1}(t), \dots)_{\substack{k < l \\ l=1, \dots, m}}$$

est une diffusion de générateur L lue dans la carte exponentielle.

En effet, il suffit d'écrire G comme solution de l'équation différentielle :

$$dG(t) = \sigma(G(t)) d\beta(t)$$

$$\text{où } \beta(t) = (\beta_1(t), \dots, \beta_m(t)) \quad \text{et } \sigma = \begin{bmatrix} I_m \\ v \end{bmatrix}$$

$$\text{avec } v \text{ matrice } (\frac{m(m-1)}{2}, m) : v(x) = \begin{bmatrix} \frac{1}{2}x_2 & -\frac{1}{2}x_1 & 0 & \dots \\ \frac{1}{2}x_3 & 0 & -\frac{1}{2}x_1 & \dots \\ \vdots & & & \end{bmatrix}$$

Son générateur différentiel est $\frac{1}{2} \text{grad}^* a \text{grad}$ avec

$$a = \sigma \sigma^* = \begin{bmatrix} I_m \\ v \end{bmatrix} \begin{bmatrix} I_m & v^* \end{bmatrix} = \begin{bmatrix} I_m & v^* \\ v & vv^* \end{bmatrix}$$

et si on calcule les coefficients de cette matrice, ce sont ceux de a donnés par l'expression de L .

2.2. Calcul de la transformée de Fourier de p_t .

Soit $p_t(0, g)$ la solution fondamentale de $\frac{\partial}{\partial t} - L = 0$ et soit $F_t(w)$ avec $w = (u_1, \dots, u_m, \dots, v_{k1}, \dots)_{k < l}$
 $= (u, v)$

sa lecture dans la carte exponentielle ; F_t est la densité de la loi de

$G(t) = (\beta_1(t), \dots, \beta_m(t), \dots, \beta_{k1}(t))$, elle vérifie pour toute fonction bornée f :

$$E_0 f(G(t)) = \int_{\mathbb{R}^{m + \frac{m(m-1)}{2}}} F_t(w) f(w) dw$$

On va calculer la transformée de Fourier de F_t et pour cela soit $\zeta = (\eta, \xi)$ les variables duales de $w = (u, v)$.

Notons que ξ (comme v) peut être identifié à une matrice antisymétrique dont les éléments au dessus de la diagonale sont les ξ_{kl} . Nous noterons également ξ cette matrice.

Avec toutes les notations ci-dessus, nous allons montrer :

2.3. PROPOSITION. La transformée de Fourier de F_t est donnée par la formule

$$\hat{F}_s(\zeta) = (2\pi s)^{-m/2} \int_{\mathbb{R}^m} \exp(i\langle \eta, u \rangle - \frac{|u|^2}{2s}) \prod_{k=1}^{\infty} \psi_s(u, \frac{\xi}{k}) du$$

où on a posé

$$\psi_s(u, \frac{\xi}{k}) = \exp \left[\frac{1}{s} (-|u|^2 + t_u (I - \frac{s^2 \xi^2}{4\pi^2 k^2})^{-1} u) \right] \times \det(I - \frac{s \xi}{2\pi k})^{-1}.$$

Preuve : nous ne donnons que les grandes lignes du calcul.

Soit à calculer :

$$\hat{F}_s(\eta, \xi) = E_0[\exp i\langle \eta, \beta(s) \rangle] E_0[\exp i \sum_{k=1}^{\infty} \xi_{kl} \beta_{kl}(s) / \beta_j(s), j=1 \dots n]$$

On calcule la loi conditionnelle en développant chaque β_{kl} en série de Fourier grâce à la formule du paragraphe 1.

Mais tout d'abord, on se ramène à $s = 2\pi$ et on renormalisera à la fin par le temps final s .

On a alors pour chaque $\beta_{kl}(2\pi)$:

$$\beta_{kl}(2\pi) = \sum_{n=1}^{\infty} \frac{1}{n} [a_n^{(k)} (b_n^{(1)} - \frac{\beta_1}{\sqrt{\pi}}) - a_n^{(1)} (b_n^{(k)} - \frac{\beta_k}{\sqrt{\pi}})]$$

L'intérêt de cette méthode est qu'il n'y a plus à conditionner car le conditionnement revient à remplacer les $\beta_i(s)$ par leur valeur en 2π : $\beta_i(2\pi) = \beta_i$.

Il s'agit de calculer :

$$E_0[\exp \{i \sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=1}^{\infty} \xi_{kl} (a_n^{(k)} (b_n^{(1)} - \frac{\beta_1}{\sqrt{\pi}}) - a_n^{(1)} (b_n^{(k)} - \frac{\beta_k}{\sqrt{\pi}})) \}]$$

toutes les variables étant indépendantes, on peut intervertir l'espérance et la sommation, ce qui donne :

$$\prod_{n=1}^{\infty} I_n,$$

$$\text{avec } I_n = E_o \left[\exp \frac{i}{n} \sum_{k < 1} \xi_{k1} \left(a_n^{(k)} \left(b_n^{(1)} - \frac{\beta_1}{\sqrt{\pi}} \right) - a_n^{(1)} \left(b_n^{(k)} - \frac{\beta_k}{\sqrt{\pi}} \right) \right) \right]$$

$$= E_o \left[\exp \frac{i}{n} \sum_k \mu_k a_n^{(k)} \right],$$

$$\text{si l'on a posé : } \mu_k = \sum_{k < 1} \xi_{k1} \left(b_n^{(1)} - \frac{\beta_1}{\sqrt{\pi}} \right) - \sum_{k > 1} \xi_{1k} \left(b_n^{(1)} - \frac{\beta_1}{\sqrt{\pi}} \right).$$

Les variables $a_n^{(k)}$ étant gaussiennes, centrées, réduites, on obtient :

$$I_n = E_o \left(\exp - \frac{1}{2n^2} \sum_k \mu_k^2 \right).$$

Désignons par $b_n'^{(1)}$ le vecteur de composantes $b_n^{(1)} - \frac{\beta_1}{\sqrt{\pi}}$ et par b_n' le vecteur de composantes $b_n'^{(1)}$; alors,

$$I_n = E_o \left(\exp - \frac{1}{2n^2} \langle \xi b_n', \xi b_n' \rangle \right)$$

$$= E_o \left(\exp + \frac{1}{2} \langle B b_n', b_n' \rangle \right),$$

$$\text{si l'on a noté : } B = \frac{\xi^2}{n^2}.$$

En utilisant le fait que b_n' est un vecteur gaussien de moyenne $-\frac{\beta}{\sqrt{\pi}}$ avec $\beta = (\beta_1, \dots, \beta_m)$ et la formule :

$$E[\exp - \langle t, Z \rangle] = \exp (- \langle t, E(Z) \rangle + \frac{1}{2} \langle t, K t \rangle)$$

si Z est un vecteur gaussien de moyenne $E(Z)$, et de covariance K ; on calcule I_n et on trouve :

$$I_n = \exp - \frac{|u|^2}{2\pi} \sqrt{\det(I-B)^{-1}} \exp \left\{ \frac{1}{2} \left\langle \frac{\beta}{\sqrt{\pi}}, (I-B)^{-1} \frac{\beta}{\sqrt{\pi}} \right\rangle \right\}$$

mais $I - B = (I - \frac{\xi}{n})(I + \frac{\xi}{n})$ avec ξ n'ayant pas de valeurs propres réelles ; donc si on pose :

$$\psi_s(u, X) = \exp \frac{1}{s} (-|u|^2 + \langle u, (I - \frac{s^2 X^2}{4\pi^2})^{-1} u \rangle) \\ \times [\det(I - \frac{sX}{2\pi})]^{-1},$$

en renormalisant par $u(2\pi) \longrightarrow u(s) \frac{s}{2\pi}$, on trouve la formule cherchée.

2.4. Réduction du calcul.

La matrice ξ étant antisymétrique, il existe une transformation orthogonale W de $\mathbb{R}^m$ telle que ${}^t W \xi W = P$ où P est une matrice antisymétrique formée de blocs :

$$P = \begin{bmatrix} 0 & P_1 & 0 & \dots \\ -P_1 & 0 & & \\ 0 & & \ddots & \\ \vdots & & & 0 & P_{2k-1} \\ & & & -P_{2k-1} & 0 & \ddots \end{bmatrix}$$

Si $m = 2p$, il y a p blocs ainsi ; si $m = 2p+1$, il y a une ligne et une colonne de zéros en plus et les P_{2k-1} sont des fonctions algébriques simples des ξ_{k1} . (voir S. Lang [3]).

2.5. LEMME. Si on pose $G_s(u, \xi) = \prod_{k=1}^{\infty} \psi_s(u, \frac{\xi}{k})$,

$$G_s(u, \xi) = \prod_{l=1}^p (\frac{s}{2} P_{2l-1}) (\text{sh}(\frac{s}{2} P_{2l-1}))^{-1} \\ \times \exp \left[\frac{({}^t W u)_{2l-1}^2 + ({}^t W u)_{2l}^2}{2s} (1 - \frac{s}{2} P_{2l-1} \coth(\frac{s}{2} P_{2l-1})) \right]$$

avec $p = [\frac{m}{2}]$.

Preuve :

$$G_s(u, \xi) = \prod_{k=1}^{\infty} \exp \left(\frac{1}{s} (-|u|^2 + {}^t u (I - \frac{s^2 \xi^2}{4\pi^2 k})^{-1} u) \right) [\det(I - \frac{s\xi}{2\pi k})]^{-1}.$$

En remarquant que si $\xi = WP {}^t W$,

$t_u(I - \frac{s^2 \xi^2}{4\pi^2 k^2})^{-1} u = t_u W(I - \frac{s^2 P^2}{4\pi^2 k^2})^{-1} t_{Wu}$ et que $(I - \frac{s^2 \xi^2}{4\pi^2 k^2})$ est une matrice diagonale qui a comme éléments diagonaux $1 - \frac{s^2 P_{21-1}^2}{4\pi^2 k^2 + s^2 P_{21-1}^2}$ comptés deux fois si $m = 2p$, et avec un 1 en plus si $m = 2p+1$, on voit que dans tous les cas :

$$t(t_{Wu})(I - \frac{s^2 P^2}{4\pi^2 k^2})^{-1} t_{Wu} = - \sum_{l=1}^p \frac{s^2 P_{21-1}^2}{4\pi^2 k^2 + s^2 P_{21-1}^2} [(t_{Wu})_{21-1}^2 + (t_{Wu})_{21}^2].$$

D'autre part, $(I - \frac{s\xi}{2\pi k})$ est une matrice formée de blocs sur la diagonale du type suivant :

$$\begin{bmatrix} 1 & -\frac{sP_{21-1}}{2\pi k} \\ \frac{sP_{21-1}}{2\pi k} & 1 \end{bmatrix}$$

$$D'où : \det(I - \frac{s\xi}{2\pi k})^{-1} = \prod_{l=1}^p \frac{4\pi^2 k^2}{4\pi^2 k^2 + s^2 P_{21-1}^2}.$$

Ce qui donne pour G_s :

$$G_s(u, \xi) = \prod_{k=1}^{\infty} [\exp(\frac{1}{s}(- \sum_{l=1}^p \frac{s^2 P_{21-1}^2}{4\pi^2 k^2 + s^2 P_{21-1}^2} [(t_{Wu})_{21-1}^2 + (t_{Wu})_{21}^2])] \times \prod_{l=1}^p \frac{4\pi^2 k^2}{4\pi^2 k^2 + s^2 P_{21-1}^2}.$$

En utilisant les deux formules (voir H. Cartan [1] p.162),

$$\frac{x}{\sinh x} = \prod_{k=1}^{\infty} \frac{k^2 \pi^2}{k^2 \pi^2 + x^2}$$

$$\text{et } x \coth x - 1 = x^2 \sum_{k=1}^{\infty} \frac{2k^2 \pi^2}{k^2 \pi^2 + x^2},$$

il vient, après avoir échangé le produit infini et l'exponentielle (on a posé $x = \frac{s}{2} P_{21-1}$) :

$$G_s(u, \xi) = \prod_{l=1}^p \left(\frac{s}{2} P_{2l-1} \right) \left(\operatorname{sh} \frac{s}{2} P_{2l-1} \right)^{-1} \\ \times \exp \left(\frac{1}{s} \left[({}^t W u)_{2l-1}^2 + ({}^t W u)_{2l}^2 \right] \left(1 - \frac{s}{2} P_{2l-1} \coth \left(\frac{s}{2} P_{2l-1} \right) \right) \right).$$

Signalons une démonstration de M. Yor [5] dans le cas $m = 2$, de la formule de l'espérance conditionnelle à l'aide du calcul stochastique.

2.6. PROPOSITION. Dans les coordonnées exponentielles, la densité $p_s(0, g)$ vaut :

[1] Si $m = 2p$,

$$F_s(w) = (2\pi)^{-\frac{m}{2} - \frac{m(m-1)}{2}} s^{-\frac{m}{2}} \int_{\mathbb{R}} \frac{m(m-1)}{2} \exp[-i\langle \xi, v \rangle] \prod_{k=1}^p \phi_s^k(u, \xi) d\xi$$

où

$$\phi_s^k(u, \xi) = \frac{s}{2} P_{2k-1} \left(\operatorname{sh} \frac{s}{2} P_{2k-1} \right)^{-1} \exp \left\{ \left[- \frac{({}^t W u)_{2k-1}^2 + ({}^t W u)_{2k}^2}{2s} \right] \frac{s}{2} P_{2k-1} \coth \frac{s}{2} P_{2k-1} \right\}$$

[2] Si $m = 2p+1$

$$F_s(w) = (2\pi)^{-\frac{m}{2} - \frac{m(m-1)}{2}} s^{-\frac{m}{2}} \int_{\mathbb{R}} \frac{m(m-2)}{2} \exp[-i\langle \xi, v \rangle] \times \exp - \frac{({}^t W u)_{2p+1}^2}{2s} \prod_{k=1}^p \phi_s^k(u, \xi) d\xi$$

Preuve :

Pour revenir à F , il faut voir que l'on a calculé seulement la transformée de Fourier en les variables de $V_2 \propto \mathbb{R}^{\frac{m(m-1)}{2}}$, soit :

$$g_s(\xi) = E_0 \left(\exp i \sum_{k \leq 1} \xi_{k1} \beta_{k1} \right) \\ = \prod_{l=1}^p \frac{s}{2} P_{2l-1} \left(\operatorname{sh} \frac{s}{2} P_{2l-1} \right)^{-1} \exp \left\{ \left[\frac{({}^t W u)_{2l-1}^2 + ({}^t W u)_{2l}^2}{2s} \right] \right. \\ \left. \times \left(1 - \frac{s}{2} P_{2l-1} \coth \left(\frac{s}{2} P_{2l-1} \right) \right) \right\}.$$

g est dans L_1 car $\sum P_i^2 = \text{trace } P^*P = \text{trace } \xi^*\xi = \sum_{k \leq 1} \xi_{k1}^2$ donc chaque P_{2k-1} se comporte comme $\sqrt{\xi_{k1}^2}$ et pour conclure on utilise le fait que $u \coth u \leq Cu$ et que ue^{-u} est dans L_1 .

D'où,

$$F_s(w) = (2\pi s)^{-\frac{m}{2}} (2\pi)^{-\frac{m(m-1)}{2}} \int_{\mathbb{R}} \frac{m(m-1)}{2} \exp(-i\langle \xi, v \rangle) \exp - \frac{|u|^2}{2s} \times$$

$$\prod_{1}^{\frac{s}{2}} P_{2k-1} (\text{sh } \frac{s}{2} P_{2k-1})^{-1} \exp \left[- \frac{(\tau_w u)_{2k-1}^2 + (\tau_w u)_{2k}^2}{2s} \right]$$

$$\times (1 - \frac{s}{2} P_{2k-1} \coth \frac{s}{2} P_{2k-1}) \}.$$

On obtient alors la formule annoncée en remarquant que $-\frac{|u|^2}{2s} + \sum \frac{(\tau_w u)_{2k-1}^2 + (\tau_w u)_{2k}^2}{2s}$ vaut zéro ou $-\frac{(\tau_w u)_{2p+1}^2}{2s}$, suivant que m est pair ou non.

2.7. Le cas du groupe d'Heisenberg.

La formule de la proposition 2.6 lorsque $m = 2$ donne la densité de transition pour le groupe d'Heisenberg (voir R. Azencott chapitre 13). Il suffit de remarquer que pour H_3 on a pris $[X, Y] = -4T$ et non comme ici $[X_i, X_j] = X_{ij}$.

On trouve tout d'abord :

$$p_s(0, (x, y, z)) = (2\pi)^{-2} s^{-1} \int_{\mathbb{R}} \exp \left[-i\xi v - \frac{x^2 + y^2}{2s} \frac{\frac{s}{2} \xi}{\text{th } \frac{s}{2} \xi} \right] \times \frac{\frac{s}{2} \xi}{\text{sh } \frac{s}{2} \xi} d\xi.$$

On change alors $\frac{s}{2} \xi$ en $2r$ et $-4v$ en z ce qui donne

$$p_s(0, (x, y, z)) = \frac{1}{(\pi s)^2} \int_{\mathbb{R}} \frac{2r}{\text{sh } 2r} \exp \left[i r \frac{z}{s} - \frac{x^2 + y^2}{2s} \frac{2r}{\text{th } 2r} \right] dr$$

BIBLIOGRAPHIE

- [1] H. Cartan, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Hermann 1961.
- [2] B. Gaveau, *Principe de moindre action. Propagation de la chaleur et estimées sous-elliptiques sur certains groupes nilpotents*. Acta Mathematica vol 139 (1977) p.95-153.

- [3] S. Lang, *Algebra*, Addison Wesley 1965.
- [4] P. Levy, *Wiener's random functions, and other laplacian random fonctions*, 2d
Berkeley symposium of probability and statistics (1950) p.171-186.
- [5] M. Yor, *Remarques sur une formule de P. Levy*, à paraître.

CHAPITRE 12

TRANSFORMÉES DE CRAMER, DIFFUSIONS EN TEMPS PETIT, ET ACTION MINIMALE

par

Robert AZENCOTT

1. Petites perturbations et diffusions en temps petit.

1.1. Introduction. Nous allons résumer sans démonstration une approche que nous avons développée dans Azencott [1] ch. III et ch. V, où nous étudions le comportement des diffusions en temps petit à partir d'une théorie "à la Ventcell-Freidlin" des petites perturbations de systèmes dynamiques. L'idée n'est pas neuve et a déjà été utilisée par Varadhan [6] et Gaveau [4]; cependant dans [1] nous avons sérieusement amélioré les résultats du type Ventcell-Freidlin sur le plan technique, ce qui permet d'englober le cas des opérateurs différentiels hypo-elliptiques sur une variété.

Dans un problème de petites perturbations de système dynamique, on considère sur une variété M un opérateur différentiel L semi-elliptique, d'ordre 2, vérifiant $L1 = 0$, et un champ de vecteurs b sur M . Pour ε tendant vers 0 on s'intéresse aux D_ε -diffusions où D_ε est "une petite perturbation" de l'opérateur différentiel d'ordre 1 défini par b , de la forme précise suivante :

$$D_\varepsilon = \varepsilon^2 L + b_\varepsilon$$

où les champs de vecteurs b_ε convergent (uniformément sur tout compact) vers b quand $\varepsilon \rightarrow 0$. On notera δ le point à l'infini de M .

Grâce à certaines hypothèses supplémentaires sur (L, b) , qui englobent les cas essentiels 1.2 et 1.3 ci-dessous, on peut alors définir sur l'espace $\mathcal{C}(M)$ des trajec-

toires (explosives éventuellement) définies sur $[0,1]$ à valeurs dans le compactifié $M \cup \delta$ de M et continues jusqu'au temps d'atteinte de δ , une fonctionnelle λ à valeurs dans $[0,+\infty]$, qui est complètement déterminée par (L,b) , et que nous appelons transformée de Cramer de (L,b) ; cette fonctionnelle a la propriété suivante : pour les parties boréliennes A de $\mathcal{Z}(M)$ "assez régulières" la trajectoire aléatoire $y^\varepsilon \in \mathcal{Z}(M)$ de la D_ε -diffusion vérifie $\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) = - \inf_{f \in A} \lambda(f)$.

En fait les théorèmes précis (cf. Azencott [1]) fournissent un encadrement de la "limite" considérée ci-dessus, par les nombres $-\Lambda(A)$ et $-\Lambda(\bar{A})$, où l'on a posé

$$\Lambda(A) = \inf_{f \in A} \lambda(f).$$

Un (très) simple changement de temps permet de ramener l'étude d'une diffusion en temps petit à la situation précédente avec en fait $b \equiv 0$, comme nous le précisons au paragraphe 3. Notons cependant que les résultats sont un peu bruts puisque (par exemple) on atteint ainsi une évaluation précise de $\lim_{t \rightarrow 0} t \log P_x(y_t \in F)$, où $F \subset M$, $x \in M$, et y_t est la position au temps t d'une diffusion donnée (fixe) sur M . Ceci ne permet donc pas d'écrire un équivalent précis de $P_x(y_t \in F)$, et encore moins de $p(t,x,y)$. Par contre la méthode a l'avantage de fournir une bonne description géométrique de ce que "doit"être le terme principal d'un équivalent éventuel de $p(t,x,y)$ lorsque $t \rightarrow 0$.

Nous allons étudier essentiellement deux cas, d'après Azencott [1]. Soit M une variété différentiable connexe, de dimension n , de classe ≥ 2 . Soit L un opérateur différentiel d'ordre 2 sur M , semi-elliptique, tel que $L1 = 0$.

1.2. Cas elliptique.

L'opérateur L est elliptique et à coefficients localement lipschitziens ; de plus, il existe un atlas de M tel que pour toute carte ϕ de cet atlas, la matrice $[\frac{1}{2} a(x)]$ des coefficients du second ordre pour l'opérateur L lu dans la carte ϕ , vérifie $a(x) = \sigma(x) \sigma^*(x)$ où $\sigma(x)$ est un champ de matrices rectangulaires, de classe C^1 en x .

1.3. Cas hypo-elliptique.

La variété M et l'opérateur L sont de classe C^∞ ; l'opérateur L vérifie localement la condition d'hypoellipticité de Hörmander (cf. Azencott ch. 2) ; enfin, dans

toute carte locale, le rang de la matrice des coefficients du second ordre de L reste constant.

2. Transformée de Cramer de L .

Renvoyons à [1] pour les preuves des assertions ci-dessous. Plaçons nous dans le cas elliptique 1.2 ou bien dans le cas hypoelliptique 1.3.

Pour f de classe 2 au voisinage de $x \in M$, l'expression $[L(f^2) - 2f Lf]$ calculée en x ne dépend que de df_x . On définit alors une forme quadratique positive (sens large) Q_x sur l'espace cotangent $T_x^*(M)$ en posant pour $w \in T_x^*(M)$

$$(1) \quad Q_x(w) = [L(f^2) - 2f Lf]_x$$

où f est une fonction quelconque de classe 2 vérifiant $df_x = w$. En coordonnées locales, si on note $[\frac{1}{2} a(x)]$ la matrice des coefficients du second ordre de L , alors Q_x est simplement la forme quadratique définie par la matrice $a(x)$. En particulier Q_x est de classe C^1 .

Définissons le champ des "formes quadratiques duales" $Q_x^* : T_x(M) \rightarrow [0, +\infty]$ par

$$(2) \quad Q_x^*(v) = \sup_{w \in T_x^*(M)} [2\langle v, w \rangle_x - Q_x(w)] \text{ pour } v \in T_x(M)$$

où $\langle \cdot, \cdot \rangle_x$ est la forme bilinéaire naturelle sur $T_x(M) \times T_x^*(M)$.

En coordonnées locales, on vérifie facilement que si $a(x)$ est inversible, alors Q_x^* est définie par la matrice $a(x)^{-1}$. Dans le cas général où $a(x)$ est non inversible, Q_x^* peut prendre la valeur $+\infty$, et en fait, l'ensemble

$$(3) \quad H_x(M) = \{v \in T_x(M) \mid Q_x^*(v) \text{ fini}\}$$

coïncide (en coordonnées locales) avec l'image de $\mathbb{R}^n$ par $a(x)$. Nous appelons ce sous espace le sous espace horizontal de $T_x(M)$ en x (terminologie qui généralise convenablement celle de Gaveau [3][4]).

La fonction $Q^* : T(M) \rightarrow [0, +\infty]$ définie par $(x, v) \rightarrow Q_x^*(v)$ est alors semi-continue inférieurement.

Soit $M \cup \delta$ le compactifié d'Alexandroff de M (avec δ isolé si M est compacte). Soit $\mathcal{E}(M)$ l'espace des trajectoires explosives à valeurs dans $M \cup \delta$, définies sur $[0, 1]$; $\mathcal{E}(M)$ est l'ensemble des fonctions continues γ sur $[0, 1]$ à valeurs dans $M \cup \delta$

telles que si le temps d'atteinte $\zeta(\gamma)$ de δ est plus petit que 1, alors γ reste en δ sur $[\zeta(\gamma), 1]$. Le temps d'atteinte de δ noté $\zeta(\gamma)$ est alors appelé temps d'explosion de γ . L'espace $\mathcal{Z}(M)$ est muni de la topologie suivante : une suite f_n converge vers f dans $\mathcal{Z}(M)$ si pour tout temps $t \leq 1$ vérifiant $t < \zeta(f)$, la suite f_n converge uniformément vers f sur $[0, t]$.

Soit $f \in \mathcal{Z}(M)$. Si f'_t existe (au sens de Lebesgue) pour presque tout $t \in [0, 1 \wedge \zeta(f)]$, nous posons

$$(4) \quad \lambda(f) = \frac{1}{2} \int_0^{1 \wedge \zeta(f)} Q_{f_t}^*(f'_t) dt$$

Si f n'est pas absolument continue sur $[0, 1 \wedge \zeta(f)]$, nous posons $\lambda(f) = +\infty$.

La fonctionnelle $\lambda : \mathcal{Z}(M) \rightarrow [0, +\infty]$ est alors appelée la transformée de Cramer de L . On peut montrer (cf. [1] ch. III et V) que λ est semi-continue inférieurement sur $\mathcal{Z}(M)$, et que pour tout $k \geq 0$ fini, l'ensemble des $f \in \mathcal{Z}(M)$ telles que $\lambda(f) \leq k$ est compact.

Nous définissons la fonctionnelle de Cramer Λ associée à L par

$$(5) \quad \Lambda(A) = \inf_{f \in A} \lambda(f) \quad \text{où } A \in \mathcal{Z}(M) \text{ est arbitraire.}$$

Remarquons que comme λ, Λ sont déterminées par Q^* et donc par les formes quadratiques Q_x , deux opérateurs L et $\tilde{L}$ tel que $\tilde{L} = L + Z$, où Z est un champ de vecteurs sur M , auront même transformée de Cramer λ .

3. Comportement asymptotique des diffusions en temps petit.

3.1. THÉOREME. (Azencott [1] ch. V). *Considérons le cas elliptique 1.2 ou bien le cas hypo-elliptique 1.3. Soit (y_t) l'unique L -diffusion sur M . Notons Λ la fonctionnelle de Cramer (cf. 2.(5)) associée à L . Soit $\mathcal{Z}(M)$ l'espace des trajectoires explosives à valeurs dans M , et posons pour $x \in M$, $\mathcal{Z}_x(M) = \{f \in \mathcal{Z}(M) \mid f_0 = x\}$.*

Pour chaque $\varepsilon > 0$ définissons le processus z^ε à valeurs dans $\mathcal{Z}(M)$ par

$$(1) \quad z_t^\varepsilon = y_{\varepsilon t} \quad 0 \leq t \leq 1$$

Alors pour toute partie borélienne A de $\mathcal{Z}_x(M)$, on a

$$(2) \quad -\Lambda(A) \leq \lim_{\varepsilon \rightarrow 0} \varepsilon \log P_x(z^\varepsilon \in A) \leq \overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon \log P_x(z^\varepsilon \in A) \leq -\Lambda(\bar{A})$$

où $\bar{A}, \bar{A}$ sont resp. l'intérieur et l'adhérence de A dans $\mathcal{Z}_x(M)$.

3.2. COROLLAIRE. *Mêmes hypothèses qu'en 3.1. Soit (y_t) l'unique L-diffusion sur M, et soit λ la transformée de Cramer associée à L. Pour $x, y \in M$ définissons $E(x, y) \in [0, +\infty]$ par*

$$(3) \quad E(x, y) = \inf \{ \lambda(f) \mid f \text{ chemin continu, } f_0 = x, f_1 = y \}$$

Alors pour toute partie borélienne F de M on a

$$(4) \quad - \inf_{z \in F} E(x, z) \leq \lim_{t \rightarrow 0} t \log P_x(y_t \in F) \leq \overline{\lim}_{t \rightarrow 0} t \log P_x(y_t \in F) \leq - \inf_{z \in \overline{F \cup \delta}} E(x, z)$$

Preuve : (cf. [1] ch. V). Pour F borélien de M, on note

$A_F = \{f \in \mathcal{X}(M) \mid f_0 = x, f_1 \in F\}$. On vérifie que A_F est un borélien de $\mathcal{X}(M)$, que $\overline{A_F} \subset A_{\overline{F \cup \delta}}$ et que $A_{\overline{F \cup \delta}} \subset (A_F)^\circ$. L'évènement $\{y_t \in F\}$ peut aussi s'écrire $\{z^t \in A_F\}$ avec

les notations de 3.1, où l'on prend $\varepsilon = t$. Il suffit alors d'appliquer (2) à A_F et de remarquer que les définitions (3) et 2.(5) impliquent $\Lambda(A_F) = \inf_{z \in F} E(x, z)$ quel que soit $F \subset M$.

3.3. Interprétation dans le cas elliptique : énergie minimale.

Plaçons nous dans le cas elliptique 1.2. En coordonnées locales, notons $[\frac{1}{2} a(x)]$ la matrice des coefficients du second ordre de L. Comme $a(x)$ est par hypothèse inversible, la forme quadratique duale Q_x^* introduite en 2.(2) est en fait définie par la matrice $a(x)^{-1}$. Par suite Q_x^* définit une métrique riemannienne sur M, qui coïncide précisément avec la métrique associée à L par Molchanov [5] (cf. L. Elie ch. 5). D'autre part l'énergie d'un chemin absolument continu $f : [0, 1] \rightarrow M$ est, pour cette métrique, définie par (cf. A. Bellaïche ch. 7)

$$(1) \quad E(f) = \frac{1}{2} \int_0^1 \|f'_t\|_{f_t}^2 dt$$

où $\| \cdot \|_x$ est la métrique en x.

Si λ est la transformée de Cramer de L, la définition 2.(4) donne donc

$$(2) \quad \lambda(f) = E(f) \text{ pour } f \text{ chemin continu.}$$

En comparant la relation classique (cf. A. Bellaïche ch.7)

$$(3) \quad \inf \{E(f) \mid f_0 = x, f_1 = y\} = \frac{1}{2} d^2(x, y)$$

avec les relations (2) et 3.2.(3) on voit que dans le cas elliptique 1.2, on a

$$(4) \quad E(x,y) = \frac{1}{2} d^2(x,y)$$

En particulier si F est une partie borélienne de M telle que $\overset{\circ}{F}$ et F aient même adhérence dans M , et vérifiant $d(x,\overset{\circ}{F}) < d(x,\delta)$ le corollaire 3.2. montre immédiatement que

$$(5) \quad \lim_{t \rightarrow 0} t \log P_x(y_t \in F) = - \frac{1}{2} d^2(x,F)$$

3.4. L'interprétation de $E(x,y)$ dans le cas hypo-elliptique est plus délicate à préciser ; nous allons la relier à l'action des bicaractéristiques de L , sans pouvoir conclure aussi agréablement que dans le cas elliptique. Bien entendu l'intérêt de cette interprétation vient de ce que au vu du corollaire 3.2 et de 3.3.(5) par exemple, il est plausible que dans toutes les bonnes (??) situations hypoelliptiques on puisse prouver que, pour $E(x,y)$ fini et $E(x,y) < E(x,\delta)$,

$$(1) \quad " \lim_{t \rightarrow 0} t \log p(t,x,y) = - E(x,y) "$$

Nous reviendrons sur cette question dans la suite.

4. Bicaractéristiques, action, et transformée de Cramer.

4.1. Cas elliptique : géodésiques et problème d'extrémum.

Pour la présentation détaillée du problème d'extrémum considérée ici, renvoyons à A. Bellaïche ch. 7. Plaçons nous dans le cas elliptique 1.2.

Une courbe $\gamma : [0,1] \longrightarrow M$ est une géodésique si et seulement si il existe une courbe (γ, q) définie sur $[0,1]$, à valeurs dans $T^*(M)$ telle que $(\gamma, q) = \phi$ soit une courbe extrémale de

$$(1) \quad J(\phi) = \int_0^1 [\langle \gamma'_t, q_t \rangle_{\gamma_t} - \frac{1}{2} Q_{\gamma_t}(q_t)] dt$$

intégrale que nous noterons

$$(2) \quad J(\phi) = \int_0^1 F(\phi_t, \phi'_t) dt.$$

En coordonnées locales, la fonction q "est à valeurs dans $\mathbb{R}^n$ " et la fonction $F(\phi_t, \phi'_t)$ s'écrit

$$(3) \quad F(\phi_t, \phi'_t) = \gamma_t^* q_t - \frac{1}{2} q_t^* a_{\gamma_t} q_t$$

où $[\frac{1}{2}a(x)]$ est la matrice des coefficients du second ordre de L.

Les extrémales ϕ de $J(\phi)$ vérifient l'équation d'Euler usuelle $\frac{d}{dt}(\frac{\partial F}{\partial \phi'}) - \frac{\partial F}{\partial \phi} = 0$, qui s'écrit en coordonnées locales, avec $\phi = (\gamma, q)$,

$$(4) \quad \begin{cases} \gamma'_t = a_{\gamma t} q_t \\ q'_t = -\frac{1}{2} q_t^* \left[\frac{\partial(a_{\gamma})}{\partial \gamma} \right]_{\gamma_t} q_t \end{cases}$$

En particulier le long d'une extrémale ϕ , $F(\phi, \phi')$ s'écrit grâce à (3) et à la 1ère relation (4)

$$(5) \quad F(\phi, \phi') = \frac{1}{2} q^* a_{\gamma} q = \frac{1}{2} \gamma'^* a_{\gamma}^{-1} \gamma'$$

d'où, en tenant compte de ce que γ est géodésique,

$$(6) \quad J(\phi) = \frac{1}{2} \int_0^1 \|\gamma'_t\|_{\gamma_t}^2 = \frac{1}{2} [\text{longueur}(\gamma)]^2$$

4.2. Cas hypo-elliptique : bicaractéristiques et action.

Plaçons nous dans le cas hypelliptique 1.2. La matrice $[\frac{1}{2} a(x)]$ associée à L en coordonnées locales est non inversible en général. Le problème d'extrémum 4.1.(1) conserve toujours un sens cependant, et l'équation d'Euler 4.1.(4) des extrémales $\phi = (\gamma, q)$ de $J(\phi)$ s'écrit comme dans la situation précédente en coordonnées locales. De plus le long d'une extrémale ϕ de $J(\phi)$, on peut encore écrire (coordonnées locales), grâce à 4.1.(3), 4.1.(4)

$$(1) \quad F(\phi, \phi') = \frac{1}{2} \gamma'^* q = \frac{1}{2} q^* a_{\gamma} q \geq 0.$$

Les courbes $\phi = (\gamma, q)$ extrémales de $J(\phi)$ sont appelées bicaractéristiques de L et $J(\phi)$ est appelée l'action de la bicaractéristique ϕ (pour plus de détails sur ces objets, renvoyons à Gaveau [3][4] et Courant-Hilbert [2]).

4.3. LEMME. (L. Elie, M. Chaleyat-Maurel). Soient M une variété, L un opérateur différentiel d'ordre 2 sur M, vérifiant les hypothèses du cas hypoelliptique 1.3. Soit λ la transformée de Cramer de L et soit $\phi = (\gamma, q)$ une bicaractéristique de L sur l'intervalle de temps $[0, 1]$. Notons $J(\phi)$ l'action de ϕ . Alors on a la relation $\lambda(\gamma) = J(\phi)$.

Preuve : Sur $\mathbb{R}^m$ soit Q une forme quadratique positive au sens large, définie par une matrice $a = \sigma \sigma^*$, avec σ rectangulaire (n, k) . La forme duale $Q^* : \mathbb{R}^n \longrightarrow [0, +\infty]$ est donnée par

$$(1) \quad Q^*(v) = \sup_{w \in \mathbb{R}^n} [2\langle v, w \rangle - Q(w)], \text{ pour } v \in \mathbb{R}^n.$$

De plus on a (cf. Azencott [1] ch. III)

$$(2) \quad Q^*(v) = \inf \{ \|w\|^2 \mid w \in \mathbb{R}^k \text{ et } \sigma w = v \} \text{ pour } v \in \mathbb{R}^n.$$

Considérons en particulier $u, v \in \mathbb{R}^n$ liés par $v = au$. Appliquons (2) à $w = \sigma^* u$ pour obtenir

$$(3) \quad Q^*(v) \leq \|w\|^2 = \|\sigma^* u\|^2 = u^* \sigma \sigma^* u = u^* au.$$

D'autre part (1) implique

$$(4) \quad Q^*(v) \geq 2 v^* u - Q(u) = 2v^* u - u^* au = u^* au.$$

De (3) (4) on déduit

$$(5) \quad Q^*(v) = u^* au \text{ si } u, v \in \mathbb{R}^n \text{ vérifiant } v = au.$$

Appliquons ce résultat à $v = \gamma'_t$, $Q = Q_{\gamma_t}$, $u = q_t$ pour conclure, grâce à 4.2.(1), que si $\phi = (\gamma, q)$ est extrémale de $J(\phi)$ on a

$$(6) \quad F(\phi, \phi') = \frac{1}{2} Q_{\gamma}^*(\gamma')$$

d'où

$$J(\phi) = \frac{1}{2} \int_0^1 Q_{\gamma_t}^*(\gamma'_t) dt = \lambda(\gamma).$$

Remarque : Notons en passant que les hypothèses sur L utilisées ici sont en fait bien plus faibles que celles du cas hypoelliptique 1.3.

4. . Cas hypoelliptique : action minimale et transformée de Cramer.

Plaçons nous dans le cas hypoelliptique 1.3. La quantité $E(x, y)$ qui représente par définition (cf. 3.2.(3)) le λ minimal pour aller de x à y en temps unité peut être reliée à l'action minimale nécessaire pour passer de x à y (en temps unité) le long d'une bicaractéristique, mais seulement par une inégalité. En effet le lemme 4.4. et la définition 3.2.(3) donnent immédiatement, pour $x, y \in M$

$$(1) \quad E(x,y) \leq \inf \{J(\phi) \mid \phi = (\gamma, q) \text{ est bicaractéristique, } \gamma_0 = x, \gamma_1 = y\}.$$

Les propriétés de semi-continuité et d'inf. compacité de λ rappelées au §2 montrent qu'il existe toujours un chemin absolument continu γ tel que

$$(2) \quad E(x,y) = \lambda(\gamma) \quad \gamma_0 = x, \gamma_1 = y$$

Mais γ n'est pas nécessairement C^1 par morceaux et à fortiori γ n'est pas nécessairement la projection sur M d'une bicaractéristique (γ, q) . En effet comme l'a remarqué Gaveau [3] il n'existe pas nécessairement (en général) de bicaractéristique (γ, q) vérifiant $\gamma_0 = x, \gamma_1 = y$ (cf. M. Chaleyat-Maurel ch.11). L'inégalité (1) peut donc fort bien être stricte. Comme la quantité liée au comportement de $t \log p(t,x,y)$ quand t tend vers 0 est évidemment $E(x,y)$ - cf. 3.4 -, il semble plus prometteur de chercher les liens de $\lim_{t \rightarrow 0} t \log p(t,x,y)$ avec $E(x,y)$ plutôt qu'avec l'action minimale des bicaractéristique joignant x à y en temps unité.

Il serait intéressant aussi de déterminer des critères garantissant l'égalité de $E(x,y)$ et de $\inf J(\phi)$ lorsque ϕ parcourt l'ensemble des bicaractéristiques brisées (en un nombre fini de points) joignant x à y en temps unité.

Nous allons détailler dans l'exposé suivant un exemple étudié par Gaveau [3], pour lequel on a en fait égalité dans (1), au moins pour des points x,y génériques. Rappelons que cette situation semble cependant être plutôt l'exception que la règle (cf. M. Chaleyat-Maurel et L. Elie ch. 14) lorsque l'hypoellipticité de L n'est pas garantie par sa seule "portion" auto-adjointe.

En fait l'étude précédente n'est satisfaisante que lorsque $E(x,y)$ est fini car elle fournit (ou suggère) des équivalents du type $[-\frac{E(x,y)}{t}]$ pour $\log p(t,x,y)$

(cf. [1] ch. VI). Lorsque $L = \sum_{i=1}^k X_i^2 + Y$, où les X_i, Y sont des champs de vecteurs C^∞ ,

et lorsque l'algèbre de Lie engendrée par $\{X_1 \dots X_k\}$ est, en chaque point de M , de dimension m égale à celle de M , un résultat de Bony garantit que pour chaque $x \in M$, $E(x,y)$ est fini au moins pour y dans une partie dense de M .

Par contre si on n'est pas dans ce cas, ce qui n'empêche pas L d'être hypoelliptique au sens de Hörmander, $E(x,y)$ sera au contraire presque partout infini pour x donné (au moins si m reste constant sur M). Il faut alors modifier convenablement la définition de l'action et de la transformée de Cramer pour obtenir des résultats précis, par exemple sur des limites du genre $\lim_{t \rightarrow 0} t^q \log \int_V p(t,x,y) dy$ avec V ouvert relativement compact. Le travail de M. Chaleyat-Maurel et L. Elie (ch. 14) traite ce problème dans le cadre des diffusions gaussiennes.

Dans les exemples étudiés par Gaveau [3], on arrive par un calcul direct de $p(t, x, y)$, à obtenir un équivalent précis de $p(t, x, y)$ pour $t \rightarrow 0$, et en particulier à calculer $\lim_{t \rightarrow 0} t \log p(t, x, y)$. La proposition suivante s'appliquera en particulier à ces exemples.

4.6. PROPOSITION. Soit M une variété, et soit L un opérateur différentiel sur M . Plaçons nous dans le cas hypoelliptique 1.3, ou bien dans le cas elliptique 1.2. Soit $E(x, y)$ le minimum de λ sur les chemins joignant x à y en temps unité, où λ est la transformée de Cramer de L .

Fixons un point x et un ouvert U de M . Soit $p(t, x, y)$ la densité de la L -diffusion sur M . Supposons que pour tout $y \in U$, la limite $\lim_{t \rightarrow 0} t \log p(t, x, y) = h(x, y)$ existe est continue en y , et que la convergence est uniforme sur les compacts de U . Alors on a $h(x, y) \equiv E(x, y)$ pour tout $y \in U$.

Preuve : Soit K un compact de U . L'hypothèse de convergence uniforme fournit l'existence de $A(t)$ tels que $\lim_{t \rightarrow 0} t \log A(t) = \lim_{t \rightarrow 0} t \log B(t) = 0$ et vérifiant pour $y \in K$

$$A(t) \exp \left[-\frac{1}{t} \sup_{y \in K} h(x, y) \right] \leq p(t, x, y) \leq B(t) \exp \left[-\frac{1}{t} \inf_{y \in K} h(x, y) \right]$$

Si (R_t) est le semi-groupe de transition de la L -diffusion, on a donc, si K est non négligeable pour la mesure de référence sur M ,

$$(1) \quad - \sup_{y \in K} h(x, y) \leq \varliminf_{t \rightarrow 0} t \log R_t^1 K(x) \leq \overline{\lim}_{t \rightarrow 0} t \log R_t^1 K(x) \leq - \inf_{y \in K} h(x, y)$$

Appliquons ce résultat à un voisinage compact V de l'ensemble $\tilde{K}$ des points de K où $h(x, \cdot)$ atteint la valeur $\inf_{y \in K} h(x, y)$, ainsi qu'à $K-V$. Ceci est possible si V et $K-V$ sont non négligeables, ce qui est sûrement réalisé lorsqu'on impose $K = \overline{\tilde{K}}$, et lorsque $h(x, \cdot)$ n'est pas constante sur K .

La relation (1) montre alors que $t \log R_t^1 K(x)$ et $t \log R_t^1 V(x)$ sont équivalents lorsque $t \rightarrow 0$, et par suite que les $\varliminf_{t \rightarrow 0}$ et $\overline{\lim}_{t \rightarrow 0}$ de $t \log R_t^1 K(x)$ sont dans l'intervalle $[- \sup_{y \in V} h(x, y), - \inf_{y \in V} h(x, y)]$. Quand V décroît vers $\tilde{K}$, on conclut donc que

$$(2) \quad \lim_{t \rightarrow 0} t \log R_t^1 K(x) = - \inf_{y \in K} h(x, y)$$

Ce résultat est encore vrai, grâce à (1) lorsque $h(x, \cdot)$ est constante sur K , de sorte que (2) est vraie dès que K est un compact de U vérifiant $K = \bar{K}$.

D'après le corollaire 3.2 qui fournit un encadrement de $\lim_{t \rightarrow 0} t \log R_t^1 K(x)$, (2) implique

$$(3) \quad \inf_{y \in K} E(xy) \leq \inf_{y \in K} h(x, y) \leq \inf_{y \in K} E(x, y)$$

Fixons un point $z \in U$ et soit K_m une suite de voisinages compacts de z , décroissant vers z . La seconde inégalité dans (3) donne $\inf_{y \in K_m} h(x, y) \leq E(x, z)$, d'où quand

$m \rightarrow +\infty$,

$$(4) \quad h(x, z) \leq E(x, z), \quad z \in U$$

Pour $\varepsilon > 0$ fixé, la première inégalité dans (3) montre l'existence d'une suite $y_m \in K_m$ telle que, pour m assez grand on ait

$$(5) \quad E(x, y_m) \leq h(x, z) + \varepsilon.$$

On a vu (cf. 4.5.(2)) qu'il existe un chemin continu γ_m joignant x à y_m en temps unité et vérifiant $E(x, y_m) = \lambda(\gamma_m)$. Comme (5) force $\lambda(\gamma_m)$ à rester borné, il existe une suite extraite de γ_m (encore notée γ_m) convergeant uniformément sur $[0, 1]$ vers un chemin continu γ (voir §2). On a alors $\lambda(\gamma) \leq \lim_{m \rightarrow \infty} \lambda(\gamma_m)$, et $\gamma(1) = \lim_{m \rightarrow \infty} y_m = z$. A fortiori, on peut majorer $E(x, z)$ par $\lambda(\gamma)$, donc par $\lim_{m \rightarrow \infty} E(x, y_m)$ et par suite (5) donne

$$(6) \quad E(x, z) \leq h(x, z) + \varepsilon.$$

Comme (6) est vraie pour tout $\varepsilon > 0$, une comparaison avec (4) entraîne $E(x, z) = h(x, z)$ pour tout $z \in U$.

4.7. Remarque : Nous avons démontré au passage que si L vérifie les hypothèses du cas hypoelliptique 1.3, ou bien celles du cas elliptique 1.2, la fonction $E(x, y)$ est s.c.i en $y \in M$. La même démonstration prouve que $E(x, y)$ est en fait s.c.i en tant que fonction de (x, y) .

4.8. Remarque : Dans la proposition 4.6 si nous avons supposé $h(x,y)$ non pas continue mais seulement s.c.i. en $y \in U$, nous aurions conclu que

$$E(x,y) = \lim_{z \rightarrow y} h(x,z), \text{ par le même argument.}$$

BIBLIOGRAPHIE

- [1] R. Azencott, *Grandes déviations et applications*. Ecole d'été St Flour 1978.
Lecture Notes in Math. vol. 774 Springer-Verlag 1980.
- [2] Courant-Hilbert, *Methods of Mathematical Physics*. Interscience 1962.
- [3] B. Gaveau, *Principe de moindre action, propagation de la chaleur, estimées sous elliptiques sur certains groupes nilpotants*. Acta Mathematica. vol. 139. 1978.
- [4] B. Gaveau, *Systèmes hamiltoniens associés à des opérateurs hypoelliptiques*.
Bull. Sci. Math. 1978.
- [5] S. Molchanov, *Diffusions et géométrie riemannienne*. Uspechi Math. Nayk. 30
(1975) p. 3-59.
- [6] S. Varadhan, *Asymptotic probabilities and differential equations*. Comm. Pure
App. Math. 1966, p.261-286.

CHAPITRE 13

DIFFUSIONS INVARIANTES SUR LE GROUPE D'HEISENBERG UNE ÉTUDE DE CAS D'APRÈS GAVEAU [1]

par

Robert AZENCOTT

1. Groupe d'Heisenberg et diffusions invariantes hypoelliptiques.

1.1. Notations.

Soit G le groupe d'Heisenberg de dimension 3 que l'on peut réaliser en munissant $\mathbb{R}^3$ du produit

$$(x, y, z)(x', y', z') = [x + x', y + y', z + z' + 2(xy' - yx')]$$

L'algèbre de Lie de G admet pour base les champs invariants

$$X = \frac{\partial}{\partial x} + 2y \frac{\partial}{\partial z} \quad Y = \frac{\partial}{\partial y} - 2x \frac{\partial}{\partial z} \quad Z = \frac{\partial}{\partial z}$$

qui vérifient $[XY] = -4Z$, $[X, Z] = [Y, Z] = 0$

Les automorphismes σ de $\mathcal{G}$ sont donnés par $\sigma X = aX + bY + cZ$; $\sigma Y = a'X + b'Y + c'Z$; $\sigma Z = (ab' - ba')Z$ avec $a, b, c, a', b', c' \in \mathbb{R}$ et $ab' - ba' \neq 0$. Le lemme suivant est immédiat

1.2. LEMME. Soit L un opérateur différentiel d'ordre 2, invariant à gauche sur G . Supposons L de la forme $L = aX^2 + bY^2 + cZ^2 + dXY + eXZ + fYZ + gZ$.

Alors pour que L vérifie la condition d'hypoellipticité d'Hörmander il faut et il suffit que l'une des deux propriétés suivantes soit vraie :

(i) L est en fait elliptique

(ii) il existe un automorphisme ϕ de G tel que l'image de L par ϕ s'écrive

$$\phi^*(L) = \frac{1}{2}(X^2 + Y^2) + hZ \text{ avec } h \in \mathbb{R}, h \text{ arbitraire.}$$

La preuve est immédiate ; nous l'omettons donc. Puisque Z et $\frac{1}{2}(X^2 + Y^2)$ sont deux opérateurs qui commutent, les semi-groupes associés commutent aussi et l'étude du semi-groupe associé à $[\frac{1}{2}(X^2 + Y^2) + hZ]$ se ramène très facilement à celle du semi-groupe associé à $\frac{1}{2}(X^2 + Y^2)$.

Désormais nous considérerons l'opérateur $L = \frac{1}{2}(X^2 + Y^2)$ et la L -diffusion associée sur G , dont le semi-groupe admet une densité $p(t, g, h)$ invariante à gauche sur G . Étudions d'après Gaveau [1] le comportement de $p(t, g, h)$ lorsque $t \rightarrow 0$. Rappelons (Gaveau [1], cf. Chaleyat-Maurel ch.11) que si $g = (x, y, z) \in G$ et si $e = (0, 0, 0)$ est l'élément neutre de G , on a

$$(0) \quad p(t, e, g) = (2\pi t)^{-2} \int_{\mathbb{R}} dr \frac{2r}{\sinh 2r} \exp\left(ir \frac{z}{t} - \frac{x^2 + y^2}{2t} \frac{2r}{\tanh 2r}\right)$$

2. Un équivalent de $p(t, e, g)$ pour $t \rightarrow 0$, d'après Gaveau [1]:

2.1. Étudions d'abord le cas $x^2 + y^2 \neq 0$. Posons

$$(1) \quad A = \frac{z}{\frac{x^2 + y^2}{2}} \quad T = \frac{2t}{\frac{x^2 + y^2}{2}}$$

$$(2) \quad f(T, A) = \int_{\mathbb{R}} dr \frac{r}{\sinh r} \exp \frac{1}{T} (ir A - r \coth r)$$

pour constater que

$$(3) \quad p(t, e, g) = \frac{1}{2} (2\pi t)^{-2} f\left(\frac{2t}{\frac{x^2 + y^2}{2}}, \frac{z}{\frac{x^2 + y^2}{2}}\right)$$

Tout revient à étudier $f(T, A)$ quand $T \rightarrow 0$.

Pour formuler le résultat introduisons la fonction

$$(4) \quad \theta(r) = \frac{2r - \sin 2r}{2\sin^2 r}$$

qui est une bijection impaire de $]-\pi, +\pi[$ sur $]-\infty, +\infty[$.

2.2. PROPOSITION. (d'après Gaveau [1]). Soit $G = (x, y, z)$ un point du groupe d'Heisenberg G , tel que $x^2 + y^2 \neq 0$. Définissons successivement $\tau(g)$, $S(g)$, $H(g)$ par

$$(5) \quad \tau = \tau(g) = \theta^{-1}\left(\frac{z}{x^2 + y^2}\right), \quad \tau \in]-\pi, +\pi[, \text{ avec } \theta \text{ comme en (4)}$$

$$(6) \quad S(g) = (x^2 + y^2) \frac{\tau^2}{\sin^2 \tau}$$

$$(7) \quad H(g) = \frac{1}{2} \frac{\tau}{\sqrt{x^2 + y^2}} \left(\frac{\sin \tau}{\sin \tau - \tau \cos \tau} \right)^{1/2}$$

Alors lorsque $t \rightarrow 0$, la densité $p(t, e, g)$ de la L -diffusion sur G , admet l'équivalent suivant

$$(8) \quad p(t, e, g) \sim \frac{1}{(2\pi t)^{3/2}} H(g) \exp\left[-\frac{S(g)}{2t}\right]$$

Par contre lorsque $x^2 + y^2 = 0$, c'est-à-dire lorsque $g = (0, 0, z)$ on a lorsque $t \rightarrow 0$

$$(9) \quad p(t, e, g) \sim \frac{1}{8t^2} \exp\left(-\frac{\pi z}{2t}\right)$$

Preuve : Etudions $f(T, A)$ quand $T \rightarrow 0$. D'après (2) on a

$$(10) \quad f(T, A) = \int_{\mathbb{R}} \exp \frac{i}{T} \phi(r) \left[\frac{r}{\sinh r} dr \right] \text{ avec}$$

$$\phi(r) = Ar + i r \coth r$$

Une simple application du théorème de Cauchy montre que si $\tau \in]-\pi, +\pi[$, on a

$$f(T, A) = \int_{\mathbb{R}} \exp\left[\frac{i}{T} \phi(r+i\tau)\right] \frac{r+i\tau}{\operatorname{sh}(r+i\tau)} dr$$

et donc que

$$(11) \quad f(T, A) = \exp\left[\frac{i}{T} \phi(i\tau)\right] \int_{\mathbb{R}} B(r) dr \text{ avec}$$

$$(12) \quad B(r) = K(r) \exp \frac{i}{T} \psi(r) ; \quad K(r) = \frac{r+i\tau}{\operatorname{sh}(r+i\tau)} ; \quad \psi(r) = \phi(r+i\tau) - \phi(i\tau).$$

Choisissons τ tel que la phase ϕ soit stationnaire au point $i\tau$. Un calcul immédiat montre que l'équation $\phi'(i\tau) = 0$ équivaut d'après (4) à $\theta(\tau) = 0$, et donc à $\tau = \theta^{-1}(A)$, en imposant $\tau \in]-\pi, +\pi[$.

Posons $\psi''(r) = \alpha(r) + i\beta(r)$. Comme $\psi(0)$ et $\psi'(0)$ sont nuls on peut écrire

$$(13) \quad \psi(r) = \frac{r^2}{2} [\alpha(r_1) + i\beta(r_2)] \text{ avec } |r_1| \leq |r|, \quad |r_2| \leq |r|.$$

Après changement de variable $r = u\sqrt{T}$, on a donc

$$(14) \quad \int_{\mathbb{R}} B(r) dr = \sqrt{T} \int_{\mathbb{R}} K(u\sqrt{T}) \exp \left\{ i \frac{u^2}{2} [\alpha(r_1) + i\beta(r_2)] \right\} du,$$

où $|r_1| \leq u\sqrt{T}$, $|r_2| \leq u\sqrt{T}$. La quantité

$$(15) \quad \frac{i}{2} \psi''(0) = \frac{i}{2} [\alpha(0) + i\beta(0)] = \frac{i}{2} \phi''(i\tau) = \frac{\tau \cos \tau - \sin \tau}{\sin^3 \tau}$$

est réelle strictement négative pour $\tau \in]-\pi, +\pi[$; elle coïncide donc avec $-\frac{1}{2}\beta(0)$, et la continuité de β montre que, puisque K est bornée, la fonction à intégrer dans le second membre de (14) est, en module, majorée par $\text{cte} \times \exp(-\frac{u^2}{4} \beta(0))$ pourvu que T soit assez petit.

Le théorème de convergence dominée garantit donc, lorsque $T \rightarrow 0$ dans (14),

$$(16) \quad \lim_{T \rightarrow 0} \frac{1}{\sqrt{T}} \int_{\mathbb{R}} B(r) dr = K(0) \int_{\mathbb{R}} \exp \left[-\frac{u^2}{2} \beta(0) \right] du = \tau \left(\frac{\pi \sin \tau}{\sin \tau - \tau \cos \tau} \right)^{1/2}$$

Comme $\exp \frac{i}{T} \phi(i\tau) = \exp(-\frac{1}{T} \frac{\tau^2}{\sin^2 \tau})$, les relations (11) (16) fournissent pour $f(T,A)$ l'équivalent

$$(17) f(T,A) \sim \tau \sqrt{T} \left(\frac{\pi \sin \tau}{\sin \tau - \tau \cos \tau} \right)^{1/2} \exp(-\frac{1}{T} \frac{\tau^2}{\sin^2 \tau})$$

Grâce à (3), on en déduit l'équivalent annoncé pour $p(t,e,g)$.

Le cas particulier où $g = (0,0,z)$ se traite plus facilement car dans cette situation, la méthode des résidus permet en fait de calculer $p(t,e,g)$ pour obtenir

$$(18) p(t,e,g) = \frac{1}{8t^2} \exp(-\frac{\pi z}{2t}) [1 + \exp(-\frac{\pi z}{2t})]^2$$

(cf. Gaveau [1]). L'équivalent annoncé pour $t \rightarrow 0$ est alors évident.

3. Interprétation géométrique.

3.1. D'après ce que nous avons vu (Azencott ch.12), pour un opérateur hyprelliptique L de "rang constant" on peut s'attendre à ce que $\frac{1}{2} S(g)$, qui représente la limite de $t \log p(t,e,g)$ lorsque $t \rightarrow 0$ d'après 2.2.(8), coïncide en fait avec l'action minimale* pour aller de e à g en temps unité, ainsi qu'avec $E(e,g)$ où $E(e,g)$ est le λ minimal pour aller de e à g en temps unité, λ désignant la transformée de Cramer de Δ .

Le fait que $\frac{1}{2} S(g)$ coïncide avec l'action minimale a été prouvé (par un argument probabiliste) par Gaveau [1][2] dans le cas étudié ici mais la partie cruciale de sa preuve (cf. [1]) nous semble un peu rapide.

3.2. LEMME. *La situation est celle décrite en 2.2. Soit U l'ouvert du groupe G défini par $U = \{g = (u,y,z) \mid x^2 + y^2 \neq 0\}$. Alors les fonctions $H(g)$, $S(g)$ définies par (6), (7) sont continues sur U . De plus on a*

$$(19) \lim_{t \rightarrow 0} \frac{p(t,e,g)}{(2\pi t)^{-3/2} H(g) \exp(-\frac{S(g)}{2t})} = 1$$

* Signalons que la définition de l'action utilisée par Gaveau diffère de celle adoptée au ch.12 (et suivants) par un coefficient $\frac{1}{2}$.

uniformément sur tout compact de U .

Preuve : La continuité de H et S sur U est évidente. Le calcul de la limite (19) a déjà été fait pour prouver 2.2. Reste à vérifier l'uniformité de cette limite. Grâce à la relation (3) entre $p(t, e, g)$ et $f(T, A)$, ceci revient à vérifier que l'équivalent (17) de $f(T, A)$ est uniforme lorsque A reste dans un petit voisinage compact de A_0 fixé (arbitraire) dans $\mathbb{R}$.

Il s'agit donc seulement de vérifier que [cf. (14)] l'intégrale

$$I = \int_{\mathbb{R}} K(u\sqrt{T}) \exp \left\{ i \frac{u^2}{2} [\alpha(r_1) + i\beta(r_2)] \right\} du \text{ où } |r_1| \leq u\sqrt{T}, |r_2| \leq u\sqrt{T}, \text{ converge}$$

uniformément pour $T \rightarrow 0$, $A \in$ voisinage compact de A_0 , vers la limite calculée en (16).

On remarque pour celà que dans la situation ci-dessus, τ décrit un compact de $]-\pi, +\pi[$, par suite :

- $K(r)$ est uniformément bornée sur $\mathbb{R}$
- $\psi''(r)$ est continue en (r, τ) et donc clairement :
- $i \operatorname{Im} \psi''(r_2)$ est réel, négatif et majoré par $\gamma < 0$ pour r_2 assez petit et τ comme ci-dessus.

L'argument est alors analogue à celui de C. Bellaïche ch.9.

3.3. PROPOSITION. Soit G le groupe d'Heisenberg, U l'ouvert des $g = (x, y, z)$ tels que $x^2 + y^2 \neq 0$. Soit $S(g)$ la fonction introduite en (6). Soit λ la transformée de Cramer de $L = \frac{1}{2}(X^2 + Y^2)$ et soit $E(g, h)$ le minimum de λ sur les chemins joignant g à h en temps unité. Alors on a $\frac{1}{2} S(g) = E(e, g)$ pour tout $g \in U$; en particulier, pour $g \in U$, on a

$$(20) \lim_{t \rightarrow 0} t \log p(t, e, g) = - E(e, g) = - \frac{1}{2} S(g)$$

Preuve : Le lemme 3.2 entraîne immédiatement

$$\lim_{t \rightarrow 0} t \log p(t, e, g) = - \frac{1}{2} S(g)$$

uniformément sur les compacts de U , et S est continue sur U . Il suffit alors d'appliquer la prop. 4.6 de Azencott ch.12 pour conclure, car L est clairement hypoelliptique de rang constant.

3.4. PROPOSITION. (d'après Gaveau [1]). Les hypothèses et notations sont celles de 3.3. Pour tout $g \in U$, la fonction $\frac{1}{2} S(g)$ coïncide avec le minimum de l'action des bicaractéristiques (brisées) de L joignant x à y en temps unité.

Preuve : On écrit explicitement L dans les coordonnées (x,y,z) introduites au §1 sur G

$$(21) \quad L = \frac{1}{2} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 4y \frac{\partial^2}{\partial x \partial z} - 4x \frac{\partial^2}{\partial y \partial z} + 4(x^2 + y^2) \frac{\partial^2}{\partial z^2} \right]$$

L'équation différentielle des bicaractéristiques (γ, q) de L s'écrit (voir Azencott, ch.12), en notant $[\frac{1}{2} a_g]$ la matrice des coefficients de L

$$(22) \quad \begin{aligned} \gamma' &= a_\gamma q \\ q' &= -\frac{1}{2} q^* \left[\frac{\partial(a_\gamma)}{\partial \gamma} \right] q \end{aligned}$$

Soit en posant $\gamma = (x,y,z)$, $q = (u,v,w)$

$$(23) \quad \begin{aligned} x' &= u + 2yw \\ y' &= v - 2xw \\ z' &= 2yu - 2xv + 4(x^2 + y^2)w \\ u' &= 2vw - 4xw^2 \\ v' &= -2uw - 4yw^2 \\ w' &= 0 \end{aligned}$$

Ceci donne $w = \text{constante}$, puis x,y,u,v par résolution d'un système linéaire à coefficients constants, puis z par quadrature. D'où une première famille de bicaractéristiques issues de $e = (0,0,0)$, dépendant de trois paramètres réels arbitraires α, β, σ avec $\sigma \neq 0$, données par (cf. Gaveau [1])

$$(24) \quad \begin{aligned} x(s) &= \alpha(1 - \cos 2\sigma s) + \beta \sin 2\sigma s \\ y(s) &= \alpha \sin 2\sigma s - \beta(1 - \cos 2\sigma s) \\ z(s) &= 2(\alpha^2 + \beta^2)(2\sigma s - \sin 2\sigma s) \\ u(s) &= \alpha\sigma \sin 2\sigma s + \beta\sigma(1 + \cos 2\sigma s) \\ v(s) &= \alpha\sigma(1 + \cos 2\sigma s) - \beta\sigma \sin 2\sigma s \\ w(s) &= \frac{\sigma}{2} \end{aligned}$$

et une deuxième (et dernière famille) de bicaractéristiques issues de e données par (α, β, σ sont alors trois paramètres réels quelconques)

$$(25) \quad \begin{aligned} x(s) &= \alpha s & u(s) &= \alpha \\ y(s) &= \beta s & v(s) &= \beta \\ z(s) &= \sigma & w(s) &= 0 \end{aligned}$$

Pour une bicaractéristique (γ, q) l'action $I(\gamma, q)$ sur l'intervalle $[0, 1]$ est donnée par (cf. Azencott ch.12) $I(\gamma, q) = \frac{1}{2} \int_0^1 (ux' + vy' + wz') ds$. Le calcul donne (Gaveau [1])

$$(26) \quad \begin{aligned} I(\gamma, q) &= 2\sigma^2(\alpha^2 + \beta^2) && \text{si } (\gamma, q) \text{ vérifie (24)} \\ I(\gamma, q) &= \frac{1}{2}(\alpha^2 + \beta^2) && \text{si } (\gamma, q) \text{ vérifie (25)} \end{aligned}$$

Soit $g = (x, y, z) \in G$ tel que $z(x^2 + y^2) \neq 0$. Alors une bicaractéristique du type (24) vérifie $\gamma(1) = g$ si

$$(27) \quad \theta(\sigma) = \frac{z}{x^2 + y^2}, \text{ où } \theta(.) \text{ est définie par (4),}$$

α et β étant alors (uniquement) déterminés par les deux premières équations de (21), qui impliquent

$$(28) \quad \alpha^2 + \beta^2 = \frac{x^2 + y^2}{4 \sin^2 \sigma}$$

En particulier (26) (27) (28) donnent alors

$$(29) \quad I(\gamma, q) = \frac{1}{2}(x^2 + y^2) \frac{\sigma^2}{\sin^2 \sigma}$$

Comparant avec la définition (6) de $S(g)$ on voit que $\frac{1}{2} S(g) = I(\gamma, q)$ où (γ, q) est une bicaractéristique convenable de la famille (24) telle que $\gamma(0) = 0$ $\gamma(1) = g$.

Une vérification analogue montre que si $g = (x, y, 0)$ avec $x^2 + y^2 \neq 0$, on a encore $\frac{1}{2} S(g) = I(\gamma, q)$ avec $\gamma(1) = g$, mais cette fois avec (γ, q) vérifiant (25).

Soit U l'ouvert des $g = (x, y, z)$ tels que $x^2 + y^2 \neq 0$. Pour chaque $g \in U$ on a déjà vu (prop. 3.3) que $\frac{1}{2} S(g) = E(e, g)$. Mais d'autre part $E(e, y)$ est majoré par l'action minimale pour joindre e à g en temps unité par des bicaractéristiques brisées (cf. Azencott ch.12 §4.5). On en conclut que $\frac{1}{2} S(g)$ coïncide avec l'action minimale pour aller de e à g , lorsque $g \in U$.

3.5. Remarques. Dans la situation ci-dessus, on voit donc que pour $g \in U$, $E(e, g)$ coïncide avec l'action minimale pour aller de e à g en temps unité, situation non générique semble-t-il. Notons que dans les démonstrations de Gaveau [1] une ambiguïté sur le terme "action minimale" l'amène pour étendre le résultat précédent au cas où $g \notin U$, a

vérifier seulement qu'alors $\frac{1}{2} S(g)$ est le minimum de l'action parmi toutes les bicaractéristiques non brisées, qui vont de e à g en temps unité. Cette vérification se fait immédiatement, mais ne prouve rien en ce qui concerne l'action minimale de e à g , qui devrait envisager au moins toutes les bicaractéristiques brisées joignant e à g .

Pour "étendre les résultats 3.3 et 3.4 au cas où $g \notin U$, c'est-à-dire où $g = (0,0,z)$ il faudrait, en utilisant notre présentation, vérifier que la convergence $\lim_{t \rightarrow 0} 2t \log p(t,e,y) = -S(g)$ est uniforme sur un voisinage de g . Nous ne l'avons pas fait, le fait même étant douteux vu le comportement de $p(t,e,g)$ et la situation étudiée nous semblant trop spéciale pour mériter ce travail supplémentaire. La rédaction de Gaveau semble ignorer ces difficultés. Il est vrai qu'il se contente de prouver une inégalité du type

$$(*) \quad -E(e,y) \leq \lim_{t \rightarrow 0} t \log p(t,e,y).$$

Son argument (cf. [1]) en ce point n'est de toutes façons pas convainquant ; en effet sans préciser un peu l'uniformité de la convergence de $t \log p(t,e,y)$ quand $t \rightarrow 0$ il nous semble très peu plausible de lier naturellement $[t \log \int_V p(t,e,g) dg]$ et $[t \int_V \log p(t,e,g) dg]$ d'autant plus que l'inégalité de Jensen donne dans les "bonnes" situations $\int \log \leq \log \int$, ce qui va dans le mauvais sens si on cherche à prouver (*) à partir de la minoration naturelle de $\lim_{t \rightarrow 0} t \log \int_V p(t,e,g) dg$ liée à la transformée de Cramer de L .

Il serait agréable de clarifier la relation entre $E(e,g)$, $\frac{1}{2} S(g)$, et action minimale parmi les bicaractéristiques brisées joignant e à g , lorsque g appartient à l'ensemble "exceptionnel" des $(0,0,z)$, $z \in \mathbb{R}$. En effet en ces points la puissance de t dans l'équivalent de $p(t,e,g)$ est en $t^{-2} > t^{-3/2}$, ce qui correspond intuitivement comme le remarquent Gaveau [1] et Molchanov [3] à l'existence d'une famille "à une dimension" de bicaractéristiques joignant e à g . Les liens entre ce type de singularité et les égalités eventuelles de $E(e,g)$, $\frac{1}{2} S(g)$ action minimale de e à g , mériteraient d'être clarifiés.

BIBLIOGRAPHIE

- [1] B. Gaveau, *Principe de moindre action, propagation de la chaleur, estimées sous elliptiques sur certains groupes nilpotents*. Acta Mathematica. vol 139 1977.
- [2] B. Gaveau, *Systèmes hamiltoniens associés à des opérateurs hypoelliptiques*. Bull. Sci. Math. 1978.
- [3] S. Molchanov. *Diffusions et géométrie riemannienne*. Usp. Mat. Nank 30 (1975) p.3-59.

CHAPITRE 14

APPROXIMATION PAR DES CHAMPS INVARIANTS

par

Laure ELIE

1. Introduction.

Dans cet exposé, on se propose, toujours dans le but d'étudier le comportement asymptotique de la densité de transition d'une diffusion de générateur hypoelliptique, de rappeler quelques résultats d'approximation d'opérateurs hypoelliptiques par des opérateurs différentiels invariants à gauche sur un groupe nilpotent, l'article de base sur ce sujet étant celui de Rothschild-Stein [6].

1.1. Cadre.

Soient M une variété réelle de classe C^∞ de dimension m , $X_1, \dots, X_r$ des éléments de $\mathcal{L}(M)$, c'est-à-dire des champs de vecteurs réels de classe C^∞ . On considère sur M

l'opérateur différentiel $L = \sum_{i=1}^r X_i^2$.

Pour toute suite $I = (i_1, \dots, i_s) \in \{1, \dots, r\}^s$, on note $|I| = s$ la longueur de I et X_I le champ :

$$X_I = [X_{i_1} [X_{i_2} \dots [X_{i_{s-1}}, X_{i_s}] \dots]].$$

Un tel champ sera appelé commutateur de longueur s , et si $\mathcal{L}(X_1, \dots, X_r)$ désigne l'algèbre de Lie engendrée par les champs $X_1, \dots, X_r$, on appellera V_s^X (ou V_s s'il n'y a pas ambiguïté) le sous-espace vectoriel de $\mathcal{L}(X_1, \dots, X_r)$ engendré par les commutateurs de longueur $\leq s$, c'est-à-dire par les X_I , $|I| \leq s$. En particulier V_1^X est l'espace vectoriel engendré par les X_i , $1 \leq i \leq r$.

Nous supposons que la famille X_i , $1 \leq i \leq r$ vérifie la condition d'Hörmander, c'est-à-dire, que pour tout x de M , l'image $\mathcal{L}(X_1, \dots, X_r)(x)$ décrit l'espace tangent $T_x(M)$ au point x . Alors l'opérateur L (et aussi l'opérateur $L - \frac{\partial}{\partial t}$) est hypoelliptique.

Nous pouvons remarquer que si en un point x de M , $\mathcal{L}(X_1, \dots, X_r)(x)$ est égal à $T_x(M)$, il existe un rang k (puisque $T_x(M)$ est de dimension finie) tel que $V_k(x) = T_x(M)$. Les champs de vecteurs étant C^∞ , cette propriété sera encore vraie sur un voisinage de x . Comme les problèmes que nous voulons étudier sont locaux, on fera (quitte à remplacer M par un voisinage de x) l'hypothèse suivante :

(H) Il existe un rang k tel que pour tout x de M , on ait $V_k(x) = T_x(M)$.

1.2. Approximation.

Des exemples naturels d'opérateurs hypoelliptiques peuvent s'obtenir de la manière suivante : Considérons un groupe de Lie G dont l'algèbre de Lie $\mathfrak{g}$ admet r générateurs $Y_1, \dots, Y_r$, alors l'opérateur $\sum_{i=1}^r Y_i^2$ est hypoelliptique et de plus l'hypothèse (H) est vérifiée. En particulier si G est un groupe nilpotent d'ordre k , pour tout g de G , $V_k^Y(g) = T_g(G)$.

On cherche à approcher $L = \sum_{i=1}^r X_i^2$ par un opérateur différentiel invariant à gauche ou plus précisément à approcher les champs de vecteurs X_i , $i \in \{1, \dots, r\}$ par des champs de vecteurs invariants à gauche Y_i , $i \in \{1, \dots, r\}$. Supposons que les champs Y_i engendrent une certaine algèbre de Lie $\mathfrak{q}$ et considérons un groupe de Lie G connexe admettant $\mathfrak{q}$ comme algèbre de Lie. L'idée d'approximation va être la suivante :

En tout point g de la variété G , l'espace tangent $T_g(G)$ est isomorphe à $\mathfrak{q}$ et chercher à "approximer" $T_g(G)$ revient en fait à transporter sur $T_x(M)$ la structure de l'algèbre de Lie $\mathfrak{q}$. Il sera donc naturel d'imposer à $\mathfrak{q}$ d'avoir une structure d'algèbre de Lie compatible avec celle de $\mathcal{L}(M)$.

On cherchera donc à construire une application linéaire λ de $\mathfrak{q}$ dans $\mathcal{L}(M)$, telle que $\lambda(Y_i) = X_i$ pour $i \in \{1, \dots, r\}$, compatible "autant que possible" avec les structures d'algèbre de $\mathfrak{q}$ et $\mathcal{L}(M)$.

Alors, pour tout x_0 de M , l'application λ_{x_0} définie sur $\mathfrak{g}$ par $\lambda_{x_0}(Y) = \lambda(Y)(x_0)$ est une application linéaire de $\mathfrak{g}$ dans $T_{x_0}(M)$, et on imposera naturellement, que λ permette au moins sur un voisinage de x_0 , de parcourir tout l'espace tangent, c'est-à-dire que λ_{x_0} soit surjective.

Quelles hypothèses imposer à $\mathfrak{g}$? On cherche une algèbre de dimension "petite" permettant cependant de décrire tout l'espace tangent. Or on sait que d'après l'hypothèse (H), on a, pour, tout x de M , $V_k^X(x) = T_x(M)$.

Supposons que λ soit "presque un homomorphisme" au sens suivant (cf. exemple 1.c ci-dessus) : λ est un "homomorphisme partiel" sur V_k^Y , c'est-à-dire que si $Z, T, [Z, T] \in V_k^Y$, alors $\lambda([Z, T]) = [\lambda(Z), \lambda(T)]$. De ce fait $\lambda(V_k^Y) = V_k^X$, et donc pour tout x de M , $\lambda_x(V_k^Y) = T_x(M)$. En conséquence la connaissance de l'image par λ des commutateurs de longueur $\leq k$ des Y_i , $i \in \{1, \dots, r\}$ nous permet alors de déterminer le fibré tangent TM . Il semble donc naturel et suffisant de rechercher une algèbre de Lie $\mathfrak{g}$ nilpotente d'ordre k et une application λ de $\mathfrak{g}$ dans $\mathcal{L}(M)$ qui soit un homomorphisme au sens précédent. Nous allons donner quelques exemples.

1.3. Exemples.

① On suppose qu'en tout point x de M , les champs $X_1, \dots, X_r$ sont linéairement indépendants et engendrent l'espace tangent $T_x(M)$. Alors si $\mathfrak{g}$ désigne l'algèbre de Lie de $\mathbb{R}^r$ de base $Y_1, \dots, Y_r$, l'application définie par $\lambda(Y_i) = X_i$ pour $i \in \{1, \dots, r\}$ se prolonge en un homomorphisme de Lie de $\mathfrak{g}$ dans $\mathcal{L}(X_1, \dots, X_r)$. De plus, pour tout x de M , l'application $\lambda_x : Y \rightarrow \lambda(Y)(x)$ est un isomorphisme linéaire de $\mathfrak{g}$ sur $T_x(M)$.

En fait cet exemple traitera exactement le cas où l'opérateur L est elliptique ; et donc, pour un opérateur elliptique, l'algèbre nilpotente "sous-jacente" est abélienne.

② On suppose que $M = \mathbb{R}^2 = \{(x, y)\}$ et que $X_1 = \frac{\partial}{\partial x}$, $X_2 = x \frac{\partial}{\partial y}$. Alors X_1 et X_2 n'engendrent pas $T_\xi(M)$ au point $\xi = (0, y)$, mais comme $[X_1, X_2] = \frac{\partial}{\partial y}$, les commutateurs de longueur ≤ 2 engendrent l'espace tangent en tout point. L'algèbre de Lie $\mathcal{L}(X_1, X_2)$ est nilpotente d'ordre 2 et admet pour base X_1, X_2 et $[X_1, X_2]$. Considérons l'algèbre de Lie d'Heisenberg $\mathcal{H}_3$ de base Y_1, Y_2 et $Y_3 = [Y_1, Y_2]$. Il existe alors un homomorphisme d'algèbre de Lie λ de $\mathcal{H}_3$ dans $\mathcal{L}(X_1, X_2)$ tel que $\lambda(Y_1) = X_1$, $\lambda(Y_2) = X_2$, $\lambda(Y_3) = [X_1, X_2]$. De plus pour tout ξ de M , $\lambda(\mathcal{H}_3)(\xi) = T_\xi(M)$, mais l'application $\lambda_\xi : Y \rightarrow \lambda(Y)(\xi)$ n'est pas un isomorphisme linéaire puisque $T_\xi(M)$ est de dimension 2.

On a remarqué qu'au point $\xi = (0, y)$, les champs X_1 et X_2 n'étaient pas linéairement indépendants ; mais si on avait considéré sur $\tilde{M} = \mathbb{R}^3 = (x, y, u)$ les champs $\tilde{X}_1 = \frac{\partial}{\partial x}$ et $\tilde{X}_2 = x \frac{\partial}{\partial y} + \frac{\partial}{\partial u}$, alors les champs $\tilde{X}_1, \tilde{X}_2$ et $[\tilde{X}_1, \tilde{X}_2] = \frac{\partial}{\partial y}$ auraient été linéairement indépendants en tout point et si on avait défini $\tilde{\lambda}$ comme ci-dessus, l'application $\tilde{\lambda}_{\tilde{\xi}}$ aurait été un isomorphisme linéaire de $\mathcal{H}_3$ sur $T_{\tilde{\xi}}(\tilde{M})$.

(C) On va légèrement modifier l'exemple précédent en supposant que M est un voisinage de 0 dans $\mathbb{R}^2$ et que $X_1 = \frac{\partial}{\partial x}$, $X_2 = \phi(x) \frac{\partial}{\partial y}$ avec ϕ fonction C^∞ vérifiant $\phi(0) = 0$ et que pour tout $n \in \mathbb{N}^*$ et $(x, y) \in M$, $\phi^{(n)}(x) \neq 0$. Alors $[X_1, X_2] = \phi'(x) \frac{\partial}{\partial y}$ et par suite les commutateurs de longueur ≤ 2 engendrent l'espace tangent en tout point. Mais dans ce cas l'algèbre de Lie $\mathcal{L}(X_1, X_2)$ n'est pas nilpotente et elle est même de dimension infinie. Mais on peut construire comme ci-dessus une application linéaire λ de $\mathcal{H}_3$ dans $\mathcal{L}(X_1, X_2)$ telle que $\lambda(Y_1) = X_1$, $\lambda(Y_2) = X_2$, $\lambda(Y_3) = [X_1, X_2]$ et pour tout ξ de M , $\lambda(\mathcal{H}_3)(\xi) = T_\xi(M)$. Mais λ ne sera pas un homomorphisme d'algèbre de Lie car par exemple $\lambda([Y_1, [Y_1, Y_2]]) = 0$ et $[X_1, [X_1, X_2]] = \phi''(x) \frac{\partial}{\partial x} \neq 0$. Il vérifiera seulement $\lambda([Y_1, Y_2]) = [X_1, X_2]$ et on dira que c'est un homomorphisme "partiel".

1.4. Admettons donc que l'on puisse construire une application linéaire de $\mathcal{g}$ dans $\mathcal{L}(M)$ vérifiant, pour tout $i \in \{1, \dots, r\}$, $\lambda(Y_i) = X_i$ et telle que $\lambda(\mathcal{g})(x_0) = T_{x_0}(M)$. Alors on peut déterminer une submersion ϕ_{x_0} d'un voisinage de 0 dans G sur un voisinage de x_0 dans M de la manière suivante :

Considérons si X est un champ de vecteurs, la coulée locale $t \rightarrow e^{tX}$ engendrée par X . Par définition, si $f \in C^\infty(M)$ et $x \in M$,

$$Xf [e^{tX}x] = \frac{d}{dt} f[e^{tX}x].$$

L'application $Y \rightarrow e^{\lambda(Y)}x_0$ existe alors pour Y proche de 0 dans $\mathcal{g}$ (d'après le théorème d'existence de solution d'équation différentielles) et comme de plus la différentielle de cette application est en 0 l'application $Y \rightarrow \lambda(Y)(x_0)$ qui est surjective puisque $\lambda(\mathcal{g})(x_0) = T_{x_0}(M)$, cette application est une submersion d'un ouvert de $\mathcal{g}$ contenant 0 sur un ouvert de M contenant x_0 .

De plus l'application exponentielle définie sur $\mathcal{g}$ à valeurs dans G est un difféomorphisme au voisinage de 0 . Par suite il existe un ouvert U dans G contenant l'élément neutre 0 et un ouvert V dans M contenant x_0 tels que l'application ϕ_{x_0}

définie, pour $g = e^Y \in U$, par $\phi[e^Y] = e^{\lambda(Y)} x_0$, soit une submersion de U sur V .

Alors si $\phi^* X_i$ désigne le champ de vecteurs défini sur $U \subset G$ par $\phi^* X_i(f \circ \phi)(g) = X_i(f)[\phi(g)]$ si $g \in U$ et $f \in C^\infty(M)$, on a $\phi^* X_i(e) = Y_i(e)$ par construction. En effet si $f \in C^\infty(M)$,

$$\begin{aligned} \phi^* X_i(f \circ \phi)(e) &= X_i f[\phi(e)] = X_i f[x_0] = \left[\frac{d}{dt} f[e^{tX_i} x_0] \right]_{t=0} \\ &= \left[\frac{d}{dt} f \circ \phi(e^{tY_i}) \right]_{t=0} = Y_i(f \circ \phi)(e) \end{aligned}$$

On cherche alors à mesurer l'écart entre $\phi^* X_i$ et Y_i sur U .

On peut remarquer qu'il serait agréable que ϕ soit un difféomorphisme pour que $\phi^* X_i$ soit une exacte copie de X_i ; cela reviendrait à imposer à l'application $Y \rightarrow \lambda(Y)(x_0)$ d'être un isomorphisme linéaire de $\mathfrak{g}$ sur $T_{x_0}(M)$.

1.5. Diverses méthodes.

On montre que quel soit le système $(X_1, \dots, X_r)$ choisi vérifiant (H) pour un certain k , on peut toujours construire un homomorphisme "partiel" λ de l'algèbre de Lie nilpotente de rang k à r générateurs $Y_1, \dots, Y_r$ notée $\mathcal{N}_{Y,k}$ dans $\mathcal{L}(M)$ tel que pour tout x de M , $\lambda(\mathcal{N}_{Y,k})(x) = T_x(M)$ (cf. Lemme 2.3). Mais les applications λ_x de $\mathcal{N}_{Y,k}$ sur $T_x(M)$ définies par $Y \rightarrow \lambda(Y)(x)$ ne sont en général pas des isomorphismes. Elles ne le sont que lorsque $\dim \mathcal{N}_{Y,k} = \dim M$. Et alors comme on l'a remarqué au paragraphe 1.4, $\phi^* X_i$ n'est pas une exacte reproduction de X_i . De plus l'exemple 1.b nous prouve qu'il n'est pas toujours possible de trouver, pour un $x_0 \in M$ donné, une algèbre nilpotente $\mathfrak{g}$ telle que λ_{x_0} soit un isomorphisme linéaire. Il était suggéré dans ce cas de grossir l'espace M en ajoutant une nouvelle variable, de transformer les variables initiales de façon à pouvoir construire un homomorphisme λ tel que $\tilde{\lambda}_{\tilde{x}_0}$ soit un isomorphisme.

Cette méthode est celle utilisée par L.P. Rothschild et E.M. Stein dans [6]. On aggrandit l'espace M en ajoutant de nouvelles variables et on prolonge à ce nouvel espace $\tilde{M}$ les vecteurs $X_1, \dots, X_r$ en vecteurs $\tilde{X}_1, \dots, \tilde{X}_r$ de manière à ce que le système $(\tilde{X}_1, \dots, \tilde{X}_r)$ soit libre, ce qui revient à dire que $\dim \mathcal{N}_{Y,k} = \dim \tilde{M}$; Alors l'homomorphisme partiel $\tilde{\lambda}$ de $\mathcal{N}_{Y,k}$ dans $\mathcal{L}(\tilde{M})$ est tel que, pour tout x de $\tilde{M}$, $\tilde{\lambda}_x$ est un isomorphisme. Par suite si on mesure l'écart entre $\phi^*(\tilde{X}_i)$ et Y_i , on saura exactement comment Y_i approche $\tilde{X}_i$. Il faut ajouter que les $\tilde{X}_i$ sont construits de façon à ce que, si f est un élément de $C^\infty(M)$ constant sur les dernières coordonnées, $\tilde{X}_i f = X_i f$, f pouvant aussi être considéré comme un élément de $C^\infty(M)$. Dans certains problèmes, cette relation

entre $\tilde{X}_i$ et X_i suffit. Mais il semble que pour notre problème d'étude en temps petit de la densité de transition de la diffusion de générateur $\sum_{i=1}^r X_i^2$, cette relation soit insuffisante. Plus précisément considérons l'exemple 1-b, alors

$$L = \sum_{i=1}^2 X_i^2 = \frac{\partial^2}{\partial x^2} + x^2 \frac{\partial^2}{\partial y^2} \text{ et } \tilde{L} = \frac{\partial^2}{\partial x^2} + x^2 \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial u^2} + 2x \frac{\partial^2}{\partial y \partial u}.$$

Notons p et $\tilde{p}$ les densités de transition respectives de la L -diffusion sur M et de la $\tilde{L}$ -diffusion sur $\tilde{M}$. Rappelons qu'en tant que fonction de (t, z) , $p(t, z, z')$ vérifie

$$(L - \frac{\partial}{\partial t}) p(t, z, z') = 0$$

et $\lim_{t \rightarrow 0} \int p(t, z, z') f(z') = f(z)$ pour $f \in C_K(M)$.

Il est alors évident que la solution $\tilde{p}[t, (x, y, u), (x', y', u')]$ n'est pas constante en u et il semble très difficile de trouver une relation entre p et $\tilde{p}$. Cette méthode consistant à rendre le système libre ne paraît donc pas appropriée à notre problème.

Une autre méthode consiste à diminuer la taille de $\mathcal{N}_{Y,k}$ de manière à ce que λ_x soit un isomorphisme. On a remarqué (exemple 1.b) qu'on ne pouvait pas toujours trouver de sous-algèbre de $\mathcal{N}_{Y,k}$ comme solution. Dans [3], R.W. Goodman considère M comme un espace homogène du groupe nilpotent connexe $N_{Y,k}$ d'algèbre de Lie $\mathcal{N}_{Y,k}$, mais n'explique pas vraiment cet espace homogène. Par contre B. Hellfer et J.F. Nourrigat dans [4] associent à chaque point x de M , une sous-algèbre de $\mathcal{N}_{Y,k}$ de $\mathcal{N}_{Y,k}$ telle que $\dim \mathcal{X}(x) = \dim \mathcal{N}_{Y,k} - \dim M$ et construisent un supplémentaire convenable $S(x)$ de $\mathcal{X}(x)$ dans $\mathcal{N}_{Y,k}$ sur lequel λ_x est un isomorphisme. Ils approximent alors au voisinage de x les champs X_i par l'image des générateurs Y_i de $\mathcal{N}_{Y,k}$ par une représentation de $\mathfrak{g}$ dans $L^2[S(x)]$.

Des hypothèses plus fortes sur la famille $X_1, \dots, X_r$ (c'est-à-dire pour $s \in \{1, \dots, k\}$, la dimension de $V_s(x)$ est indépendante de x dans M et $\dim V_k(x) = \dim M$) assurent que la sous-algèbre $\mathcal{X}(x)$ est un idéal de $\mathcal{N}_{Y,k}$ et permettent à G. Métivier dans [5] d'approximer au voisinage de x les champs X_i par des champs invariants à gauche sur le groupe $N_{Y,k}/\exp \mathcal{X}(x)$. Il y est aussi obtenu des résultats fins de dépendance en x du difféomorphisme ϕ_x .

Pour notre problème, il semble donc que ce deuxième type d'approximation soit le meilleur. Mais en fait pour un opérateur hypoelliptique $L = \sum X_i^2$, les seuls résultats sur le comportement en temps petit de la densité p_t associée à L ont été obtenus par B. Gaveau pour le noyau de la chaleur sur le groupe libre $N_{n,2}$ à n générateurs de rang 2.

Il est donné dans [2] une majoration en temps petit de $-\log p_t$ sur $N_{n,2}$ et dans [1] une estimée de p_t sur le groupe d'Heisenberg.

Et si on voulait utiliser les résultats d'approximation de [5] (ou de [4]), il faudrait avoir des informations sur le p_t associé à un générateur invariant hypoelliptique sur un groupe nilpotent quelconque (ou sur un espace homogène de groupe nilpotent)!!!

Nous allons donc dans cet exposé nous restreindre pour obtenir le théorème d'approximation au cas où le système $X_1, \dots, X_r$ est libre, c'est-à-dire au cas où "l'homomorphisme" λ de $\mathcal{D}_{Y,r}$ dans $\mathcal{L}(M)$ est tel que pour tout x , λ_x est isomorphisme linéaire de $\mathcal{D}_{Y,r}$ sur $T_x(M)$. Dans ce cadre, les démonstrations sont particulièrement simples et sont données au paragraphe 2.

Nous expliquerons enfin au paragraphe 3 ce que nous espérons déduire d'une telle approximation pour l'étude en temps petit des densités de transition.

2. Théorème d'approximation pour un système $X_1, \dots, X_r$ libre.

2.1. DÉFINITIONS. Une algèbre de Lie $\mathcal{g}$ est dite graduée d'ordre k si $\mathcal{g} = \bigoplus_{j=1}^k \mathcal{g}_j$ avec $[\mathcal{g}_i, \mathcal{g}_j] \subset \mathcal{g}_{i+j}$ si $i+j \leq k$ et $[\mathcal{g}_i, \mathcal{g}_j] = 0$ si $i+j > k$. Une telle algèbre de Lie est nilpotente de rang k .

. Soit $X_1, \dots, X_r$, r champs de vecteurs C^∞ sur M . Une application linéaire λ de $\mathcal{g}$ dans $\mathcal{L}(X_1, \dots, X_r)$ sera dit un homomorphisme partiel d'ordre s si pour tout $Y \in \mathcal{g}_i$ et tout $T \in \mathcal{g}_j$ avec $i+j \leq s$, alors $\lambda([Y, T]) = [\lambda(Y), \lambda(T)]$. Si de plus, en tout point x de M , l'application λ_x qui à $Y \in \mathcal{g}$ associe $\lambda(Y)(x)$ est un isomorphisme linéaire de $\mathcal{g}$ sur $T_x(M)$, on dira que λ est un isomorphisme partiel d'ordre s .

Exemple : Le groupe d'Heisenberg est une algèbre de Lie graduée : $\mathcal{H}_3 = \mathcal{g}_1 \oplus \mathcal{g}_2$ où $\mathcal{g}_1$ (resp. $\mathcal{g}_2$) est l'espace vectoriel engendré par Y_1 et Y_2 (resp. Y_3). Dans l'exemple $\textcircled{C}$, l'homomorphisme λ construit est un homomorphisme partiel d'ordre 2.

2.2. Groupe nilpotent libre.

Soient $y_1, \dots, y_r$ r éléments et soit E l'algèbre libre à r générateurs $y_1, \dots, y_r$. L'algèbre E peut être munie d'une structure d'algèbre de Lie par le crochet $[u, v] = uv - vu$ et l'algèbre de Lie libre F à r générateurs $y_1, \dots, y_r$ est la sous-algèbre de Lie de E engendrée par $y_1, \dots, y_r$. Alors $F = \bigoplus_n F_n$ où F_n désigne l'espace vectoriel

engendré par les commutateurs de longueur n . Intuitivement F est l'algèbre de Lie de générateurs $y_1, \dots, y_r$ ayant le moins de relations possibles entre les commutateurs. (Existent toujours l'antisymétrie, l'identité de Jacobi...).

On considère $H_k = \bigoplus_{n>k} F_n$; H_k est un idéal de F et $F/H_k = \mathcal{N}_{Y,k}$ est une algèbre de Lie nilpotente de rang k . Si $Y_1, \dots, Y_r$ désignent les images de $y_1, \dots, y_r$, on dit que $\mathcal{N}_{Y,k}$ est l'algèbre de Lie nilpotente libre de rang k à r générateurs $Y_1, \dots, Y_r$. Par exemple $\mathcal{N}_3$ est une algèbre de Lie nilpotente de rang 2 à 2 générateurs.

Comme $F = \bigoplus_{n=1}^k F_n \oplus H_k$, $\mathcal{N}_{Y,k}$ est une algèbre de Lie graduée : $\mathcal{N}_{Y,k}^\cap = \bigoplus_{n=1}^k g_n$ où g_n désigne l'image de F_n . De plus il existe un homomorphisme partiel ψ_k d'ordre k de $\mathcal{N}_{Y,k}$ dans F qui à Y_i associe y_i pour $i \in \{1, \dots, r\}$.

On appellera groupe libre $N_{Y,k}$ le groupe nilpotent connexe simplement connexe d'algèbre de Lie $\mathcal{N}_{Y,k}$. L'application exponentielle est alors un difféomorphisme de $\mathcal{N}_{Y,k}$ sur $N_{Y,k}$.

2.3. LEMME. Soient $X_1, \dots, X_r \in \mathcal{L}(M)$. Alors pour tout $s \geq 1$, il existe un et un seul homomorphisme partiel λ_s de l'algèbre de Lie $\mathcal{N}_{Y,s}$ à r générateurs $Y_1, \dots, Y_r$ dans $\mathcal{L}(X_1, \dots, X_r)$ tel que $\lambda_0(Y_i) = X_i$ pour $i \in \{1, \dots, r\}$.

Preuve : Elle est immédiate à l'aide de la propriété universelle de l'algèbre de Lie libre F . Il existe un homomorphisme d'algèbre de Lie χ de F dans $\mathcal{L}(X_1, \dots, X_r)$ vérifiant $\chi(y_i) = X_i$ il suffit de poser $\lambda_s = \chi \circ \psi_s$.

On peut remarquer que $\lambda_s[\mathcal{N}_{Y,s}^\cap] = V_s$ où V_s désigne l'espace vectoriel engendré par les commutateurs de longueur $\leq s$ des X_i , $i \in \{1, \dots, r\}$. En conséquence en tout point x de M , $\dim V_s(x) \leq \dim \mathcal{N}_{F,s}^\cap$. On pose alors la définition suivante :

2.4. DÉFINITION. On notera r_s la dimension de l'algèbre de Lie libre $\mathcal{N}_{Y,s}$ de rang s à r générateurs. On dira que le système de champs de vecteurs $\{X_1, \dots, X_r\}$ est libre jusqu'à l'ordre s en x si $\dim V_s(x) = r_s$.

En fait puisque les champs $X_1, \dots, X_r$ sont C^∞ , si le système est libre en x jusqu'à l'ordre s , il est libre sur un voisinage de x jusqu'à l'ordre s . Comme les problèmes que l'on se pose sont locaux, on peut supposer (quitte à remplacer M par un voisinage de x) que le système est libre partout jusqu'à l'ordre s .

2.5. COROLLAIRE. Si le système $(X_1, \dots, X_r)$ vérifie l'hypothèse (H) pour un certain entier k (c'est-à-dire $V_k(x) = T_x(M)$, $\forall x \in M$) et est libre jusqu'à l'ordre k , en tout point de M , alors l'homomorphisme λ_k construit en 2.3 est un isomorphisme partiel d'ordre k .

Dans toute la suite, nous allons supposer que les hypothèses du corollaire sont vérifiées.

2.6. Notation. Si X est un champ de vecteurs. On note $t \rightarrow e^{tX}$ la coulée locale engendrée par X . (cf. 1.4).

2.7. PROPOSITION. Supposons que le système $X_1, \dots, X_r$ vérifie (H) pour un entier k et est libre jusqu'à l'ordre k . Considérons l'isomorphisme partiel λ_k de $\mathcal{M}_{Y,k}$ dans $\mathcal{L}(X_1, \dots, X_r)$ construit en 2.3. Alors il existe un voisinage ouvert V de x_0 dans M et un voisinage ouvert U de 0 dans $N_{Y,k}$ tels que

. l'application $(x, u) \rightarrow \phi_x(u)$ définie pour $x \in V$ et $u = e^Y \in U$ ($Y \in \mathcal{M}_{Y,k}$) par $\phi_x(u) = \phi_x(e^Y) = e^{\lambda_k(Y)} x$ soit C^∞ de $V \times U$ dans M .

. l'application $\phi_x : u \rightarrow \phi_x(u)$ pour $x \in V$ soit un difféomorphisme de U sur un ouvert de M contenant V .

Preuve : Suivant [3], considérons l'application $(Y, x) \rightarrow e^{\lambda_k(Y)} x$. Celle-ci existe et est C^∞ pour $(Y, x) \in \mathcal{U}_1 \times V_1$, $\mathcal{U}_1$ (resp. V_1) étant un voisinage ouvert de 0 dans $\mathcal{M}_{Y,k}$ (resp. de x_0 dans M), d'après les résultats d'existence de solutions d'équations différentielles. Comme de plus, pour x fixé $\in V_1$, cette application admet pour différentielle en 0 , l'isomorphisme $Y \rightarrow \lambda_k(Y)(x)$, il existe un voisinage de 0 inclus dans $\mathcal{U}_1$ (dépendant de x) sur lequel l'application $Y \rightarrow e^{\lambda_k(Y)} x$ est un difféomorphisme. En conséquence, par un argument d'uniformité (cf. [3]), il existe des voisinages ouverts $\mathcal{U}$ et V de 0 dans $\mathcal{M}_{Y,k}$ et de x_0 dans M tels que $\mathcal{U} \subset \mathcal{U}_1$ et $V \subset V_1$ et tels que, pour $x \in V$, l'application $Y \rightarrow e^{\lambda_k(Y)} x$ soit un difféomorphisme de $\mathcal{U}$ sur un ouvert de M contenant V .

Comme l'application exponentielle de $\mathcal{M}_{Y,k}$ dans $N_{Y,k}$ est un difféomorphisme, il suffit de poser $U = e^{\mathcal{U}}$ et si $x \in V$, et $u = e^Y \in U$ ($Y \in \mathcal{U}$), $\phi_x(u) = \phi_x[e^Y] = e^{\lambda_k(Y)} x$ pour obtenir la proposition.

2.8. COROLLAIRE. Sous les notations de la proposition 2.7, il existe une application θ de classe C^∞ de $V \times V$ dans U telle que pour $(x, y) \in V \times V$

$$\theta(x, y) = \theta_x(y) = e^Y \text{ si } y = e^{\lambda_k(Y)} x.$$

De plus, pour tout x de V , l'application θ_x est un difféomorphisme sur V .

Preuve : Elle est immédiate : il suffit de poser $\theta_x(y) = \phi_x^{-1}(y)$ pour $(x, y) \in V \times V$. On peut remarquer que si $M = \mathcal{W}_{Y, r}^p$ et $X_i = Y_i$, $\theta(x, y) = yx^{-1}$; et donc θ généralise ce type de fonctions.

2.9. Coordonnées canoniques.

Choisissons une base (e_n) , $1 \leq n \leq r_k$ de $\mathcal{W}_{Y, k}^1$ telle que les e_j , $r_{s-1} \leq j \leq r_s$, forment une base de l'espace vectoriel des commutateurs de longueur s ($s \in \{1, \dots, k\}$). On suppose de plus que $e_i = Y_i$ pour $i \in \{1, \dots, r\}$.

On dira que $u \in \mathcal{W}_{Y, k}^p$ admet pour coordonnées canoniques $(u_n)_{1 \leq n \leq r_k}$ si $u = e^{\sum u_n e_n}$.

On appellera p la carte de $N_{Y, k}$ définie par $p(u) = (u_n)_{1 \leq n \leq r_k}$.

L'application θ_x permet de transporter ce système de coordonnées sur V et pour tout x de V , l'application $k_x = p \circ \theta_x : V \rightarrow \mathbb{R}^{r_k}$ définit une carte locale. Plus précisément si $y = e^{\lambda(\sum u_n e_n)} \in V$, $k_x(y) = (u_n)_{1 \leq n \leq r_k}$.

On peut donc représenter ainsi les éléments de $V \subset M$ et $U \subset N_{Y, k}$ par le même système de coordonnées.

2.10. Notion de degré d'homogénéité.

Si $\mathcal{G} = \bigoplus_{j=1}^k \mathcal{G}_j$ est une algèbre de Lie graduée, il existe une famille d'automorphismes de $\mathcal{G}$, $(\delta_t)_{t \in \mathbb{R}^+}$, appelés dilatations, définis pour $Y \in \mathcal{G}_j$ par $\delta_t(Y) = t^j Y$.

En particulier, si $\mathcal{G} = \mathcal{W}_{Y, k}^p$ et si on définit la fonction poids $[]$ de $\{1, \dots, r_k\}$ dans $\{1, \dots, k\}$ par $[j] = s$ si $r_{s-1} \leq j \leq r_s$, on a $\delta_t(e_j) = t^{[j]} e_j$.

Par l'application exponentielle, ces automorphismes δ_t de $\mathcal{G}$ sont transformés en automorphismes, notés encore δ_t , de G , groupe nilpotent connexe et simplement connexe d'algèbre de Lie $\mathcal{G}$. Dans le système de coordonnées canoniques de $N_{Y, k}$ défini en 2.9,

$$\delta_t((u_j)) = (t^{[j]} u_j).$$

Une fonction f définie sur G est dite homogène de degré ρ si $f \circ \delta_t = t^\rho \cdot f$. Par exemple p_j l'application $j^{\text{ème}}$ -coordonnée de $N_{Y,k}$ dans $\mathbb{R}$ est de degré $[j]$.

Un opérateur différentiel sur G sera dit homogène de degré ρ si $D(f \circ \delta_t) = t^\rho (Df) \circ \delta_t$; les opérateurs e_j et $\frac{\partial}{\partial u_j}$ sont homogènes de degré $[j]$ sur $N_{Y,k}$.

Pour un multi-indice $\alpha = (j_1, \dots, j_m)$, on pose $|\alpha| = \sum_{i=1}^m [j_i]$, $u_\alpha = u_{j_1} \dots u_{j_m}$ et $\partial_u^\alpha = \frac{\partial^m}{\partial j_1 \dots \partial j_m}$. Alors $u_\alpha \partial_u^\beta$ est un opérateur homogène de degré $|\beta| - |\alpha|$ sur $N_{Y,k}$.

2.11. Degré local.

On munit $N_{Y,k}$ de la norme homogène, $|u| = \sum_{j=1}^{r_k} |u_j|^{1/[j]}$. Suivant [3], on pose, si U est un voisinage de 0 dans $N_{Y,k}$, pour $m \in \mathbb{Z}$,

$$C_m(U) = \{f \in C^\infty(U), |f(u)| = O(|u|^m) \text{ lorsque } |u| \rightarrow 0\}.$$

. Un opérateur T de $C^\infty(U)$ dans $C^\infty(U)$ est alors dit d'ordre inférieur ou égal à p en 0 si, pour tout m , on a :

$$T C_m(U) \subset C_{m-p}(U).$$

Soit $h_\alpha \in C^\infty(U)$, l'opérateur $h_\alpha(u) \partial_u^\alpha$ est de degré inférieur ou égal à p en 0 si $h_\alpha \in C_{[\alpha]-p}(U)$. En d'autres termes $\partial_u^\beta h_\alpha(0) = 0$ pour $\beta < [\alpha] - p$.

Plus généralement $T = \sum_\alpha h_\alpha(u) \partial_u^\alpha$ est de degré $\leq p$ en 0 si chaque terme $h_\alpha(u) \partial_u^\alpha$ l'est.

. On peut remarquer que si T et S sont de degrés respectifs en 0 inférieurs ou égaux à p et q , alors les opérateurs TS et $[T, S]$ sont en 0 de degré $\leq p+q$.

2.12. Théorème d'approximation.

Supposons que le système $(X_1, \dots, X_r)$ vérifie (H) pour un entier k et est libre jusqu'à l'ordre k . Considérons l'isomorphisme partiel λ_k de $\mathcal{A}_{Y,k}$ dans $\mathcal{L}(X_1, \dots, X_r)$ construit en 2.3 et le difféomorphisme ϕ_x défini sur le voisinage ouvert U de 0 dans $N_{Y,k}$ et à valeurs dans un ouvert de M , obtenu en 2.7 pour x appartenant à un voisinage de x_0 dans M .

Alors, pour $i \in \{1, \dots, r\}$,

$$\phi_x^*(X_i) = Y_i + T_i$$

où T_i est un opérateur de degré en 0 négatif ou nul.

Plus précisément dans le système de coordonnées canoniques,

$$T_i = \sum_{j=1}^{n_k} \psi_j^i(u) \frac{\partial}{\partial u_j} \quad \text{avec } \psi_j^i \in C_k(U).$$

Preuve :

. Il suffit de montrer que, si $f \in C^\infty(U)$,

$$(*) \quad \phi_x^*(X_i)f - Y_i f \in C_k(U).$$

En effet comme $\phi_x^*(X_i) - Y_i$ est un champ de vecteurs sur U , et que

$\frac{\partial}{\partial u_j}$, $j \in \{1, \dots, n_k\}$ est une base de $T_u(U)$, pour $u \in U$, nous pouvons alors écrire dans le système de coordonnées canoniques,

$$T_i = \phi_{x_0}^*(X_i) - Y_i = \sum_{j=1}^{n_k} \psi_j^i(u) \frac{\partial}{\partial u_j}, \quad \psi_j^i \in C^\infty(U).$$

Appliquant alors la relation (*) à l'application p_j désignant la $j^{\text{ème}}$ -coordonnée ($p_j(u) = u_j$), on en déduit que

$$\phi_{x_0}^*(X_i)(p_j) - Y_i(p_j) = \psi_j^i$$

est un élément de $C_k(U)$.

L'opérateur $\frac{\partial}{\partial u_j}$ étant toujours de degré $\leq k$ en 0, l'opérateur $\psi_j^i(u) \frac{\partial}{\partial u_j}$ est de degré ≤ 0 pour $j \in \{1, \dots, n_k\}$. Il en est donc de même de l'opérateur T_i .

. Nous allons donc prouver la relation (*). Posons si $u \in N_{Y,k}$, $u = e^{uY}$ où $uY = \sum_{i=1}^{n_k} u_i e_i$.

Alors si $f \in C^\infty(U)$,

$$\phi_x^*(X_i)(f)(u) = X_i(f \circ \phi_x^{-1})(\phi_x(u)) = X_i(f \circ \phi_x^{-1})(e^{\lambda_k(uY)} x).$$

Nous supprimons l'indice x de ϕ et k de λ , d'où

$$\phi^*(X_i)f(u) = \frac{d}{dt} (f \circ \phi^{-1})(e^{tX_i} e^{\lambda(uY)_x}) \Big|_{t=0}.$$

Et pour le champ invariant à gauche Y_i , nous avons

$$Y_i f(u) = \frac{d}{dr} f[e^{uY} e^{tY_i}] \Big|_{t=0}.$$

Considérons la fonction $u \rightarrow h[e^{\lambda(uY)_x}]$, pour $h \in C^\infty(M)$; cette fonction admet pour série de Taylor en 0, la série formelle $e^{\lambda(uY)_x} h(x)$. En effet si on considère pour $t \in \mathbb{R}$ petit la fonction $\gamma(t) = e^{t\lambda(uY)_x}$, on a par définition de l'exponentielle,

$$\frac{d}{ds}(g \circ \gamma) = uY(g \circ \gamma)$$

et ainsi de suite $\frac{d^n}{ds^n}(g \circ \gamma) = (uY)^n(g \circ \gamma)$.

Donc $h[e^{s\lambda(uY)_x}]$ admet pour série de Taylor, pour s petit, $e^{t\lambda(uY)_x} h(x)$. Comme λ est une application linéaire, il est aisé de conclure pour u petit.

Par suite la série de Taylor de $\phi^*(X_i)(f)(u)$ est, pour u proche de 0, $e^{\lambda(uY)_x} X_i(f \circ \phi^{-1})(x)$. Ce dernier terme est lui-même le coefficient de t de la série formelle $e^{\lambda(uY)_x} e^{tX_i}(f \circ \phi^{-1})(x)$.

2.13. LEMME. Pour $f \in C^\infty(U)$, on a formellement, si u et t sont proches de 0,

$$e^{\lambda(uY)_x} e^{tX_i}(f \circ \phi^{-1}) - e^{uY} e^{tY_i} f(0) = O[|u|^k, t^2]$$

A l'aide de ce lemme, nous déduirons que la série formelle $e^{\lambda(uY)_x} X_i(f \circ \phi^{-1})(x) - e^{uY} Y_i f(0) = O[|u|^k]$. Et donc la fonction C^∞ , $\phi^*(X_i)f(u) - Y_i f(u)$ qui admet pour série de Taylor en 0, la série formelle ci-dessus, est un élément de $C_k(U)$. D'où le théorème.

Preuve du Lemme 2.13. : La formule de Campbell-Hausdorff va nous permettre d'explicitement l'opérateur $e^{\lambda(uY)_x} e^{tX_i}$: Il existe des constantes c_p , $p \in \mathbb{N}$ telles que

$$e^{\lambda(uY)_x} e^{tX_i} = e^{\lambda(uY)_x + tX_i} + t \left(\sum_{p>0} c_p [\text{Ad} \lambda(uY)]^p(X_i) \right) + O(t^2)$$

où $\text{Ad}[\lambda(uY)](X_i) = [\lambda(uY), X_i]$.

De même

$$e^{uY} e^{tY_i} = e^{uY + t(Y_i + \sum_{p \leq k} \text{Ad}(uY)^p(Y_i)) + O(t^2)}$$

puisque $N_{Y,k}$ est nilpotent de rang k . On posera $Z_i = Y_i + \sum_{p \leq k} \text{Ad}(uY)^p(Y_i)$.

$$\text{Comme } uY = \sum_{j=1}^{n_k} u_j e_j \text{ et } X_i = \lambda(Y_i) = \lambda(e_i),$$

$$\text{Ad}[\lambda(uY)]^p(X_i) = [\lambda(uY), [\dots, [\lambda(uY), X_i] \dots]]$$

$$= \sum_{\beta=(j_1, \dots, j_{p-1})} u_{\beta} X_{\beta}$$

où $u_{\beta} = u_{j_1} \dots u_{j_{p-1}}$ est homogène de degré $[\beta] = \sum_{s=1}^{p-1} [j_s]$ et où

$X_{\beta} = [\lambda(e_{j_1}), [\dots, [\lambda(e_{j_{p-1}}), \lambda(e_i)] \dots]]$ est un commutateur de longueur $[\beta]+1$ puisque $i \in \{1, \dots, r\}$.

L'homomorphisme λ étant partiel jusqu'à l'ordre k , on en déduit que si $[\beta]+1 \leq k$, $X_{\beta} = \lambda([e_{j_1} [\dots [e_{j_{p-1}}, e_i] \dots]]) = \lambda(Y_{\beta})$.

$$\text{Or } (\text{Ad} uY)^p(Y_i) = \sum_{\substack{\beta=(j_1, \dots, j_{p-1}) \\ [\beta]+1 \leq k}} u_{\beta} Y_{\beta}$$

puisque $N_{Y,k}$ est nilpotent de rang k .

$$\text{D'où } \text{Ad}[\lambda(uY)]^p(X_i) = \lambda[(\text{Ad} uY)^p(Y_i)] + \sum_{\substack{\beta=(j_1, \dots, j_{p-1}) \\ [\beta]+1 > k}} u_{\beta} Y_{\beta}.$$

Or si $[\beta] \geq k$, la norme homogène $|u_{\beta}| \leq |u|^k$ et par suite

$$\text{Ad}[\lambda(uY)]^p(X_i) - \lambda[(\text{Ad} uY)^p(Y_i)] = O[|u|^k].$$

D'où

$$e^{\lambda(uY)} e^{tX_i} (f \circ \phi^{-1}(x)) = e^{\lambda(uY + tZ_i) + O(t)^2 + O[|u|^k]} (f \circ \phi^{-1})(x).$$

Et donc si on développe le terme exponentiel de droite, on obtient formellement

$$e^{\lambda(uY)} e^{tX_i} (f \circ \phi^{-1}(x)) = e^{\lambda(uY+tZ_i)} (f \circ \phi^{-1})(x) + O[t^2, |u|^k],$$

puisque $X O[|u|^k]_{X'} = O[|u|^k]$ et $X O(t^2)_{X'} = O(t^2)$.

De même

$$e^{uY} e^{tY_i} f(0) = e^{uY+tZ_i+O(t^2)} f(0) = e^{uY+tZ_i} f(0) + O(t^2).$$

Or par construction, on a formellement,

$$e^{\lambda(uY+tZ_i)} f \circ \phi^{-1}(x) = e^{uY+tZ_i} f(0).$$

En effet ces deux séries formelles sont les séries de Taylor en $(u,t) = 0$, de la fonction $f(e^{uY+tZ_i})$ puisque par définition de ϕ , $f[e^{uY+tZ_i}] = f \circ \phi^{-1}(e^{\lambda(uY+tZ_i)} x)$.

En conséquence

$$e^{\lambda(uY)} e^{tX_i} (f \circ \phi^{-1})(x) = e^{uY} e^{tY_i} f(0) + O[|u|^k, t^2].$$

D'où le lemme.

2.14. COROLLAIRE. Sous les hypothèses du théorème 2.12, si $Y \in V_s^Y$, c'est-à-dire est un commutateur de longueur $s \leq k$,

$\phi^*(\lambda(Y)) - Y$ est un opérateur de degré local en 0 inférieur ou égal à s .

Preuve : Il suffit de considérer le cas $Y = [Y_{i_1}, \dots, [Y_{i_{s-1}}, Y_{i_s}] \dots]$ avec

$i_1, \dots, i_s \in \{1, \dots, r\}$ et $s \leq k$. Il est alors clair, puisque $\phi^*\lambda(Y_{i_1}) = Y_{i_1} + T_{i_1}$, T_{i_1} de degré ≤ 0 en 0 pour $i \in \{1 \dots r\}$, et puisque λ est un homomorphisme partiel jusqu'à l'ordre k , que

$$\begin{aligned} \phi^*\lambda(Y) &= [\phi^*\lambda(Y_{i_1}), \dots, [\phi^*\lambda(Y_{i_{s-1}}), \phi^*\lambda(Y_{i_s})] \dots] \\ &= [Y_{i_1}, \dots, [Y_{i_{s-1}}, Y_{i_s}] \dots] + T = Y + T \end{aligned}$$

où T est de degré local $\leq s$. En effet le degré du crochet de deux opérateurs est inférieur ou égal à la somme des degrés des opérateurs.

3. Intérêt de cette approximation pour l'étude des diffusions en temps petit.

Considérons sur M un système de champ de vecteurs $X_1, \dots, X_r$ libre jusqu'à l'ordre k tel que pour tout x de M , $V_k(x) = T_x(M)$.

Notons $p(t, x, y)$ la densité par rapport à un volume riemannien π du semi-groupe de transition de générateur $L = \sum_{i=1}^r X_i^2$. Cherchons à étudier $p(t, x, y)$ pour x et y proches de x_0 fixé dans M . Soit W un voisinage de x_0 et ψ une carte locale de W dans le groupe libre $N_{Y,k}$; considérons la densité de transition $p_W(t, x, y)$ de la densité induite sur W . On peut penser (cf. chap. 5) que les comportements asymptotiques de p et p_W sont similaires au voisinage de x_0 , et on est alors amené à étudier sur $\psi(W)$ l'image de la diffusion induite. On pourra prolonger à l'espace $N_{Y,k}$ le générateur de cette diffusion de façon à ce que en dehors d'un voisinage W_1 de 0 , $L = \sum_{i=1}^r Y_i^2$, et par le même argument on peut imaginer comparer les densités de transition de la diffusion sur $N_{Y,k}$ et de la diffusion sur $\psi(W)$. On est donc amené à supposer que $M = N_{Y,k}$ et L du type ci-dessus.

On sait (2.7) que pour tout y_0 de $M = N_{Y,k}$, il existe un voisinage ouvert V de y_0 dans $M = N_{Y,k}$ et un voisinage ouvert U de 0 dans $N_{Y,k}$ tel que, pour $y \in V$, l'application ϕ_y (définie en 2.7) soit un difféomorphisme de U sur un ouvert de M contenant V . De plus $\phi_y^*(X_i) = Y_i + T_i$ où T_i est de degré local négatif ou nul en 0 . Par suite $\phi_y^*(L) = \sum_{i=1}^r Y_i^2 + R_y$ où R_y est un opérateur de degré local en 0 inférieur ou égal à 1 .

Considérons $(Y_i)_{1 \leq i \leq n_k}$ la base canonique (2.9) de champs invariants à gauche de $N_{Y,k}$ et prolongeons, pour chaque y de V , ϕ_y en un difféomorphisme ϕ_y de $M = N_{Y,k}$ sur $N_{Y,k}$ de façon à ce que l'opérateur $\phi_y^*(L)$, écrit à l'aide de la base $(Y_i)_{1 \leq i \leq n_k}$, ait des coefficients uniformément bornés, ainsi que leurs différentielles. Nous ferons en fait ce prolongement pour tout y appartenant à un voisinage compact de 0 contenant strictement le voisinage $\overline{W_1}$. En dehors de $\overline{W_1}$, $L = \sum_{i=1}^r Y_i^2$ et ϕ_y est seulement la translation à gauche par y . Les majorants obtenus ci-dessus seront donc uniformes pour $y \in N_{Y,k}^k$.

Appelons alors $r(t, x, y)$ la densité, par rapport à la mesure de Haar m_D sur $N_{Y,k}$, du semi-groupe de transition de générateur $\sum_{i=1}^r Y_i^2$. Posons $q(t, x, y) = r[t, \phi_y^{-1}(x), \phi_y^{-1}(y)] = r(t, \phi_y^{-1}(x), 0)$ pour x et $y \in N_{Y,k}$. L'idée est de considérer $q(t, x, y)$ comme une paramétrix dans l'étude de la densité $p(t, x, y)$, du semi-

groupe de transition de générateur L , densité que l'on calculera par rapport au volume π vérifiant, $\pi(y) = \phi_y(m_D)$. Cette méthode est classique dans le cas de $\mathbb{R}^m$ (cf. chapitre 4).

En effet considérons sur N_Y^k l'opérateur gelé en y , $L_y = (\phi_y^{-1})^* (\sum_{i=1}^r Y_i^2)$, alors $(L_y - \frac{\partial}{\partial t})q(t, x, y) = 0$.

Comme de plus $(L - \frac{\partial}{\partial t})p(t, x, y) = 0$, on en déduit que

$$\begin{aligned} (L - \frac{\partial}{\partial t})(p-q)(t, x, y) &= (L_y - L)q(t, x, y) \\ &= \phi_y^*(L_y - L)r(t, \phi_y^{-1}(x), \phi_y^{-1}(y)) \\ &= S(t, x, y). \end{aligned}$$

Par suite, si on pose $p * S = \int_{N_{Y,k}} \int_0^t p(t-s, x, z) S(s, z, y) d\pi(z) ds$, $p-q$ doit être

de la forme $p * S$ et on est amené à chercher p sous la forme $q + \sum_{n \geq 1} q * S^{*n}$.

Il s'agit donc d'étudier $S(t, x, y)$ et ses itérés $S^{*n}(t, x, y)$ et de s'assurer qu'ils sont suffisamment "petits" pour t petit. En fait $S(t, x, y) = R_y r(t, \phi_y^{-1}(x), 0)$ où $R_y = \sum_{i=1}^r Y_i^2 - \phi_y^*(\sum_{i=1}^r X_i^2)$ est de degré local inférieur ou égal à 1 en 0. On peut penser que ce degré local, contrebalançant les singularités des différentielles de r permettra de trouver un majorant de $S(t, x, y)$ pour x proche de y ; quant aux points x éloignés de y , l'uniformité bornitude des coefficients de R_y et de leurs différentielles, ainsi que la décroissance très rapide de r et de ses différentielles lorsque x s'éloigne, devrait permettre de répondre !

Des travaux ont été entrepris dans cette direction par C. Bellaïche et M. Chaleyat-Maurel sur le cas du Laplacien de Kohn sur le groupe d'Heisenberg où r est donné "explicitement". Mais les méthodes classiques d'itération sur $\mathbb{R}^m$ ne sont pas applicables et les calculs sont difficiles à mener.

BIBLIOGRAPHIE,

- [1] B. Gaveau, *Principe de moindre action, propagation de la chaleur, estimées sous elliptiques sur certains groupes nilpotents*. Acta Mathematica. vol. 139. (1977).
- [2] B. Gaveau, *Systèmes dynamiques associés à des opérateurs hypoelliptiques*. Bull. Sc. Math 102 p.203-229 (1978).

- [3] R.W. Goodman, *Nilpotent Lie Groups*. Lecture Notes n°562.
- [4] B. Hellfer et J.F. Nourrigat, *Approximation d'un système de champ de vecteurs et applications à l'hypoellipticité*. Séminaire Goulaouic-Schwartz (1978).
- [5] G. Métivier, *Fonction spectrale et valeurs propres d'une classe d'opérateurs non elliptiques*. Comm. in P.D.E. 1 (1976) p.467-519.
- [6] L.P. Rothschild et E.M. Stein, *Hypoelliptic differential operators and nilpotent groups*. Acta Mathematica 137 p.248-315.

CHAPITRE 15

DIFFUSIONS GAUSSIENNES

par

Mireille CHALEYAT-MAUREL et Laure ELIE

0. Introduction.

Dans ce chapitre nous allons étudier un exemple de diffusions à générateur hypoelliptique sur $\mathbb{R}^m$: les diffusions gaussiennes. Cet exemple est intéressant dans la mesure où il permet d'explorer l'étude d'un certain type d'opérateurs hypoelliptiques.

Plus précisément, soit M une variété C^∞ de dimension m et $X_1, \dots, X_r, Y$, des champs de vecteurs C^∞ sur M . On considère l'opérateur semi elliptique :

$$L = \frac{1}{2} \sum_{i=1}^r X_i^2 + Y.$$
 Nous savons que cet opérateur est hypoelliptique au sens de Hörmander si l'algèbre de Lie $\mathcal{L}(X_1, \dots, X_r, Y)$ engendrée par $X_1, \dots, X_r, Y$ est de dimension m en tout point.

Nous pouvons considérer les trois classes d'opérateurs suivantes :

① La dimension de $\mathcal{L}(X_i, [X_i, Y], \dots, \underbrace{[[X_i, Y], \dots, Y]}_{m-1 \text{ fois}}, i = 1, \dots, r)$ est m en tout point de M . Cette hypothèse garantit l'hypoellipticité de L et de $L - \frac{\partial}{\partial t}$ et donc l'existence de la densité de transition associée à la diffusion de générateur L . (cf. R. Azencott. Chap.2).

② $\mathcal{L}(X_1, \dots, X_r)$ est de dimension m en tout point; cet exemple ayant été traité par B. Gaveau sur certains groupes nilpotents ([6]).

③ Le cas involutif : on suppose que pour tout $k \in \{0, 1, \dots, m-1\}$, l'espace vecto-

riel $\mathcal{H}_k$ engendré par les champs $\{X_i, [X_i, Y], \dots, [X_i, \underbrace{Y, Y, \dots, Y}_{k \text{ fois}}]\}$ $i = \{1, \dots, r\}$ est involutif en tout point (c'est-à-dire est égal en tout point de M à l'algèbre de Lie engendrée par ces champs), et on suppose qu'il existe $p \in \mathbb{N}$ (évidemment inférieur ou égal à $m-1$), tel que la dimension de $\mathcal{H}_p$ soit m .

Autrement dit, si on regarde l'algèbre de Lie engendrée par les X_i et tous les crochets des X_i et des Y , $i \in \{1, \dots, r\}$, elle est en tout point égale à $\mathcal{H}_p$ et seuls les crochets du type suivant $[[X_i, Y], \dots, Y]$, $k \in \{0, \dots, m-1\}$ sont susceptibles de contribuer à la construction de l'espace k fois tangent en tout point. Les classes ② et ③ sont des sous-classes de ① et les éléments communs à ces deux classes sont les opérateurs elliptiques.

Les exemples les plus simples d'opérateurs du dernier type sont les générateurs différentiels de diffusions gaussiennes; et ce sont en fait ces diffusions qui motivent l'introduction de cette classe ③ car elles sont un bon outil d'approximation des opérateurs de cette classe.

Par exemple, considérons un opérateur $L = \sum_{i,j=1}^m a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x) \frac{\partial}{\partial x_i}$ sur $\mathbb{R}^m$, élément de la classe ③. On dira que L est défini par la matrice $a(x) = [a_{ij}(x)]_{1 \leq i,j \leq m}$ et le vecteur $b(x) = [b_i(x)]_{1 \leq i \leq m}$. Alors on peut remarquer que l'opérateur L_ξ figé en ξ défini par la matrice constante $a(\xi)$ et le vecteur $b(\xi) + \text{Jac}b(\xi) [x-\xi]$, appartient à la classe ③ et est le générateur d'une diffusion gaussienne. Ces opérateurs L_ξ approchent correctement L (on a tout simplement fait un développement du 1er ordre de b) et leurs densités de transition peuvent servir de paramétrix dans l'étude de la densité de transition associée à L . Ceci sera développé dans (L. Elie [5]). Comme on peut calculer explicitement les densités de transition associées aux L_ξ , on obtient des majorants et certains équivalents de la densité de transition de L .

Dans ce chapitre, nous allons étudier tout simplement les diffusions gaussiennes. Le calcul explicite de leurs densités de transition lorsqu'elles existent (c'est-à-dire exactement lorsque leurs générateurs appartiennent à la classe ③) suscitera l'introduction au paragraphe 3 d'une fonctionnelle de Cramer permettant d'interpréter ces densités.

Au paragraphe 4, nous évaluerons, si X_t désigne la trajectoire de la diffusion, l'expression $\log P_x[X_t \in A]$ où A est une partie de $\mathbb{R}^m$. Nous verrons que cette expression peut être approchée par des termes de la forme $\frac{C(A)}{k(A)}$ où $C(A) \in \mathbb{R}^+$ et où $k(A)$ est un entier impair dépendant de A . Les estimations du type Ventcell et Freidlin donnent une approximation de $t \log P[X_t \in A]$, qui peut être éventuellement infinie (Azencott [1]). Nous avons recherché ici la bonne puissance $k(A)$ de t qui permette d'approximer $t^{k(A)} \log P[X_t \in A]$ de façon satisfaisante. Les résultats que nous obtenons seront des prototypes dans l'étude des opérateurs du type ③ (M. Chaleyat-Maurel et L. Elie [2]).

1. Diffusions Gaussiennes.

1.1. Cadre. Soit sur $\mathbb{R}^m$ l'opérateur $L = \frac{1}{2} \sum_{i=1}^r X_i^2 + Y$ où les champs de vecteurs

$X_1, \dots, X_r$ et Y sont de classe C^∞ .

Si $(x_1, \dots, x_m)$ désigne un système de coordonnées, on désigne par

$\sigma(x) = [\sigma_{i,j}(x)]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq r}}$ le champ C^∞ de matrices rectangulaires défini par :

$$X_j = \sum_{i=1}^m \sigma_{i,j}(x) \frac{\partial}{\partial x_i}, \quad 1 \leq j \leq r.$$

Alors si on pose $a(x) = \sigma(x)^t \sigma(x) = [a_{i,j}(x)]_{1 \leq i,j \leq m}$, on obtient :

$$L = \frac{1}{2} \sum_{i,j=1}^m a_{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x) \frac{\partial}{\partial x_i}$$

$$\text{où } b_i(x) = y_i(x) + \sum_{j,k=1}^m \sigma_{j,k}(x) \frac{\partial \sigma_{i,k}(x)}{\partial x_j} \quad \text{si } Y = \sum_{i=1}^m y_i(x) \frac{\partial}{\partial x_i}.$$

1.2. PROPOSITION. Les deux assertions suivantes sont équivalentes.

$$(i) \quad L = \frac{1}{2} \sum_{i,j=1}^m a_{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x) \frac{\partial}{\partial x_i} \quad \text{est le générateur}$$

différentiel d'une diffusion gaussienne homogène.

(ii) $a(x) = a$, $b(x) = bx + c$ avec a et b deux matrices $m \times m$ constantes et c un vecteur constant.

On notera $L(a,b,c)$ un tel générateur.

Pour la démonstration de cette proposition, on pourra par exemple consulter E. Nelson ([9]).

1.3. PROPOSITION. Soit $L(a,b,c)$ le générateur d'une diffusion gaussienne homogène sur $\mathbb{R}^m$. Alors L vérifie la condition d'hypoellipticité ① si et seulement si la matrice $m \times m^2$ suivante

$$\begin{bmatrix} a & ba & b^2 a & \dots & b^{m-1} a \end{bmatrix}$$

est de rang m .

De plus l'opérateur L est alors hypoelliptique au sens ③.

Démonstration. Notons $\tilde{X}_i$, $1 \leq i \leq m$, le $i^{\text{ème}}$ vecteur colonne de la matrice a , c'est-à-dire que $\tilde{X}_i = \sum_{j=1}^m a_{j,i} \frac{\partial}{\partial x_j}$. Comme les champs de vecteurs $X_1, \dots, X_r$ et

$\tilde{X}_1, \dots, \tilde{X}_m$ sont constants et comme $a = \sigma^t \sigma$, l'espace vectoriel $\mathcal{H}_0$ engendré par $X_1, \dots, X_r$ est identique à celui engendré par $\tilde{X}_1, \dots, \tilde{X}_m$. Par suite il est équivalent de vérifier l'hypothèse ① pour les champs $X_1, \dots, X_r, Y$ ou pour les champs $\tilde{X}_1, \dots, \tilde{X}_m, Y$.

Or on remarque :

$$\begin{aligned} [\tilde{X}_i, Y] &= \left[\sum_{j=1}^m a_{j,i} \frac{\partial}{\partial x_j}, \sum_{k=1}^m \left(\sum_{p=1}^m b_{kp} x_p + c_k \right) \frac{\partial}{\partial x_k} \right] \\ &= \sum_{k=1}^m \left(\sum_{j=1}^m b_{kj} a_{ji} \right) \frac{\partial}{\partial x_k}, \end{aligned}$$

et donc $[\tilde{X}_i, Y]$ est le $i^{\text{ème}}$ vecteur colonne de ba , et ainsi de suite $\underbrace{[\dots [\tilde{X}_i, Y], Y] \dots Y]}_{k \text{ fois}}$ est le $i^{\text{ème}}$ vecteur colonne de $b^k a$.

Comme ces champs de vecteurs-ci sont constants, l'espace vectoriel $\mathcal{H}_k$ engendré par $\{\tilde{X}_i, [\tilde{X}_i, Y], \dots, \underbrace{[\tilde{X}_i, Y], \dots, Y}_{k \text{ fois}}\}$ $i = 1, \dots, m$ est involutif et l'opérateur L vérifie la condition ① si et seulement si il existe un p tel que la dimension de $\mathcal{H}_p$ soit m . Nécessairement ce p sera inférieur ou égal à $m-1$ puisque si $\mathcal{H}_k = \mathcal{H}_{k+1}$, on a $\mathcal{H}_k = \mathcal{H}_j$ pour tout $j \geq k$. D'où la proposition.

Comme de plus les espaces vectoriels $\tilde{\mathcal{H}}_k$ et $\mathcal{H}_k$ (définis dans l'introduction à l'aide des $X_1, \dots, X_r, Y$) sont identiques, on en déduit que si L vérifie l'hypothèse ①, il est hypoelliptique au sens ③.

1.4. Forme des diffusions gaussiennes. Une diffusion gaussienne homogène associée à L vérifie l'équation différentielle stochastique :

$$(1) \quad dX_t = a dB_t + (bX_t + c)dt$$

où $a = \sigma^t \sigma$ et B_t est un brownien r -dimensionnel si σ est une matrice rectangulaire (m, r) .

Il y a existence et unicité globale sur $\mathbb{R}^m$ de la solution de (1) partant de x , élément de $\mathbb{R}^m$, au temps zéro. Cette solution s'explicite par la formule suivante :

$$(2) \quad X_t = (\exp tb) \cdot x + \int_0^t (\exp(t-s)b) \cdot \sigma dB_s + \int_0^t (\exp sb) \cdot c ds.$$

En dimension 1, on retrouve le processus d'Ornstein-Uhlenbeck (voir E. Nelson [9]).

1.5. Exemples.

α). Si $\det(a) \neq 0$, le générateur L est elliptique.

β). Plaçons-nous sur $\mathbb{R}^2$ et posons $a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $c \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Alors on peut choisir $\sigma = (1, 0)$ et la diffusion X_t s'exprime à l'aide d'un brownien uni-dimensionnel B_t^1 et s'écrit sous la forme

$$\begin{cases} X_t^1 = B_t^1 + X_0^1 \\ X_t^2 = \int_0^t B_s^1 ds + X_0^2 + t X_0^1 \end{cases}$$

Cet exemple est l'exemple le plus simple de diffusion gaussienne dégénérée hypoelliptique.

Plus généralement, nous allons voir que cette forme particulière d'expression de la diffusion à l'aide de brownien et d'intégrale de brownien est liée au caractère nilpotent de la matrice b .

1.6. PROPOSITION. Soit $L(a, b, c)$ le générateur d'une diffusion gaussienne homogène X_t sur $\mathbb{R}^m$. Si la matrice a est de rang n_0 , alors X_t ne dépend que d'un brownien n_0 -dimensionnel B_t .

Si de plus la matrice b est nilpotente d'ordre p , alors X_t s'exprime comme application linéaire de B_t , $\int_0^t B_s ds$, $\int_0^t ds_1 [\int_0^t B_s ds]$, ..., $\int_0^t \dots \underbrace{\int_0^{s_2} ds_1 [\int_0^{s_1} B_s ds]}_{p-1}$ et d'une fonction $c(t)$ non aléatoire.

Démonstration. Comme a est de rang n_0 , il en est de même de σ ; et par suite σB_s est l'image par une application de matrice σ_0 rectangulaire $m \times m_0$ d'un brownien n_0 -dimensionnel. La première assertion est alors immédiate à l'aide de (2).

Si la matrice b est nilpotente d'ordre p , nous avons $b^p = 0$ et d'après (2)

$$X_t = (\exp tb) \cdot x_0 + \int_0^t (\exp sb) \cdot c ds + \int_0^t \sum_{k=0}^{p-1} \frac{(t-s)^k}{k!} b^k \sigma dB_s.$$

Or d'après une application simple de la formule d'Ito,

$$\int_0^t \frac{(t-s)^k}{k!} b^k \sigma dB_s = \underbrace{\int_0^t \dots \int_0^{s_2} ds_1}_{k} \left(\int_0^{s_1} b^k \sigma B_s ds \right),$$

$$\text{d'où} \quad X_t = c(t) + \underbrace{\sum_{k=0}^{p-1} \int_0^t \dots \int_0^{s_2} ds_1}_{k} \left(\int_0^{s_1} b^k \sigma B_s ds \right).$$

1.7. Remarque. Notons $H_m = a(\mathbb{R}^m) + ba(\mathbb{R}^m) + \dots + b^{m-1}a(\mathbb{R}^m)$.

Alors la variable $\int_0^t (\exp(t-s)b)\sigma dB_s$ prend ses valeurs dans H_m . Par suite si c est un vecteur de H_m , le processus X_t s'il part d'un point x de H_m reste dans H_m , et le générateur de ce processus induit sur H_m vérifie la condition d'hypoellipticité (1). En d'autres termes, on pourrait dire que toute diffusion gaussienne sur $\mathbb{R}^m$ s'écrit

comme somme de processus non aléatoire et d'un processus porté par un sous-espace vectoriel de $\mathbb{R}^m$ dont la projection sur ce sous-espace est une diffusion hypoelliptique (vérifiant ①).

1.8. Calcul de la matrice de covariance et de la densité de transition.

Soit L le générateur d'une diffusion gaussienne X_t . Il suffira pour connaître la densité $p(t, x, y)$, lorsqu'elle existe, du semi-groupe de transition associé, de déterminer la matrice de covariance du processus X_t :

$$\begin{aligned} K(t) &= E [(X_t - E(X_t)) {}^t(X_t - E(X_t))] \\ &= E \left(\left(\int_0^t (\exp(t-s)b) \cdot \sigma dB_s \right) {}^t \left(\int_0^t (\exp(t-s)b) \cdot \sigma dB_s \right) \right) \\ &= (\exp tb) \left(\int_0^t (\exp sb) \cdot a \cdot (\exp -s {}^tb) ds \right) (\exp {}^tb.) \end{aligned}$$

$$\text{D'où } K(t) = \int_0^t (\exp sb) \cdot a \cdot (\exp s {}^tb) ds.$$

Par suite on obtient le résultat suivant :

1.9. PROPOSITION. Soit $L(a, b, c)$ le générateur d'une diffusion gaussienne sur $\mathbb{R}^m$. Si la matrice $K(t) = \int_0^t (\exp sb)a(\exp s {}^tb)ds$ est inversible pour tout $t > 0$, le semi-groupe de transition de générateur L admet pour densité par rapport à la mesure de Lebesgue la fonction $p(t, x, y)$ définie sur $\mathbb{R}^+ \times \mathbb{R}^m \times \mathbb{R}^m$ définie par :

$$p(t, x, y) = \frac{1}{(\sqrt{2\pi})^{m/2}} \frac{1}{\sqrt{\det K(t)}} \exp - S_t^b(x, y)$$

$$\text{où } S_t^b(x, y) = \frac{1}{2} {}^t \varphi(t, x, y) K^{-1}(t) \varphi(t, x, y)$$

$$\text{avec } \varphi(t, x, y) = y - (\exp tb) \cdot x - \int_0^t (\exp sb) \cdot c ds.$$

1.10. LEMME. La matrice $K(t) = \int_0^t (\exp sb)a(\exp s {}^tb)ds$ est inversible pour $t > 0$ si et seulement si la matrice $A = [a \ b a \ \dots \ b^{m-1} \ a]$ est de rang m .

Preuve : Soit $u \in \mathbb{R}^m$ tel que $K(t)u = 0$. Comme $\exp sb \cdot a \exp s {}^tb$ est une matrice symétrique de type positif, on en déduit que pour tout $s \in]0, t[$, $\exp sb \cdot a \cdot \exp s {}^tb \cdot u$ est nul. Comme $\exp sb$ est une matrice inversible, il en résulte que, pour tout $s \in]0, t[$, $a \exp s {}^tb \cdot u$ est nul, et donc que pour tout entier k positif ou nul $a {}^tb^k u = 0$. Par suite u appartient au noyau de la matrice tA . Réciproquement soit u tel que ${}^tAu = 0$. Alors, pour tout entier k compris entre 0 et $m-1$, $a {}^tb^k u = 0$. Il découle du théorème de Cayley-Hamilton, qu'en fait $a {}^tb^k u = 0$ pour tout k et par suite u appartient au noyau de $K(t)$.

En conséquence $K(t)$ est inversible si et seulement si le noyau de tA est réduit à $\{0\}$, c'est-à-dire si et seulement si le rang de A est m .

2. Diffusions gaussiennes hypoelliptiques.

2.1. Soit $L = \frac{1}{2} \sum_{i=1}^r X_i^2 + Y$ le générateur différentiel d'une diffusion gaussienne sur $\mathbb{R}^m$. On considère la suite croissante des espaces vectoriels $\mathcal{H}_k$, $k \in \mathbb{N}$, engendrés par les champs constants $\{X_i, [X_i, Y], \dots, [[X_i, Y] \dots Y], i \in \{1, \dots, r\}\}$. En fait $\mathcal{H}_k$ est pour $k > 0$ l'espace vectoriel engendré par $\mathcal{H}_{k-1}$ et $[X_{k-1}, Y]$. On notera n_k la dimension de $\mathcal{H}_k$.

Si L est un générateur hypoelliptique, il existe un entier $p \leq m-1$ pour lequel $\mathcal{H}_p$ est de dimension m .

2.2. DÉFINITIONS. Soit L le générateur différentiel hypoelliptique d'une diffusion gaussienne sur $\mathbb{R}^m$. Une base de $\mathbb{R}^m$ sera dite adaptée à L si, dans le système de coordonnées associé, pour tout $i \in \{0, \dots, p\}$, $\{\frac{\partial}{\partial x_k}, 1 \leq k \leq n_i\}$ est une base de $\mathcal{H}_i$. On appellera N_i l'espace engendré par les éléments de la base compris entre les rangs n_{i-1} et n_i . Alors $\mathbb{R}^m = \bigoplus_{i=0}^p N_i$. De plus pour $k \in \{1, \dots, m\}$, on notera $n(k)$ l'entier $i \in \{1, \dots, p\}$ tel que $n_{i-1} \leq k < n_i$ si $k > n_0$ et $n(k) = 0$ si $k \leq n_0$.

2.3. LEMME. Soit L le générateur différentiel hypoelliptique d'une diffusion gaussienne sur $\mathbb{R}^m$. Dans une base adaptée à L , l'opérateur L s'écrit $L(a, b, c)$ où

$$a = \begin{bmatrix} a_0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{avec } a_0 \text{ matrice carrée } n_0 \times n_0 \text{ inversible}$$

et où

$$b = \begin{bmatrix} & n_0 & & n_1 & & n_2 & & \\ & \vdots & & \vdots & & \vdots & & \\ n_0 & * & & * & & * & & * \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & b_1 & & * & & * & & * \\ n_1 & - & - & - & - & - & - & - \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & 0 & & b_2 & & * & & * \\ n_2 & - & - & - & - & - & - & - \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & 0 & & 0 & & b_3 & & * \\ n_3 & - & - & - & - & - & - & - \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & 0 & & 0 & & 0 & & \vdots \end{bmatrix}$$

On appellera d la matrice $m \times m$ du type ci-dessus où les matrices $*$ sont remplacées par des matrices nulles. La matrice d est donc seulement fonction de b_k , $k \in \{1, \dots, p\}$ et est une matrice nilpotente d'ordre $p+1$.

Preuve. Comme l'espace vectoriel $\mathcal{H}_0$ engendré par $X_1, \dots, X_r$ a pour base

$\left\{ \frac{\partial}{\partial x_i} , 1 \leq i \leq n_0 \right\}$, la forme de a est immédiate.

Pour tout $k \geq 1$, $\mathcal{H}_k$ est l'espace vectoriel engendré par

$$\left\{ \frac{\partial}{\partial x_i} , \left[\frac{\partial}{\partial x_i} , Y \right] , i \in \{1, \dots, n_0 + \dots + n_{k-1}\} \right\} . \text{ Or}$$

$$\left[\frac{\partial}{\partial x_i} , Y \right] = \left[\frac{\partial}{\partial x_i} , \sum_{j=1}^m (bx+c)_j \frac{\partial}{\partial x_j} \right] = \sum_{j=1}^m b_{ji} \frac{\partial}{\partial x_j} .$$

Donc, puisque $\left[\frac{\partial}{\partial x_i} , Y \right]$ appartient à $\mathcal{H}_k$ si $i \leq n_{k-1}$, on doit avoir $b_{ji} = 0$ pour $j > n_k$. D'où la forme de b . On note, pour $k \in \{1, \dots, p\}$, b_k la matrice $n_k \times n_{k-1}$ de coefficient $c_{i,j} = b_{n_{k-1}+i, n_{k-2}+j}$ avec $n_1 = 0$. On peut remarquer que de par la construction des espaces $\mathcal{H}_k$, la suite des entiers $n_k - n_{k-1}$ est décroissante et que chaque matrice b_k est de rang $n_k - n_{k-1}$.

2.4. PROPOSITION. Soit L le générateur différentiel hypoelliptique d'une diffusion gaussienne sur $\mathbb{R}^m$. Dans une base de $\mathbb{R}^m$ adaptée à L , la matrice de covariance $K(t)$ de la diffusion s'écrit $U(t)B(t)U(t)$ où $U(t)$ est la matrice diagonale de termes diagonaux $t^{(2n(i)+1)/2}$ pour $i \in \{1, \dots, m\}$ et où la matrice $B(t)$ tend lorsque $t \rightarrow 0$ vers une matrice B inversible ne dépendant que de a et de d . Plus précisément, $B(t) = B + tC(t)$ où $C(t)$ est une matrice bornée au voisinage de 0 .

Preuve. Considérons $\tilde{X}_t$ la diffusion gaussienne de générateur $\tilde{L}(a, d, 0)$ et notons $\tilde{K}(t)$ sa matrice de covariance. La matrice d est nilpotente d'ordre $p+1$. Alors

$$\begin{aligned} \tilde{K}(t) &= \int_0^t (I + sd + \dots + \frac{s^p}{p!} d^p) a (I + s^t d + \dots + \frac{s^p}{p!} t^p d) ds \\ &= ta + \frac{t^2}{2} (da + a^t d) + \dots + \frac{t^k}{k} \left(\sum_{i=0}^{k-1} \frac{d^{k-1-i} a^t d^i}{(k-1-i)! i!} \right) + \dots + \frac{t^{2p+1}}{2p+1} \frac{d^p a^t d^p}{(p!)^2} \end{aligned}$$

Or de par la forme de d et a , nous avons pour $i, j \in \{0, \dots, p\}$,

$$d^i a^t d^j = \begin{matrix} & & n_{j-1} & & n_j & & \\ & & \vdots & & \vdots & & \\ n_{i-1} & \begin{bmatrix} 0 & \vdots & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & D_{i,j} & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & 0 & \vdots & 0 \end{bmatrix} & & \\ n_i & \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} & & \end{matrix} \quad \text{avec } D_{i,j} = b_i \dots b_1 a_0^t b_1 \dots b_j$$

Comme les blocs $D_{i,j}$ des différentes matrices $d^i a^t d^j$ sont tous disjoints pour $0 \leq i, j \leq p$, nous obtenons

$$\tilde{K}(t) = U(t) \tilde{K}(1) U(t),$$

si $U(t)$ est la matrice diagonale de termes diagonaux $t^{(2n(i)+1)/2}$ pour $i \in \{1, \dots, m\}$.

De plus les blocs $D_{k,0}$, matrice $(n_k - n_{k-1}) \times n_0$ sont de rang $n_k - n_{k-1}$ pour tout $k \in \{0, \dots, p\}$. On en déduit aisément que la matrice $m \times m^2$ $[a \ da \ \dots \ d^{m-1} \ a]$ est de rang m et par suite $\tilde{L}$ est un opérateur hypoelliptique. Les matrices $\tilde{K}(t)$, et en particulier $\tilde{K}(1)$ sont donc inversibles.

Enfin si $\tilde{K}_{i,j}$ désigne le bloc de $\tilde{K}(1)$ formé des coefficients $\tilde{k}_{\ell,s}$ avec $n_{i-1} \leq \ell < n_i$ et $n_{j-1} \leq s < n_j$, on a

$$\tilde{K}_{i,j} = \frac{b_i \dots b_1 a_0 t b_1 \dots t b_j}{i! j! (i+j+1)}.$$

Nous allons maintenant approximer $K(t)$ par $\tilde{K}(t)$ ce qui revient à "approcher" la diffusion X_t par $\tilde{X}_t$. Nous pouvons écrire $b = d + f$, alors $b^i a t b^j = d^i a t d^j + F^{i,j}$ où la matrice $F^{i,j}$ est de la forme suivante

$$F^{i,j} = \begin{matrix} & & n_{j-1} & & n_j & & \\ & & * & | & * & | & 0 \\ n_{i-1} & & \dots & | & \dots & | & \dots \\ & & * & | & 0 & | & 0 \\ n_i & & \dots & | & \dots & | & \dots \\ & & 0 & | & 0 & | & 0 \end{matrix}$$

et en particulier tous les blocs $F_{k,\ell}^{i,j}$ sont nuls si $k + \ell \geq i + j$.

Par suite si on cherche à calculer le bloc $K_{i,j}(t)$ de $K(t)$ (c'est-à-dire la matrice des coefficients $k_{\ell,s}$ avec $n_{i-1} \leq \ell < n_i$ et $n_{j-1} \leq s < n_j$), on aura, pour $i, j \in \{0, \dots, p\}$,

$$K_{i,j}(t) = \frac{d^i a t d^j}{i! j!} \frac{t^{i+j+1}}{i+j+1} + t^{i+j+2} C_{i,j}(t) \text{ et } C_{i,j}(t) \text{ converge}$$

lorsque $t \rightarrow 0$. D'où $K(t) = U(t) [\tilde{K}(1) + tC(t)] U(t)$ si $C(t)$ est la matrice de blocs $C_{i,j}(t)$. Il suffit alors pour conclure de poser $B = \tilde{K}(1)$.

2.5. COROLLAIRE. Soit L le générateur différentiel hypoelliptique d'une diffusion gaussienne sur $\mathbb{R}^m$; et soit dans une base adaptée à L , la densité continue $p(t, x, y)$ définie sur $\mathbb{R}^+ \times \mathbb{R}^m \times \mathbb{R}^m$ du semi-groupe de transition de générateur L , alors, lorsque $t \rightarrow 0$,

$$\log p(t, x, y) \sim -\frac{1}{2} t (y-x) U^{-1}(t) B^{-1} U^{-1}(t) (y-x)$$

où les matrices B et $U(t)$ ont été définies dans la proposition 2.4.

Plus précisément si $y_i - x_i$ désigne la projection sur N_i de $y - x$, nous obtenons

$$\log p(t, x, y) \sim -\frac{1}{2} \frac{1}{t^{2i+1}} t^{2i+1} (y_i - x_i) c_i (y_i - x_i)$$

avec c_i matrice $n_i - n_{i-1} \times n_i - n_{i-1}$, si $i = \sup \{k \in \{0, \dots, p\}, y_k - x_k \neq 0\}$.

Preuve. D'après la proposition 1.9, $\log p(t, x, y) \sim -S_t^b(x, y)$ lorsque $t \rightarrow 0$.

Or $S_t^b(x, y) = \frac{1}{2} t \left[y - (\exp tb).x - \int_0^t (\exp sb).c \, ds \right] K^{-1}(t) \left[y - (\exp tb).x - \int_0^t (\exp sb).c \, ds \right]$

et $y - (\exp tb).x - \int_0^t (\exp sb).c \, ds = y - x + t h(t)$

où $h(t)$ est bornée au voisinage de 0. Le corollaire est alors immédiat, les matrices c_i se calculant à partir de B^{-1} .

Signalons une note de D. Manankiandrianana ([7]) traitant un exemple de ces diffusions gaussiennes.

3. Transformée de Cramer. Hamiltonien. Interprétation de la densité en temps petit à l'aide de l'action.

3.1. Transformée de Cramer sur $[0, t]$.

Considérons sur $\mathbb{R}^m$ un opérateur différentiel du second ordre L , annulant les constantes et à coefficients C^∞ . Cet opérateur s'écrit donc sous la forme

$$L = \frac{1}{2} \sum_{1 \leq i, j \leq m} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1 \leq i \leq m} b_i(x) \frac{\partial}{\partial x_i}.$$

On notera $a(x)$ la matrice $[a_{ij}(x)]_{1 \leq i, j \leq m}$ et $b(x)$ le vecteur de coordonnées $(b_i(x))_{1 \leq i \leq m}$.

On considère alors sur $\mathbb{R}^m$ la forme quadratique Q_x définie par la matrice $a(x)$ et on appelle forme quadratique duale de Q_x (cf. Azencott chap. 12), l'application :

$Q_x^* : \mathbb{R}^m \rightarrow [0, \infty]$ définie, pour $v \in \mathbb{R}^m$, par

$$Q_x^*(v) = \sup_{w \in \mathbb{R}^m} [2 \langle v, w \rangle - Q_x(w)]$$

si $\langle, \rangle$ désigne le produit scalaire euclidien sur $\mathbb{R}^m \times \mathbb{R}^m$.

L'espace des trajectoires explosives définies sur $[0, t]$ à valeurs dans $\mathbb{R}^m \cup \delta$, où δ est le point à l'infini du compactifié d'Alexandroff, sera noté $\mathcal{E}_t(\mathbb{R}^m)$ et $\xi(\gamma)$ sera le temps d'explosion de la trajectoire γ sur $[0, t]$.

On appelle transformée de Cramer sur $[0, t]$ de (L, b) la fonctionnelle :

$$\lambda_t^b : \mathcal{E}_t(\mathbb{R}^m) \rightarrow [0, \infty]$$

définie par :

$$\lambda_t^b(\gamma) = \frac{1}{2} \int_0^{t \wedge \xi(\gamma)} Q_{\gamma_s}^* [\gamma'_s - b(\gamma_s)] \, ds \quad \text{si } \gamma'_s \text{ existe presque partout,}$$

et $\lambda_t^b(\gamma) = +\infty$ si γ n'est pas absolument continue.

3.2. Hamiltonien.

On appellera hamiltonien associé à (L, b) , la fonctionnelle définie sur $\mathbb{R}^m \times \mathbb{R}^m$ par :

$$H(\gamma, q) = \sup_{\xi \in \mathbb{R}^m} \left[-\frac{1}{2} Q_{\gamma}^*(\xi - b(\gamma)) + \langle q, \xi \rangle \right]$$

Remarquons que c'est dans [7] qu'à été, pour la première fois, introduit ce type de Hamiltonien.

3.3. LEMME. Q, Q^* et H sont liés par les formules suivantes :

$$(i) \quad H(\gamma, q) = \frac{1}{2} Q_{\gamma}(q) + \langle q, b \rangle$$

$$(ii) \quad \frac{1}{2} Q_{\gamma}^*(\gamma' - b(\gamma)) = \sup_{q \in \mathbb{R}^m} [-H(\gamma, q) + \langle q, \gamma' \rangle]$$

Preuve. On montre (i) en utilisant la définition de $H(\gamma, q)$

et en remarquant que si $\xi - b(\gamma) \notin \text{Im } a$, $Q^*(\xi - b(\gamma))$ vaut $+\infty$ et n'intervient pas dans le sup; et si $\xi - b(\gamma) \in \text{Im } a$, il existe w tel que $\xi - b(\gamma) = aw$.

Alors, $-\frac{1}{2} Q_{\gamma}^*(\xi - b(\gamma)) + \langle q, \xi \rangle = -\frac{1}{2} \langle aw, w \rangle + \langle q, aw + b(\gamma) \rangle$ et cette fonctionnelle atteint son sup pour $w = q$ d'où le résultat.

On montre (ii) avec le même type d'arguments.

3.4. Action. On appellera action de la courbe (γ, q) élément de l'espace $C_t^1(\mathbb{R}^m \times \mathbb{R}^m)$ des courbes C^1 définies sur $[0, t]$ à valeurs dans $\mathbb{R}^m \times \mathbb{R}^m$, la fonctionnelle

$$J_t^b(\gamma, q) = \int_0^t [\langle q_s, \gamma'_s \rangle - H(\gamma_s, q_s)] ds.$$

Les bicaractéristiques sont par définition les extrémales de l'action, c'est-à-dire les solutions des équations d'Euler :

$$\begin{cases} \gamma' = H_q \\ q' = -H_{\gamma} \end{cases}$$

La propriété suivante exprime le lien qui existe entre l'action et la transformée de Cramer et plus précisément entre l'action des bicaractéristiques et la transformée de Cramer.

3.5. PROPOSITION.

Soit (γ, q) un élément de $C_t^1(\mathbb{R}^m \times \mathbb{R}^m)$.

Alors :

$$J_t^b(\gamma, q) \geq \lambda_t^b(\gamma)$$

et si (γ, q) est une bicaractéristique :

$$J_t^b(\gamma, q) = \lambda_t^b(\gamma)$$

Preuve. L'inégalité est évidente avec la formule (ii) du lemme 3.3.

Pour démontrer l'égalité dans le cas des bicaractéristiques, remarquons tout d'abord que si u et v vérifient $v = au$, v et u vecteurs de $\mathbb{R}^m$,

$$(*) \quad Q^*(v) = \langle u, a(\cdot)u \rangle = Q_*(u)$$

Soit (γ, q) une bicaractéristique. Il découle de l'équation d'Euler :

$$\gamma' = H_q \quad ,$$

que

$$\gamma' = a(\gamma)q + b(\gamma).$$

Et donc

$$\begin{aligned} \langle \gamma', q \rangle - H(\gamma, q) &= \langle a(\gamma)q + b(\gamma), q \rangle - \frac{1}{2} Q_\gamma(q) - \langle b(\gamma), q \rangle \\ &= \frac{1}{2} Q_\gamma(q) = \frac{1}{2} Q_\gamma^*(\gamma' - b(\gamma)) \quad \text{d'après } (*). \end{aligned}$$

Par suite

$$\begin{aligned} J_t^b(\gamma, q) &= \int_0^t [\langle \gamma'_s, q_s \rangle - H(\gamma_s, q_s)] \, ds \\ &= \frac{1}{2} \int_0^t Q_{\gamma_s}^* [\gamma'_s - b(\gamma_s)] \, ds = \lambda_t^b(\gamma). \end{aligned}$$

3.6. Interprétation dans le cas des diffusions gaussiennes :

Nous allons maintenant étudier ces fonctionnelles lorsque L est le générateur d'une diffusion gaussienne hypoelliptique. Nous verrons qu'il existe alors une unique bicaractéristique joignant x et y dans le temps t et que l'action le long de cette bicaractéristique est en fait le terme $S_t(x, y)$, exposant de l'exponentielle dans la forme explicite de $p(t, x, y)$ (cf. paragraphe 1). De plus nous montrerons que l'action le long de cette bicaractéristique est égale à la borne inférieure des transformées de Cramer de toutes les courbes de $\mathcal{C}_t(\mathbb{R}^m)$ joignant x à y . Dans ce cas $b(x) = bx + c$; nous utiliserons alors la notation $\lambda_t^{b,c}$ et $J_t^{b,c}$ ou pour simplifier, lorsqu'il n'y aura pas ambiguïté, λ_t^b et J_t^b car c'est la matrice b qui joue un rôle fondamental.

3.7. PROPOSITION : Soit $L(a, b, c)$ le générateur d'une diffusion gaussienne sur $\mathbb{R}^m$ vérifiant la condition d'hypoellipticité ③. Alors pour tout $(x, y) \in \mathbb{R}^m \times \mathbb{R}^m$, il existe une unique bicaractéristique $(\gamma, q)_{x, y} \in C_t^1(\mathbb{R}^m \times \mathbb{R}^m)$ vérifiant $\gamma(0) = x$, $\gamma(t) = y$.

Si on note $J_t^b(x, y)$ l'action de cette bicaractéristique, on a

$$J_t^b(x, y) = S_t^b(x, y)$$

où S_t^b est la fonction définie en 1.9

Preuve. Le Hamiltonien associé à L est $H(\gamma, q) = \frac{1}{2} \langle q, aq \rangle + \langle b\gamma + c, q \rangle$, et les

équations d'Euler s'écrivent

$$\begin{cases} \frac{d\gamma}{ds} = aq + b\gamma + c \\ \frac{dq}{ds} = -{}^tbq. \end{cases}$$

La deuxième équation s'intègre en :

$$q(s) = (\exp - s{}^tb).q_0.$$

Par suite

$$\frac{d\gamma}{ds} = a(\exp - s{}^tb).q_0 + b\gamma + c$$

et alors

$$\gamma(s) = (\exp sb).\gamma_0 + \int_0^s (\exp (s-u)b).a(\exp -u{}^tb).q_0 du + \int_0^s (\exp ub).c du.$$

Il existe alors une unique bicaractéristique partant de x au temps zéro et atteignant y au temps t si $K(t)$ est inversible; en effet : $\gamma(0) = x$ et $\gamma(t) = y$ imposent :

$$\begin{aligned} \gamma_0 &= x \\ \text{et} \quad q_0 &= (\exp t{}^tb) K^{-1}(t) [y - (\exp tb).x - \int_0^t (\exp ub).c du] . \end{aligned}$$

Le long d'une bicaractéristique l'action vaut :

$$\begin{aligned} J_t^b(\gamma, q)_{x,y} &= \frac{1}{2} \int_0^t \langle q_s, a(\gamma_s)q_s \rangle ds \\ &= \frac{1}{2} \int_0^t {}^tq_0(\exp - sb).a(\exp - s{}^tb).q_0 ds \\ &= \frac{1}{2} \varphi(t, x, y) K^{-1}(t) \varphi(t, x, y) = S_t^b(x, y). \end{aligned}$$

3.8. PROPOSITION : Les hypothèses étant celles de la proposition 3.7,

$$\begin{aligned} J_t^b(x, y) &= \inf_{\gamma \in \mathcal{E}_t(\mathbb{R}^m)} \lambda_t^b(\gamma) \\ &\quad \gamma(0) = x \\ &\quad \gamma(t) = y \end{aligned}$$

Preuve. Soient x et $y \in \mathbb{R}^m$ et (γ, q) la bicaractéristique vérifiant $\gamma(0) = x$, $\gamma(t) = y$. (Nous omettrons l'indice (x, y)).

Nous savons (proposition 3.5) que

$$J_t^b(x, y) = J_t^b(\gamma, q) = \lambda_t^b(\gamma).$$

Soit γ un élément de $\mathcal{E}_t(\mathbb{R}^m)$ tel que $\lambda_t^b(\tilde{\gamma})$ soit fini. Nous allons montrer que

$$\lambda_t^b(\tilde{\gamma}) \geq \lambda_t^b(\gamma).$$

Posons $h = \tilde{\gamma} - \gamma$, alors $h(0) = h(t) = 0$. Comme $\lambda_t^b(\gamma)$ et $\lambda_t^b(\tilde{\gamma})$ sont finis,

$\gamma'_s - b\gamma_s - c$ et $\tilde{\gamma}'_s - b\tilde{\gamma}_s - c$ appartiennent p.s à $\text{Im } a$, et par suite il en est

de même de $h'_s - bh_s$.

$$\text{Or } \lambda_t^b(\tilde{\gamma}) - \lambda_t^b(\gamma) = \frac{1}{2} \int_0^t Q^*(\gamma'_s - b\gamma_s - c + h'_s - bh_s) ds - \frac{1}{2} \int_0^t Q^*(\gamma'_s - b\gamma_s - c) ds.$$

De la formule évidente suivante :

$$Q^*(au + bv) = Q^*(au) + 2 \langle u, av \rangle + Q^*(av)$$

et de l'équation d'Euler :

$$\gamma'_s - b\gamma_s - c = aq_s,$$

il résulte que

$$\lambda_t^b(\tilde{\gamma}) - \lambda_t^b(\gamma) = \int_0^t \langle q_s, h'_s - bh_s \rangle ds + \int_0^t Q^*(h'_s - bh_s) ds.$$

Le dernier terme est positif et nous allons montrer que

$$\int_0^t \langle q_s, h'_s - bh_s \rangle ds = 0.$$

En effet, il découle de l'équation d'Euler

$$q'_s = -{}^t b q_s,$$

que

$$\begin{aligned} \int_0^t \langle q_s, h'_s - bh_s \rangle ds &= \int_0^t [\langle q_s, h'_s \rangle + \langle -{}^t b q_s, h_s \rangle] ds = \int_0^t [\langle q_s, h'_s \rangle + \langle q'_s, h_s \rangle] ds \\ &= [\langle q_s, h_s \rangle]_0^t. \end{aligned}$$

Ce dernier terme est nul puisque $h(0) = h(t) = 0$. Et la proposition est démontrée.

3.9. Notation : Nous poserons $E_t^{b,c}(x,y) = \inf_{\substack{\gamma \in \mathcal{E}_t(\mathbb{R}^m) \\ \gamma(0) = x \\ \gamma(t) = y}} \lambda_t^{b,c}(\gamma)$ pour $(x,y) \in \mathbb{R}^m \times \mathbb{R}^m$.

Nous pouvons donc interpréter le $\text{Log } p(t,x,y)$, où $p(t,x,y)$ est la densité de transition d'une diffusion gaussienne, à l'aide de l'action $J_t^{b,c}(x,y)$ ou de $E_t^{b,c}(x,y)$.

Mais nous avons associé au paragraphe 2 (lemme 2.3) une matrice nilpotente d et l'opérateur $L(a,d,0)$ à l'opérateur $L(a,b,c)$ et nous avons remarqué que ce dernier opérateur $L(a,d,0)$ suffisait à déterminer un équivalent de $S_t^{b,c}(x,y)$ et donc de $\text{Log } p(t,x,y)$ en temps petit ; plus précisément nous savons que $S_t^{b,c}(x,y) \sim S_t^{d,0}(x,y)$ pour t proche de 0. Nous obtenons alors immédiatement le corollaire suivant :

3.10 COROLLAIRE. Soit $L(a,b,c)$ le générateur différentiel d'une diffusion gaussienne vérifiant la condition d'hypoellipticité ③ et d la matrice nilpotente associée (2.3). Alors pour t proche de 0, et $(x,y) \in \mathbb{R}^m \times \mathbb{R}^m$,

$$\text{Log } p(t,x,y) \sim J_t^{b,c}(x,y) \sim J_t^{d,0}(x,y)$$

$$\text{et } \text{Log } p(t,x,y) \sim E_t^{b,c}(x,y) \sim E_t^{d,0}(x,y)$$

3.11. Remarquons que la fonctionnelle de Cramer et l'action étudiées ici ne sont pas les mêmes que celles considérées au chapitre 13 où le terme $b(x)$ n'est pas pris en considération. On peut se demander quels liens existent entre ces fonctionnelles.

Considérons le cas des diffusions gaussiennes. Il est clair que dès que $y-x \notin \text{Im } a$, $E_t^0(x,y)$ est infini et ne peut être comparé à $E_t^b(x,y)$. Par contre nous venons de voir qu'il existe une matrice nilpotente d telle que $E_t^b(x,y)$ et $E_t^d(x,y)$ soient équivalents pour t petit. En fait nous avons approché l'opérateur $L(a,b,c)$ par l'opérateur simplifié $L(a,d,0)$ et nous avons remarqué (paragraphe 2) que les logarithmes de leurs densités de transition avaient un comportement équivalent. Le cas où la matrice nilpotente d est nulle est exactement celui où l'opérateur $L(a,b,c)$ est elliptique ($\dim a = m$). Sinon l'opérateur $L(a,0,0)$ ne vérifie même pas la condition d'hypoellipticité (3) et il semble fondamental d'approcher l'opérateur de départ par un opérateur hypoelliptique.

Dans le cas d'un opérateur du second ordre elliptique sur $\mathbb{R}^m$ défini, pour tout x de $\mathbb{R}^m$, par une matrice inversible $a(x)$ et un vecteur $b(x)$, nous savons que, pour x et y proches,

$$E_t^0(x,y) \sim \frac{d^2(x,y)}{t} \quad \text{en temps petit, où } d \text{ est la distance}$$
 riemannienne associée à la matrice $a^{-1}(x)$. Il est facile de voir que, pour des points proches,

$$E_t^b(x,y) \sim E_t^0(x,y) \quad \text{en temps petit, la forme quadratique}$$
 duale Q_x^* étant simplement la forme quadratique définie par $a^{-1}(x)$. Sachant (cf. [8] et chap. 5) que

$$\text{Log } p(t,x,y) \sim E_t^0(x,y),$$

on en déduit que

$$\text{Log } p(t,x,y) \sim E_t^b(x,y).$$

Pour l'exemple du Laplacien de Kohn sur le groupe d'Heisenberg, il a aussi été prouvé (cf. [6] et chap. 13) que

$$\text{Log } p(t,x,y) \sim E_t^0(x,y).$$

Dans cet exemple, il n'existe pas de terme $b(x)$, mais il semblerait naturel de se demander, lorsque ce terme est non nul et lorsque l'opérateur $\sum_{i=1}^r X_i^2 + Y$ vérifie la condition d'hypoellipticité ② ($\mathcal{L}(X_1, \dots, X_r)$ est de dimension m en tout point) si on a encore

$$E_t^b(x,y) \sim E_t^0(x,y).$$

On sait dans ce cas que $\inf_{y \in U} E_t^0(x,y)$ est fini pour tout point x et tout ouvert U de $\mathbb{R}^m$ et on approcherait l'opérateur L par l'opérateur formé des termes du se-

cond ordre, opérateur encore hypoelliptique au sens ②.

Sous la condition d'hypoellipticité ③ il semble tout à fait envisageable de prouver que

$$\text{Log } p(t, x, y) \sim E_t^b(x, y) \sim E_t^d(x, y)$$

où d est une matrice nilpotente obtenue à l'aide de la matrice Jacobienne de $b(x)$ par des méthodes identiques à celles du cas gaussien. En quelque sorte la fonction $E_t^d(x, y)$ serait une généralisation de la fonction $E_t^o(x, y)$ considérée dans le cas elliptique.

Evidemment la fonction $E_t^b(x, y)$ semble être attirante pour conjecturer que $\text{Log } p(t, x, y) \sim E_t^b(x, y)$, mais ne serait-il pas préférable de faire appel aux fonctions simplifiées $E_t^o(x, y)$, $E_t^d(x, y)$ pour ce type de conjectures ?

Nous avons regardé le problème en terme de transformée de Cramer et non en termes d'action, car il existe des exemples (cf. [6] sur le groupe $N_{n,2}$, $n > 2$) pour lesquels $J_t^b(x, y)$ est infini et ou par contre $E_t^b(x, y)$ est fini; et il semble que la notion de transformée de Cramer soit plus adéquate pour notre problème probabiliste, dans la mesure où c'est cette notion qui intervient dans les estimations du type Ventcell et Freidlin de $\int_V \text{Log } p(t, x, y) dy$. [1]

4. Diffusions gaussiennes hypoelliptiques : estimations du type Ventcell et Freidlin et comportement en temps petit :

4.1. Lorsque l'on veut étudier le comportement en temps petit de la densité $p(t, x, y)$ d'une diffusion X_t , une méthode consiste, lorsque l'on a une forme explicite, à calculer directement un équivalent de $p(t, x, y)$ (cf. S. Molchanov [8] pour le cas elliptique et B. Gaveau [6] pour un exemple hypoelliptique du type ②). Une autre méthode envisageable consiste à étudier le $\text{Log } P_x(X_t \in V)$ pour certains voisinages et de faire tendre le volume de V vers zéro. Nous allons dans le cas des diffusions gaussiennes étudier $\text{Log } P_x(X_t \in V)$ par des méthodes du type Ventcell et Freidlin, et nous verrons qu'il est très difficile d'en déduire des résultats sur le $\text{Log } p(t, x, y)$.

Il semble en effet que cette méthode soit utilisable lorsque

$$\text{Log } p(t, x, y) \sim \frac{c(x, y)}{t}$$

où c ne dépend pas de t . C'est le cas elliptique et le cas du Laplacien de Kohn sur le groupe d'Heisenberg. On peut prévoir que c'est aussi le cas des générateurs

$L = \frac{1}{2} \sum_{i=1}^r X_i^2 + Y$ vérifiant l'hypothèse ② c'est-à-dire $\mathcal{L}(X_1, \dots, X_r)(x) = T_x(\mathbb{R}^m)$ pour tout $x \in \mathbb{R}^m$.

Or dans notre cas $\text{Log } p(t, x, y) \sim \frac{c(x, y)}{t^k(x, y)}$ où la puissance k dépend de x et

y et les estimations du type Ventcell et Freidlin vont faire aussi intervenir ces diverses puissances; les passages à la limite sont alors plus délicats, puisque $k(x,y)$ varie selon les points du voisinage V .

4.2. Nous nous proposons donc d'obtenir des évaluations de $\text{Log } P_x(X_t \in V)$ pour certains voisinages et pour cela, comme dans le cours de R. Azencott [1], nous utilisons le lien entre ces questions et les estimations à la Ventcell et Freidlin.

Il suffit d'abord de remarquer qu'étudier la diffusion X_s sur $[0,t]$ lorsque t tend vers zéro, revient à étudier la diffusion $Y_s^\varepsilon = X_{\varepsilon s}$ sur $[0,1]$ lorsque ε tend vers zéro.

De plus, si X_s a pour générateur différentiel L et vérifie l'équation différentielle stochastique :

$$\begin{cases} dX_s = \sigma dB_s + (bX_s + c)ds \\ X_0 = x \end{cases}$$

Y_s^ε a pour générateur différentiel εL et vérifie l'équation différentielle stochastique :

$$S_\varepsilon \quad \begin{cases} dY_s^\varepsilon = \sqrt{\varepsilon} \sigma dB_s + \varepsilon (bY_s^\varepsilon + c)ds \\ Y_0^\varepsilon = x \end{cases}$$

L'extension des estimations de Ventcell et Freidlin au cas hypoelliptique (cf. R. Azencott [1]) donne un encadrement de $\text{Log } P_x(Y^\varepsilon \in A)$ pour un paquet de trajectoires A , où Y^ε désigne la trajectoire de Y_s^ε sur $[0,1]$; mais avant de rappeler ce résultat en 4.5. nous énonçons quelques définitions nécessaires.

4.3. L'espace des trajectoires.

Désignons par $C(\mathbb{R}^m)$ l'espace des fonctions continues de $[0,1]$ dans $\mathbb{R}^m$ muni de la topologie de la convergence uniforme. $C(\mathbb{R}^m)$ est alors l'espace des trajectoires.

Nous allons étudier les projections du processus sur certains sous-espaces N_i (voir le paragraphe 2) et nous aurons besoin des espaces de trajectoires correspondant que nous noterons $C(N_i)$, $0 \leq i \leq p$.

Pour $x = (x_0, \dots, x_i, \dots, x_p)$, $C_x(\mathbb{R}^m)$ sera le sous-espace des fonctions f de $C(\mathbb{R}^m)$ telles que $f(0) = x$; et $C_{x_i}(N_i)$ le sous-espace de $C(N_i)$ tel que ses éléments vérifient $f_i(0) = x_i$.

4.4. Soit $L(a,b,c)$ le générateur différentiel d'une diffusion gaussienne hypoelliptique. On notera Q_0 la forme quadratique associée à a et Q_0^* la forme quadratique duale. L'ensemble des $v \in \mathbb{R}^m$ tels que $Q_0(v)$ soit fini coïncide avec $a(\mathbb{R}^m) = N_0$; on l'appellera le sous-espace horizontal de $\mathbb{R}^m$ (relativement à Q_0). Sur N_0 , Q_0^* définit une norme et donc une distance; on désignera la norme par

$\|v_o\|_o = \langle v_o, a_o^{-1} v_o \rangle = \langle v_o, D_{o,o}^{-1} v_o \rangle$ avec les notations de la proposition 2.4., où $\langle \rangle$ désigne le produit scalaire usuel sur $\mathbb{R}^{n_o}$, et on notera la distance par $d_o(x_o, y_o) = \|y_o - x_o\|_o$.

L'introduction de cette norme prendra tout son sens dans sa généralisation aux différents sous-espaces N_i .

On considère alors la fonctionnelle $\lambda_o : C(\mathbb{R}^m) \rightarrow [0, \infty]$ définie par

$$\lambda_o(g) = \frac{1}{2} \int_0^1 Q_o^*(g'_t) dt \quad \text{si } g'_t \text{ existe au sens de Lebesgue p.s. sur } [0,1]$$

$$\lambda_o(g) = +\infty \quad \text{sinon.}$$

Cette fonctionnelle n'est autre que la transformée de Cramer sur $[0,1]$ de (L, o) .

On définit alors la fonctionnelle de Cramer Λ_o par

$$\Lambda_o(A) = \inf_{g \in A} \lambda_o(g) \quad \text{pour } A \subset C(\mathbb{R}^m),$$

4.5. Soit Y_s^ε la diffusion de générateur $\varepsilon L(a, b, c)$. Alors cf. Azencott [1], on a l'encadrement suivant :

$$-\Lambda_o(\tilde{A}) \leq \lim_{\varepsilon \rightarrow 0} \varepsilon \log P_x(Y^\varepsilon \in A) \leq \overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon \log P_x(Y^\varepsilon \in A) \leq -\Lambda_o(\tilde{A}).$$

où Λ_o est la fonctionnelle de Cramer de S_ε et A une partie de $C_x(\mathbb{R}^m)$.

Ce résultat donne pour des ensembles assez réguliers A un équivalent du $\log P_x(Y^\varepsilon \in A)$ en $-\frac{cte}{\varepsilon}$ si $\Lambda_o(A) < \infty$, c'est-à-dire si le paquet de trajectoires contient des courbes horizontales; mais si $\Lambda_o(A) = \infty$, le théorème n'est pas assez précis, le problème étant alors de trouver la "bonne" puissance k telle que

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^k \log P_x(Y^\varepsilon \in A) \text{ soit finie.}$$

Nous allons donner des exemples d'ensembles A pour lesquels nous trouvons les bons exposants mais avant, nous allons introduire des nouvelles fonctionnelles, attachées à chaque sous-espace N_i et qui, comme Λ_o pour N_o caractérisent en un certain sens le comportement de la projection des trajectoires sur N_i .

4.6. DÉFINITION. Pour chaque $i \in \{0, \dots, p\}$ (pour $i = 0$ on retrouvera la définition du sous-paragraphe 4.4), on appelle Q_i la forme quadratique associée à la matrice carrée positive $(m-n_{i-1}, m-n_{i-1})$ $\overline{D}_{ii}$ obtenue comme suit :

$$\overline{D}_{ii} = \begin{bmatrix} & n_i - n_{i-1} & m - n_{i-1} \\ & D_{i,i} & 0 \\ n_i - n_{i-1} \dots & 0 & 0 \end{bmatrix}$$

avec $D_{i,i} = b_i \dots b_1 a_o t_{b_1} \dots t_{b_i}$ (et $D_{o,o} = a_o$) et soit Q_i^* sa forme quadratique duale à valeurs dans $[0, +\infty]$ définie par la formule habituelle :

$$\frac{1}{2} Q_i^*(w) = \sup_{t \in \mathbb{R}} \sup_{m-n_i-1} [\langle t, w \rangle - \frac{1}{2} Q_i(t)] \quad , \quad w \in \mathbb{R}^{m-n_i-1}$$

(on notera également $\langle, \rangle$ tous les produits scalaires usuels dans les différents $\mathbb{R}^k$, la signification étant claire pour chacun).

De même que pour Q_0 , l'ensemble des vecteurs w de $\mathbb{R}^{m-n_i-1}$ tels que $Q_i^*(w)$ soit fini coïncide avec N_i , ce sera le sous-espace horizontal associé à Q_i .

Sur N_i , Q_i^* définit une norme et une distance, on désignera par $\|w_i\|_i$ la norme et $d_i(x_i, y_i) = \|y_i - x_i\|_i$ la distance.

On a donc : $\|w_i\|_i = \langle w_i, D_{ii}^{-1} w_i \rangle$, le produit scalaire usuel étant sur $\mathbb{R}^{n_i - n_{i-1}}$.

Soit alors la fonctionnelle $\lambda_i : C(\mathbb{R}^{m-n_i-1}) \rightarrow [0, \infty]$ définie par :

$$\lambda_i(g) = \frac{1}{2} \int_0^1 Q_i^*(g_t^{(i+1)}) dt \quad \text{si } g_t^{(i+1)} \text{ existe au sens de Lebesgue p.s. sur } [0, 1] .$$

$$\lambda_i(g) = +\infty \quad \text{sinon;}$$

et on posera $\Lambda_i(A) = \inf_{g \in A} \lambda_i(g)$ pour toute partie de $C(\mathbb{R}^{m-n_i-1})$.

Donnons maintenant le résultat pour les paquets de trajectoires où l'on n'impose rien aux coordonnées avant N_i et où l'on assujettit les suivantes à rester dans un ensemble fixe.

4.7. THÉOREME. Soit $A = C_{x_0}(\mathbb{R}^{n_i-1}) \times A_i$ avec A_i borélien de $C_{x_i, \dots, x_p}(\mathbb{R}^{m-n_i})$, on a alors avec les notations ci-dessus, et $x = (x_0, x_1, \dots, x_i, \dots, x_p)$:

$$-\Lambda_i(\dot{A}_i) \leq \lim_{\varepsilon \rightarrow 0} \varepsilon^{2i+1} \text{Log } P_x(Y^\varepsilon \in A) \leq \overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon^{2i+1} \text{Log } P_x(Y^\varepsilon \in A) \leq -\Lambda_i(\bar{A}_i)$$

Nous allons démontrer ce théorème dans le cas $i = 1$, et pour ce faire, si nous notons Y_1^ε les coordonnées à partir du rang n_0+1 , il suffit d'étudier Y_1^ε car :

$$P_x[Y^\varepsilon \in A] = P_{(x_1, \dots, x_p)}[Y_1^\varepsilon \in A_1] .$$

La démonstration se fait alors en deux temps comme dans le cours de R. Azencott [1].

D'abord on prouve la continuité de l'application mesurable qui à $\varepsilon \sqrt{\varepsilon} B$ associe Y_1^ε , où B est la trajectoire du mouvement brownien sur $[0, 1]$. Pour cela on montre que si $\varepsilon \sqrt{\varepsilon} B$ est proche d'une fonction f donnée, Y_1^ε est proche de g_1 où g_1 est la solution du système déterministe suivant :

$$\begin{cases} g_1'' = \overline{b_1 \sigma_0} f' \\ g_1'(0) = 0 \\ g_1(0) = (x_1, \dots, x_p) \end{cases}$$

$$\text{avec } \overline{b_1 \sigma_o} = \begin{pmatrix} b_1 \sigma_o \\ \hline 0 \end{pmatrix}^{n_1 - n_o}_{m - n_o}$$

(on a prolongé $b_1 \sigma_o$ qui est une matrice $(n_1 - n_o, n_o)$ par une matrice de zéro de dimension $(m - n_1, n_o)$).

On notera également $(x_1, \dots, x_p) = \bar{x}_1$ pour plus de commodité.

4.8. LEMME Pour tout η positif, il existe ϵ_o et α tels que si $\epsilon \leq \epsilon_o$,

$$d(\epsilon \sqrt{\epsilon} B, f) \leq \alpha \Rightarrow d(Y_1^{\epsilon}, g_1) \leq \eta$$

où d désigne à la fois les deux distances uniformes sur $[0, 1]$, la première pour les fonctions à valeurs dans $\mathbb{R}^{n_o}$, la deuxième pour les fonctions à valeurs dans $\mathbb{R}^{m - n_o}$.

Preuve. Notons par un indice supérieur les coordonnées de toutes les fonctions qui interviennent ainsi que celles du processus.

Remarquons tout d'abord qu'il suffit d'évaluer $\|Y_t^{\epsilon, i} - g_t^{\epsilon, i}\|$ pour $i > n_o$, $t \in [0, 1]$ où g^{ϵ} est la solution du système déterministe :

$$\begin{cases} g'^{\epsilon} = \frac{1}{\epsilon} \sigma f' + \epsilon (b g^{\epsilon} + c) \\ g^{\epsilon}(0) = x. \end{cases}$$

En effet montrons que $\|g^{\epsilon, i} - g_1^i\|$ tend vers zéro lorsque ϵ tend vers zéro pour tous les $i > n_o$.

On a pour $i > n_o$:

$$g''^{\epsilon, i} = (b \sigma f')^i + (\epsilon^2 b^2 g^{\epsilon} + \epsilon^2 b c + \epsilon c)^i.$$

En remarquant que la matrice $b \sigma$ s'écrit :

$$\begin{bmatrix} * & 0 \\ \hline b_1 \sigma_o & \\ \hline 0 & 0 \end{bmatrix}$$

On obtient :

$$g''^{\epsilon, i} - g_1''^i = (\epsilon^2 b^2 g^{\epsilon} + \epsilon^2 b c + \epsilon c)^i.$$

En appliquant une première fois le lemme de Gronvall à g^{ϵ} , on voit que ϵg^{ϵ} reste borné; puis la formule simple de calcul différentiel (J. Dieudonné [3]) :

$$\|g^{\epsilon, i} - g_1^i\| \leq \|g^{\epsilon, i} - g_1^i\| + \frac{1}{2} \|g''^{\epsilon, i} - g_1''^i\|$$

suivie d'une nouvelle application du lemme de Gronvall montre alors que $\|g^{\epsilon, i} - g_1^i\|$ tend vers zéro pour tous les $i > n_o$.

Évaluons maintenant $\|Y_t^{\varepsilon,i} - g_t^{\varepsilon,i}\|$; en appliquant la formule de Ito à la fonction $h(t,x) = t(bx+c)$ pour chaque composante $i > n_0$, on trouve :

$$Y_t^{\varepsilon,i} = \varepsilon t(bY_t^{\varepsilon} + c)^i - \int_0^t s(b\sigma \varepsilon \sqrt{\varepsilon} dB_s)^i - \varepsilon^2 \sum_k \int_0^t s b_{ik} (bY_s^{\varepsilon} + c)^k ds + x^i$$

et en transformant $g_t^{\varepsilon,i}$, $i > n_0$, de la même manière, il vient :

$$g_t^{\varepsilon,i} = \varepsilon t(bg_t^{\varepsilon} + c)^i - \int_0^t s(b\sigma f')^i ds - \varepsilon^2 \sum_k \int_0^t s b_{ik} (bg_s^{\varepsilon} + c)^k ds + x^i.$$

D'où :

$$\begin{aligned} Y_t^{\varepsilon,i} - g_t^{\varepsilon,i} &= \varepsilon t [b(Y_t^{\varepsilon} - g_t^{\varepsilon})]^i - \int_0^t s [b\sigma(\varepsilon \sqrt{\varepsilon} dB_s - f'(s)ds)]^i \\ &\quad - \varepsilon^2 \sum_k \int_0^t s b_{ik} [b(Y_s^{\varepsilon} - g_s^{\varepsilon})]^k ds. \end{aligned}$$

Majoration du premier terme :

$$\|\varepsilon t(b(Y_t^{\varepsilon} - g_t^{\varepsilon}))^i\| \leq \varepsilon M \|Y^{\varepsilon} - g^{\varepsilon}\|,$$

mais :

$$\begin{aligned} \varepsilon(Y_t^{\varepsilon} - g_t^{\varepsilon}) &= \int_0^t (\varepsilon \sqrt{\varepsilon} dB_s - \sigma f'(s)ds) + \varepsilon^2 \int_0^t b(Y_s^{\varepsilon} - g_s^{\varepsilon}) ds \\ &= t\sigma(\varepsilon \sqrt{\varepsilon} B_t - f(t)) + \varepsilon^2 \int_0^t b(Y_s^{\varepsilon} - g_s^{\varepsilon}) ds. \end{aligned}$$

Donc, si $d(\varepsilon \sqrt{\varepsilon} B, f) \leq \alpha$,

$$\|\varepsilon(Y_t^{\varepsilon} - g_t^{\varepsilon})\| \leq c\alpha + \varepsilon \int_0^t M \|\varepsilon(Y_s^{\varepsilon} - g_s^{\varepsilon})\| ds$$

et on conclut grâce au lemme de Gronvall que $\varepsilon(Y_t^{\varepsilon} - g_t^{\varepsilon})$ peut être rendu petit lorsque α est petit, pourvu que ε soit borné.

Deuxième terme :

$$\int_0^t s(b\sigma(\varepsilon \sqrt{\varepsilon} dB_s - f'(s)))^i ds = t(b\sigma(\varepsilon \sqrt{\varepsilon} B_t - f(t)))^i - \int_0^t (b\sigma(\varepsilon \sqrt{\varepsilon} B_s - f(s)))^i ds.$$

Donc si $d(\varepsilon \sqrt{\varepsilon} B, f) \leq \alpha$, le deuxième terme est majoré en norme par $c\alpha$.

. quant au troisième terme, on le majore facilement par $\varepsilon c\alpha$ grâce aux remarques faites pour les autres termes.

4.9. Fin de la démonstration du théorème 4.7.

Une fois le lemme démontré, c'est-à-dire la "continuité" au sens du lemme de l'application $\varepsilon \sqrt{\varepsilon} B \rightarrow Y_1^{\varepsilon}$ prouvée, on conclut très rapidement comme l'a remarqué H. Doss [4].

En effet, $P[Y_1^{\varepsilon} \in A_1] = P[\varepsilon \sqrt{\varepsilon} B \in \phi^{-1}(A_1)]$

où ϕ est l'application continue qui à f fait correspondre g_1 telle que :

$$\left\{ \begin{array}{l} g_1'' = \overline{b_1} \sigma_0 f' \\ g_1'(0) = 0 \\ g_1(0) = \overline{x_1} \end{array} \right.$$

On utilise alors les résultats sur la transformée de Cramer du mouvement brownien pour obtenir :

$$- \inf_{\phi^{-1}(A_1)} \tilde{\lambda}(f) \leq \lim_{\varepsilon \rightarrow 0} \varepsilon^3 \text{Log } P[Y_1^\varepsilon \in A_1] \leq \overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon^3 \text{Log } P[Y_1^\varepsilon \in A_1] \leq - \inf_{f \in \phi^{-1}(A_1)} \tilde{\lambda}(f)$$

avec $\tilde{\lambda}(f) = \int_0^1 \|f'\|^2 ds$, transformée de Cramer du mouvement brownien; on remarque enfin que $\tilde{\lambda}(f) = \lambda_1(g_1)$ si $g_1'' = \overline{b_1} \sigma_0 f'$, ce qui achève la démonstration du théorème.

Le théorème permet alors d'énoncer un corollaire pour les diffusions gaussiennes hypoelliptiques en temps petit.

4.10 COROLLAIRE . Soit X_t la diffusion gaussienne hypoelliptique de générateur différentiel $L(a,b,c)$ partant de $x = (x_0, x_1, \dots, x_i, \dots, x_p)$ et soit un autre point $y = (y_0, \dots, y_i, \dots, y_p)$; soit enfin un voisinage de y de la forme $N_0 \times \dots \times N_{i-1} \times V_i \times W$ où V_i est inclus dans $\mathbb{R}^{n_i - n_{i-1}}$ et W dans $\mathbb{R}^{m - n_i}$ est un voisinage de $\overline{x_{i+1}} = (x_{i+1}, \dots, x_p)$. Alors,

$$- \frac{1}{2}(i!)^2 (2i+1) d_i^2(x_i, \overline{V_i}) \leq \lim_{t \rightarrow 0} t^{2i+1} \text{Log } P_x(X_t \in V) \leq \overline{\lim}_{t \rightarrow 0} t^{2i+1} \text{Log } P_x(X_t \in V) \leq - \frac{1}{2}(i!)^2 (2i+1) d_i^2(x_i, \overline{V_i})$$

Remarquons que si V_i contient x_i , tous les termes valent zéro et si V_i ne contient pas x_i , cette expression donne un équivalent de $\text{Log } P_x(X_t \in V)$ pour les ouverts réguliers du type ci-dessus en $-\frac{\text{cte}}{2i+1}$.

Preuve. Nous allons donner la démonstration dans le cas $i = 1$, c'est-à-dire $V = N_0 \times V_1 \times W$.

Montrons tout d'abord la formule suivante qui est une conséquence immédiate du théorème 4.7 :

$$- \Lambda_1(\bar{U}) \leq \lim_{t \rightarrow 0} t^3 \text{Log } P_x(X_t \in V) \leq \overline{\lim}_{t \rightarrow 0} t^3 \text{Log } P_x(X_t \in V) \leq - \Lambda_1(\bar{U})$$

si $U = \{g_1 \in C_{\overline{x_1}}^{m-n_0}(\mathbb{R}^{m-n_0}) : g_1(0) = \overline{x_1}, g_1'(0) = 0, g_1(1) \in V_1 \times W\} = U_{V_1 \times W}$

En effet, les trois événements $\{X_t \in V\}$, $\{X_{1,t} \in V_1 \times W\}$, $\{Y_1^\varepsilon \in U_{V_1 \times W}\}$ ont même probabilité, avec $X_{s\varepsilon} = Y_s^\varepsilon$ et où $X_{1,t}$ désigne les coordonnées de X_s après le rang n_0 .

De plus, $\bar{U}_{V_1 \times W} \subset U_{V_1 \times W}$ et $U_{\frac{o}{V_1 \times W}} \subset \bar{U}_{V_1 \times W}$ donc les deux limites $\lim$ et $\lim$ de $t^3 \text{Log } P(X_t \in V)$ sont dans $[-\lambda_1(\bar{U}), -\lambda_1(\bar{U})]$.

Pour terminer la démonstration, il faut montrer que : $\inf_A \lambda_1(g_1) = \frac{3}{2} d_1^2(x_1, y_1)$ où A est l'ensemble suivant :

$$A = \begin{cases} g_1(0) = \bar{x}_1 \\ g_1(1) = \bar{y}_1 \\ g'_1(0) = 0 \end{cases}$$

Soit à minimiser la fonctionnelle $\int_0^1 Q_1^*(g_1'') ds$ sur A .

Appelons $h_1 = (g_1^{n_0+1}, \dots, g_1^{n_1})$ les coordonnées de g_1 qui sont dans N_1 et montrons que il est équivalent de calculer : $\inf_{A_1} \int_0^1 \|h_1''\|_1^2 ds$ où A_1 désigne l'ensemble :

$$A_1 = \begin{cases} h_1(0) = x_1 \\ h_1(1) = y_1 \\ h'_1(0) = 0 \end{cases}$$

En effet, on remarque tout d'abord que d'après la définition de Q_1^* qui vaut $+\infty$ pour les éléments qui ne sont pas dans le sous-espace horizontal correspondant à Q_1 , il suffit de minimiser $\int_0^1 \|h_1''\|_1^2 ds$ en imposant aux coordonnées après n_1 de g_1 de rester au point de départ ce qui est possible puisque la seule condition exigée concernant les coordonnées après n_1 des éléments de A est de rester dans un voisinage W du point de départ.

L'équation d'Euler s'écrit alors :

$$\frac{d^2}{ds^2}(F_{x''}) = 0 \text{ avec } F(s, x, x', x'') = \|x''\|_1^2$$

la différentielle appliquée à u vaut

$$2 \langle u, h_1^{(4)} \rangle_1 = 2 \langle u, D_{1,1}^{-1} h_1^{(4)} \rangle$$

elle s'annule si $h_1^{(4)} = 0$, soit :

$$h_1 = \frac{k}{6} s^3 + \frac{\lambda}{2} s^2 + \mu s + \nu,$$

les conditions initiales :

$$A_z = \begin{cases} h_1(0) = x_1 \\ h_1(1) = y_1 \\ h'_1(0) = 0 \\ h'_1(1) = z \end{cases}$$

entraînent : $v = x_1$, $\mu = 0$, $k = 6z - 12(y_1 - x_1)$, $\lambda = 6(y_1 - x_1) - 2z$.

D'où

$$\inf_{A_z} \int_0^1 \|h''\|_1^2 ds = \frac{1}{3} \|6z - 12y_1\|_1^2 + \langle 6z - 12y_1, 6y_1 - 2z \rangle_1 + \|6y_1 - 2z\|_1^2 .$$

Pour obtenir le résultat cherché, on prend l'inf sur les z en calculant d'abord la dérivée par rapport à z qui vaut :

$$u \rightarrow 2 \langle u, 6z - 12y_1 \rangle_1 + 2 \langle u, 6y_1 - 2z \rangle_1$$

l'inf est alors atteint pour $z = \frac{3}{2} (y_1 - x_1)$; soit en reportant dans l'expression de la fonctionnelle,

$$\inf_{A_z} \int_0^1 \|h''\|_1^2 ds = 3 \|y_1 - x_1\|_1^2 = 3d_1^2(x_1, y_1) .$$

Remarque : Considérons la matrice $K_{i,i}(t)$ introduite en 2.4 et qui n'est rien d'autre que la matrice de covariance de la projection de la diffusion sur N_i .

De l'équivalence de $K_{i,i}(t)$ à $\frac{1}{(i!)^2} \frac{t^{2i+1}}{2i+1} D_{ii}$, il résulte que

$$\langle y_i - x_i, K_{i,i}^{-1}(t)(y_i - x_i) \rangle \sim (i!)^2 \frac{(2i+1)}{t^{2i+1}} d_i^2(x_i, y_i) ,$$

et par suite $\log P_x[X_t \in V] \sim -\frac{1}{2} \langle y_i - x_i, K_{i,i}^{-1}(t)(y_i - x_i) \rangle$ ce qui est naturel car intuitivement dans ces évaluations seule la projection du processus sur N_i intervient.

BIBLIOGRAPHIE

- [1] R. Azencott, *Grandes déviations et applications* - Ecole d'Eté de Probabilités de Saint-Flour 1978, Springer 774 .
- [2] M. Chaleyat-Maurel et L. Elie, à paraître.
- [3] J. Dieudonné, *Fondements de l'analyse moderne*, Gauthier-Villars Paris 1968.
- [4] H. Doss, *Quelques formules asymptotiques pour les petites perturbations de systèmes dynamiques*, à paraître.
- [5] L. Elie, *Une paramétrix pour l'étude de la densité de transition de certaines diffusions à générateur hypoelliptique*, à paraître.
- [6] B. Gaveau, *Principe de moindre action. Propagation de la chaleur. Estimées sous elliptiques sur certains groupes nilpotents*. Acta Math. 139 (1977) p. 96-153.
- [7] D. Manankiandrianana, *Noyau de la chaleur d'un opérateur hypoelliptique dégénéré (comportement pour des temps petits)*. C.R. Acad. Sci. Paris t 288 n° 23 (25 juin 1979).
- [8] S. Molchanov, *Diffusions et géométrie riemannienne*. Uspetchi. Mat. Nayk. 30 (1975) p. 3-59.
- [9] E. Nelson, *Dynamical theories of brownian motion*. Mathematical notes Princeton University press (1967).

ABSTRACT

Geodesics and diffusions in small time intervals

On a differentiable manifold M , consider a second order hypoelliptic differential operator L , such that $L1 = 0$. Call $p(t, x, y)$, $t > 0$, $x, y \in M$, the minimal fundamental solution of $(\frac{d}{dt} - L)$ associated to an arbitrary riemannian volume v on M . If $\mathcal{P}$ is the unique diffusion on M with infinitesimal generator L , then p is the density (with respect to v) of the transition semi-group of $\mathcal{P}$.

This volume, which springs from a two year seminar held at the Université Paris VII, is concerned with the behaviour of $p(t, x, y)$ as $t \rightarrow 0$ and its geometric significance in terms of geodesics, or more generally, minimal action curves from x to y .

The aim was to clarify and improve the probabilistic approach of these questions, as presented in the work of VARADHAN, MOLCHANOV, GAVEAU. The link with path-space integrals and estimates à la VENTCELL-FREIDLIN has been emphasized.

The text, to be as selfcontained as possible, includes summary but systematic presentations of the main technical tools borrowed from diffusion theory, riemannian geometry, hypoelliptic operators, slightly perturbed dynamic systems.