# GRANDES DEVIATIONS ET APPLICATIONS

PAR R. AZENCOTT

# CHAPITRE ZERO

## 1. DECLARATION D'INTENTIONS

Ce cours présente les principaux résultats probabilistes sur les grandes déviations des sommes de variables aléatoires indépendantes (à valeurs dans un espace vectoriel topologique) et les points essentiels de la théorie des petites perturbations (aléatoires) des systèmes dynamiques.

Nous nous sommes efforcés d'unifier la présentation de ces deux courants d'idées, et de prouver les résultats de base sur les petites perturbations, à partir des théorèmes de grandes déviations sur les gaussiennes. Ce point de vue s'est révélé fécond, et nous a permis d'obtenir de sérieuses améliorations techniques comme l'extension des théorèmes de petites perturbations au cadre des opérateurs différentiels hypoelliptiques (du second ordre) sur des variétés non compactes.

Nous avons ébauché aussi les liens de la théorie des petites perturbations avec celle du comportement des diffusions en temps petit, liens qui méritent d'être exploités plus profondément à notre avis, particulièrement en conjonction avec le point de vue adopté ici (voir la conclusion de ce cours).

Nous avons enfin esquissé les applications au contrôle des équilibres perturbés.

L'espace restreint du cours d'été a forcé l'omission de plusieurs thèmes intéressants ; nous renvoyons au paragraphe 3 pour un bilan de ces manques.

## 2. INFLUENCES, CONNIVENCES ET MONDANITES

Commençons par remercier A. Badrikian et P.L. Hennequin pour la chaleur de leur accueil, qui a failli compenser le côté plutôt monastique de Saint-Flour-by-night. Nous gardons aussi un très bon souvenir des échanges avec l'ensemble des participants à l'école d'été.

Sur le plan mathématique, quelques conversations avec A. Badrikian nous ont aidé à clarifier de désagréables questions de mesurabilité concernant le chapitre I. Les séances du séminaire de statistiques d'Orsay (1977-78) [47] ont été pour nous un utile stimulant pendant la préparation de ce cours. Le séminaire de probabilité de Paris VII (1978-79) [46] a joué le même rôle pendant la rédaction de ce travail. Enfin, tout au long de notre réflexion sur ces thèmes, nous avons souvent bénéficié à retardement de l'impact d'une collaboration antérieure avec G. Ruget [3] sur des thèmes voisins.

Remercions enfin Madame Courageot et Madame Fontaine qui ont assuré avec beaucoup d'efficacité la dactylographie de ce travail, permettant ainsi de rattraper un retard considérable dû au rédacteur.

## 3. SURVOL DU TEXTE

Soit $(X_n)$ une suite de v.a. (à valeurs dans un espace vectoriel topologique E) indépendantes et de même loi $\mu$. Posons $m = \int_E x \, d\mu(x)$ et $\overline{X}_n = \frac{1}{n}(X_1 + \ldots + X_n)$. Les théorèmes "à la Cramer-Chernoff" étudient le comportement asymptotique de $\frac{1}{n} \log P(\overline{X}_n \in A)$ pour $n \to +\infty$, avec $A \subset E$. Comme "en général" $\overline{X}_n \to m$ presque sûrement, les évènements $\{\overline{X}_n \in A\}$ ont des probabilités tendant vers 0 lorsque $n \to +\infty$ dès que $m \notin \overline{A}$, et sont considérés alors comme réalisant de "grandes déviations" de $\overline{X}_n$ par rapport à m. Dans ce contexte terminologique, les "petites déviations" correspondraient à des évènements du type $\{\overline{X}_n \in A_n\}$ avec $A_n = m + \frac{1}{\sqrt{n}} B$, $B \subset E$, évènements dont les probabilités tendent "en général" vers des limites non nulles décrites par le théorème limite central.

## Chapitre I : Théorème à la Cramer-Chernoff

Nous étendons les résultats de Cramer [14] Chernoff [10] (qui traitent $E = \mathbb{R}$, $A$ = demi-droite) en dimension infinie : nous présentons ainsi les importants résultats de Donsker-Varadhan [15] sur le cas des Banach, en nous inspirant d'une formalisation ultérieure de Bahadur-Zabell [7].

Introduisons la transformée de Laplace $\hat{\mu} : E' \to [0, +\infty]$ définie sur le dual $E'$ de $E$ par

$$(1) \qquad \hat{\mu}(t) = \int_E e^{<t,x>} d\mu(x) \qquad\qquad t \in E'$$

Définissons la transformée de Cramer $\lambda : E \to [0, +\infty]$ de la loi $\mu$ par

$$(2) \qquad \lambda(x) = \sup_{t \in E'} [< t, x > - \log \hat{\mu}(t)] \quad , \qquad x \in E$$

et la fonctionnelle de Cramer $\Lambda$ : {parties de $E$} $\longrightarrow [0, +\infty]$ par

$$(3) \qquad \Lambda(A) = \inf_{x \in A} \lambda(x) \qquad\qquad A \subset E$$

Le résultat de base [§ 7] fournit l'encadrement, dû à Donsker-Varadhan [15]

$$(4) \quad - \Lambda(\mathring{A}) \leqslant \underline{\lim}_{n\to\infty} \frac{1}{n} \log P(\overline{X}_n \in A) \leqslant \overline{\lim}_{n\to\infty} \frac{1}{n} \log P(\overline{X}_n \in A) \leqslant - \Lambda(\overline{A})$$

pour tout borélien $A$ de $E$. La minoration est pratiquement vraie ; la majoration utilise, dans le cas où $E$ est un espace de Banach, l'hypothèse $\{ \int e^{s||x||} d\mu(x)$ fini pour tout $s \in \mathbb{R}\}$. Nous en déduisons (§ 8) le théorème de Sanov [43], lequel affirme "pratiquement" que, si $Z_n$ est la loi empirique d'un n-échantillon de v.a. de même loi $\pi$, indépendantes, à valeurs dans un espace topologique polonais $\Gamma$, "on a"

$$\lim_{n\to+\infty} \frac{1}{n} \log P(Z_n \in A) = - \inf_{\nu \in A} I_\pi(\nu)$$

pour tout ensemble borélien A de probabilités sur $\Gamma$, où

$I_\pi(\nu) = \int_\Gamma \frac{d\nu}{d\pi}(x) \log \frac{d\nu}{d\pi}(x)\, d\pi(x)$ est l'information de Kullback de $\nu$ par rapport à $\pi$.

Le § 9 décrit d'un point de vue botanique l'allure des transformées de Cramer $\lambda$ en dimension finie et précise les liens de la correspondance $\{\log \hat{\mu} \to \lambda\}$ avec la dualité de Legendre et la dualité des fonctions convexes.

## <u>Chapitre II</u> : <u>Application aux mesures gaussiennes et processus gaussiens</u>

Nous calculons les transformées de Cramer des lois gaussiennes en dimension infinie ; nous traitons en particulier le cas où $\mu$ est la loi de la trajectoire d'un processus gaussien à trajectoires continues. Le résultat (4) fournit l'encadrement

$$(5) \quad -\Lambda(\mathring{A}) \leqslant \underline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P(\varepsilon X \in A) \leqslant \overline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P(\varepsilon X \in A) \leqslant -\Lambda(\bar{A})$$

où X est une v.a. gaussienne centrée à valeurs dans un Banach E, A un borélien de E, $\Lambda$ la fonctionnelle de Cramer de la loi $\mu$ de X. De (5) on déduit (Donsker-Varadhan [15]) le calcul des limites (quand $r \to +\infty$) de $\frac{1}{r^2} \log \mu \{x \in E \mid \|x\| \geqslant r\}$ pour une loi gaussienne $\mu$, et de $\frac{1}{r^2} \log P \{\sup_{s\in[0,1]} |X_s| \geqslant r\}$ pour un processus gaussien continu $X_s$.

Enfin nous calculons la transformée de Cramer de la loi des trajectoires du brownien sur l'intervalle de temps $[0, 1]$ . La transformée de Cramer $\lambda$ est définie sur l'espace E des fonctions continues sur $[0, 1]$ nulles en 0, et vaut

(6) $\lambda(f) = \frac{1}{2} \int_0^1 f'(t)^2\, dt$ si f est absolument continue

$\lambda(f) = +\infty$ dans le cas contraire.

Chapitre III : Petites perturbations de systèmes dynamiques

Nous considérons (Ventsel [52] Freidlin [23]) le système dynamique $y'_t = b\ (y_t)$ où b est un champ de vecteurs sur $\mathbb{R}^n$ et le système perturbé

$$\frac{d\ y_t^\varepsilon}{dt} = b\ (y_t^\varepsilon) + \varepsilon\ \sigma\ (y_t^\varepsilon)\ X_t \qquad \varepsilon > 0$$

où $(X_t)$ est un processus gaussien continu de loi trajectorielle $\mu$ sur $C_{0,1} = \{$espace des fonctions continues sur $[0,\ 1]$ à valeurs dans $\mathbb{R}^n\}$.
Nous évaluons $\lim\limits_{\varepsilon \to 0} \varepsilon^2 \log P\ (y^\varepsilon \in A)$, où $A \subset C_{0,1}$, en prenant "l'image" de la transformée de Cramer de $\mu$ par une application déterministe.

Nous abordons ensuite le cas plus important, et plus délicat, du modèle de perturbations

$$(7) \qquad d\ y_t^\varepsilon = b_\varepsilon\ (y_t^\varepsilon) + \varepsilon\ \ \sigma\ (y_t^\varepsilon)\ \ d\ \beta_t \qquad \varepsilon > 0$$

défini sur l'ouvert U de $\mathbb{R}^n$, où $\beta$ est un brownien k-dimensionnel, $\sigma\ (x)$ est un champ $C^1$ de matrices (n, k), et les $b_\varepsilon\ (x)$, $\varepsilon > 0$ sont des champs de vecteurs tels que $\lim\limits_{\varepsilon \to 0} b_\varepsilon = b$ uniformément sur tout compact.

Dans un travail fondamental [53], Ventsel et Freidlin considéraient un modèle analogue sur une variété compacte M, en s'intéressant à la diffusion $(y_t^\varepsilon)$ de générateur infinitésimal $b + \varepsilon^2\ \Delta$, où b est un champ de vecteurs et $\Delta$ un opérateur différentiel elliptique (strict) d'ordre 2.
L'introduction de $b_\varepsilon$ au lieu de b dans (7) n'est pas gratuite : elle permet une écriture locale invariante par changements de coordonnées.

Notre méthode est différente de celle de Ventsel et Freidlin : la solution $y^\varepsilon$ de (7) peut s'écrire $y^\varepsilon = F\ (\varepsilon\beta)$ où F est une fonctionnelle qui en général est seulement mesurable. Pour x fixé nous introduisons l'application G définie sur une partie dense $C^\circ_{0,1}$ de $C_{0,1}$ par (détails omis ...)

$$(8) \quad G\ f = g \iff \begin{cases} g'_t = b\ (g_t) + \ \varepsilon\ \sigma\ \ (g_t)\ f'_t & \text{p.p.} \quad t \in [0,\ 1] \\ g_o = x \end{cases}$$

Nous montrons que G est une bonne approximation de F au sens suivant :

(9) La probabilité conditionnelle

$$P\left[F\ (\varepsilon\beta)\ \text{"proche" de } Gf \mid \varepsilon\beta \ \text{"proche" de } f\right]$$

tend vers 1 à vitesse exponentielle $\exp\left(-\frac{C}{\varepsilon^2}\right)$ quand $\varepsilon \to 0$, où C peut être pris arbitrairement grand.

On introduit alors la transformée de Cramer de (7) par

(10) $$\lambda\ (g) = \inf_{Gf=g} \tilde{\lambda}\ (f)$$

où $\tilde{\lambda}$ est la transformée de Cramer (cf. (6)) du mouvement brownien k-dimensionnel. Comme nous traitons le cas général où les trajectoires de $y^\varepsilon$ peuvent atteindre l'infini $\delta$ de U en temps fini, nous considérons le temps d'explosion $\tau$ (g) d'un chemin g (i.e. le temps d'atteinte de l'infini) et l'espace topologique $\mathcal{E}_x$ (U) des trajectoires "explosives" issues de x , à valeurs dans $U \cup \delta$ , continues sur $[0,\ 1]$ qui restent en $\delta$ sur $[\tau\ (g),\ 1]$. La transformée de Cramer $\lambda$ définie par (10) peut alors s'écrire

(11) $$\lambda\ (g) = \int_0^{1\wedge\tau(g)} Q^*_{g_t}\ \left[g'_t - b\ (g_t)\right]\ dt$$

où $Q^*_x : \mathbb{R}^n = T_x\ (U) \longrightarrow [0,\ +\ \infty]$ est la "forme quadratique duale" de $Q_x = \sigma\ (x)\ \sigma^*\ (x)$, définie par

(12) $$\frac{1}{2}\ Q^*_x\ (v) = \sup_{t\in\mathbb{R}^n} \left[< t,\ v > - \frac{1}{2}\ Q_x(t)\right] = \inf_{\sigma(x)w=v} ||w||^2$$

Nous prouvons alors (théorème 2.13) que pour $A \subset \mathcal{E}_x$ (U), les limites inf. et sup. de $\varepsilon^2 \log P_x\ (y^\varepsilon \in A)$ quand $\varepsilon \to 0$, sont dans l'intervalle $[-\Lambda(\mathring{A}),\ -\Lambda\ (\bar{A})]$ où la fonctionnelle de Cramer $\Lambda$ se déduit du $\lambda$ de (11) par la formule (3). Ici, $P_x$ est la loi de $y^\varepsilon$ quand $y^\varepsilon_0 \equiv x$.

Chapitre IV : Goulots de sortie et contrôle des équilibres stables

Nous précisons l'uniformité "en A" et "en x" des estimations obtenues pour $\lim_{\varepsilon\to 0} \varepsilon^2 \log P_x (y^\varepsilon \in A)$. Ces points techniques sont cruciaux pour le recollement des résultats précédents sur une variété, ou pour l'estimation des probabilités conditionnelles du type $P\,[y^\varepsilon \in A \mid \mathcal{F}_t]$ . La formulation est plus proche de celle de Ventsel-Freidlin [53] , mais la présence de temps d'explosion et la non inversibilité de $\sigma$ (x) rendent nécessaire l'utilisation méticuleuse de l'approximation G de F , via (8) et (9).

Nous prouvons alors le résultat de Ventsel-Freidlin [53] sur le "goulot de sortie" : si 0 est un point d'équilibre "très stable" de $y'_t = b\,(y_t)$ dans l'ouvert U, le processus $y^\varepsilon$ ne peut sortir de U qu'en suivant un tube d'axe $\varphi$ , où $\varphi$ est la trajectoire (supposée unique) joignant 0 à $\partial U$ dans l'intervalle de temps $[-\infty, T]$ et minimisant $\int_{-\infty}^{T} Q^*_{\varphi_t} (\varphi'_t - b\,(\varphi_t))\, dt$.

Nous construisons de façon heuristique les équations différentielles vérifiées par les courbes minimisantes (telles que $\varphi$) et par le "quasi-potentiel"

$$V\,(x,\,y) = \inf_{g_0=x,\,g_T=y} [\lambda_{0,T}\,(g)]$$

Nous décrivons ensuite, d'après Ventsel-Freidlin [54] , quelques problèmes de contrôle (choix de U avec volume (U) = constante, choix de b pour U donné) associés à la situation précédente : il s'agit soit de faciliter au maximum la détection d'une perturbation (minimiser le temps de sortie de U pour $y^\varepsilon$), soit de préserver au mieux la stabilité de l'équilibre perturbé (maximiser le temps de sortie de U pour $y^\varepsilon$).

## Chapitre V : Systèmes dynamiques perturbés sur une variété ; diffusions en temps petit.

Nous considérons une variété différentiable connexe M quelconque, un opérateur différentiel $\Delta$ du second ordre, semi-elliptique sur M, et un champ de vecteurs b. Nous traitons deux cas

(13) Cas elliptique : $\Delta$ est elliptique, b et $\Delta$ sont à coefficients localement lipschitziens, et dans suffisamment de cartes locales, la matrice a(x) des "coefficients du second ordre" de $\Delta$ se factorise par $a(x) = \sigma(x)\, \sigma^*(x)$ avec $\sigma(x)$ champ $C^1$ de matrices rectangulaires.

(14) Cas hypoelliptique : b, $\Delta$, M sont $C^\infty$ , b + $\Delta$ vérifie localement la condition d'hypoellipticité de Hormander, et la forme quadratique définie par "les termes du second ordre" de $\Delta$ est de rang constant.

Dans la situation (13) ou (14), nous notons $y_t^\varepsilon$ la $(b + \varepsilon^2 \Delta)$-diffusion sur M. Nous construisons la transformée de Cramer $\lambda$ par une formule (intrinsèque) analogue à (11) et nous démontrons par "recollement" que les limites inf. et sup. de $\varepsilon^2 \log P_x (y^\varepsilon \in A)$ sont dans l'intervalle $[- \Lambda (\mathring{A}), - \Lambda (\overline{A})]$ , pour $A \subset \mathcal{C}_x (M)$. Le travail de recollement est non-trivial et s'inspire en partie de Azencott-Ruget [3] qui traitaient un problème du même type pour certains processus à temps discret.

Nous appliquons ce résultat à l'étude des diffusions en temps petit. On se donne un champ b et un opérateur différentiel $\Delta$ sur une variété M, vérifiant soit (13), soit (14). Notons $(y_t)$ la $(b + \Delta)$-diffusion sur M ; par un simple changement de temps nous ramenons l'étude des trajectoires $y_{[0,t]}$ de la $(b + \Delta)$-diffusion lorsque $t \to 0$, à celle des trajectoires sur $[0, 1]$ de la $(\varepsilon^2 b + \varepsilon^2 \Delta)$-diffusion lorsque $\varepsilon \to 0$.

Nous évaluons ainsi $\lim_{t\to 0} t \log P_x (y_t \in F)$ et $\lim_{t\to 0} t \log P_x (\tau_F \leq t)$, où $F \subset M$ , et où $\tau_F$ est le temps de première sortie de F. Ces limites se calculent à partir de $\Delta$ et F seulement , sans faire intervenir b ; nous donnons les interprétations géométriques simples de ces résultats dans le cas elliptique et mentionnons quelques "pathologies" du cas hypoelliptique.

Chapitre VI : Questions sans réponses

## 4. COMPLEMENTS BIBLIOGRAPHIQUES

Faute de place, nous avons omis un bon nombre de développements intéressants. Donnons une liste (non exhaustive) des thèmes ignorés dans ce cours, mais fortement liés aux théories exposées. Pour chaque question nous indiquons un germe de bibliographie, qui sera loin d'épuiser la littérature sur le sujet.

1. Application en statistique : voir Séminaire Orsay [47], Bahadur [5] [6] Chernoff [10] ...

2. Raffinements de la théorie unidimensionnelle :

   - travaux russes, qui classent méticuleusement (et péniblement) les équivalents de $P\left[\overline{X}_n \geq g(n)\right]$ suivant l'allure des fonctions $n \to g(n)$ et $t \to \mu\ [t, +\infty]$ : Linnik [31] Nagaev [38] Petrov [39] ...
   - développements du type :

   $$P(\overline{X}_n \geq a) \sim e^{-n\lambda(a)} \left[f(a) + \frac{1}{n} f_1(a) + \frac{1}{n^2} f_2(a) + \dots\right]$$

   Feller [20] [21] Bahadur [6] Cramer [14] ...

3. V.a. à valeurs dans $\mathbb{R}^k$

   - comportement de $\lambda$ au bord du support de $\mu$ et robustesse du calcul de $\lim\limits_{n\to\infty} \frac{1}{n} \log P(\overline{X}_n \in A)$ quand $\mu$ varie. Azencott-Ruget [3] ...

- pathologies de $\lambda$ ; survol de la théorie ; bibliographies Bahadur-Zabell [7] ...

4. Cas où les $X_n$ ne sont pas indépendantes équidistribuées

- chaînes de Markov récurrentes en temps très long : Donsker-Varadhan [15]...
- contrôle de telles chaînes : Maigret [47]...
- diffusions récurrentes en temps très long : Donsker-Varadhan [15]
- suites de chaînes de Markov à sauts tendant vers 0 : Ventsel [51]
- mélange d'équations différentielles ; processus d'apprentissage lent : Varadhan [50] Azencott-Ruget [3]...

5. Développements asymptotiques du type

$$E\left[e^{1/\varepsilon^2\,\theta(y^\varepsilon)}\right] = e^{\frac{K}{\varepsilon^2}}\,(C + C_1\,\varepsilon + C_2\,\varepsilon^2 + \ldots)\,, \qquad \varepsilon \to 0$$

où $y^\varepsilon$ est la trajectoire d'une $(b + \varepsilon^2\,\Delta)$-diffusion : Schilder [44] Doss [17]...

6. Raffinement des résultats sur les diffusions en temps petit

- développement asymptotique précis de la densité $p\,(t, x, y)$ d'une diffusion quand $t \to 0$, avec $\Delta$ elliptique. Molchanov [37] ...
- même problème sur quelques exemples fondamentaux avec $\Delta$ hypoelliptique. Gaveau [24] [25] Séminaire Paris 7 [46] ...

# CHAPITRE I

# THEOREMES A LA CRAMER-CHERNOFF

## 1. LE PROBLEME DES GRANDES DEVIATIONS

Soit E un espace vectoriel, $X_n$ une suite de variables aléatoires (v.a.) indépendantes, de même loi $\mu$, à valeurs dans E. Soit $\overline{X}_n = \frac{1}{n}(X_1 + \ldots + X_n)$. On étudie le comportement asymptotique de $P(\overline{X}_n \in A)$ quand $n \to +\infty$. Quand $E = \mathbb{R}^k$ et $\int |x| \, d\mu(x)$ est fini, la loi des grands nombres montre que $\lim\limits_{n\to+\infty} P(\overline{X}_n \in A) = 1$ dès que A contient un voisinage de $m = \int x \, d\mu(x)$, tandis que $\lim\limits_{n\to+\infty} P(\overline{X}_n \in A) = 0$ lorsque $m \notin \overline{A}$. Pour n grand, un évènement du type $\{\overline{X}_n \in A\}$, avec $m \notin \overline{A}$, représente une situation de "grande déviation par rapport à la loi des grands nombres". Le qualificatif "grande" traduit le fait que $\overline{X}_n$ évite alors un voisinage de m dont la taille ne tend pas vers 0 quand $n \to +\infty$.

Les résultats de Cramer [14] et de Chernoff [10] traitent le cas où $E = \mathbb{R}$, et calculent $\lim\limits_{n\to+\infty} \frac{1}{n} \log P(\overline{X}_n \in A)$ lorsque A est une demi-droite. L'extension de leurs résultats aux espaces de Banach est due à Donsker-Varadhan [15], et a été formalisée pour des espaces vectoriels plus généraux par Bahadur-Zabell [7] qui ont fait bon usage d'une idée de Lanford [30].

## 2. LE THEOREME DE CRAMER-CHERNOFF SUR $\mathbb{R}$

2.1. Soit $\mu$ une probabilité sur $\mathbb{R}$. Soit $\hat{\mu} : \mathbb{R} \to ]0, +\infty]$ sa transformée de Laplace, définie par $\hat{\mu}(t) = \int_{\mathbb{R}} e^{tx} \, d\mu(x)$. Définissons la transformée de Cramer $\lambda : \mathbb{R} \to [0, +\infty]$ de la mesure $\mu$ par

2.2 $$\lambda(x) = \sup_{t\in\mathbb{R}} [tx - \log \hat{\mu}(t)] \qquad , \quad x \in \mathbb{R}$$

En tant qu'enveloppe supérieure d'une famille de fonctions linéaires, $\lambda$ est donc <u>convexe</u> et <u>semi-continue inférieurement</u> (s.c.i.)

Notons que les valeurs infinies de $\hat{\mu}$ (t) ne sont pas exclues. En particulier si $\hat{\mu}$ (t) est infinie pour tout $t \neq 0$, on a $\lambda \equiv 0$.

2.3. <u>Théorème</u> (Cramer [14] Chernoff [10] ). Soit $X_n$ une suite de v.a. indépendantes de même loi $\mu$. Soit $\overline{X}_n = \frac{1}{n}(X_1 + \ldots + X_n)$. Soit $\lambda$ la transformée de Cramer de $\mu$. On a alors pour $a \in \mathbb{R}$

(i) $$-\lambda(a) \leqslant \underline{\lim}_{n\to+\infty} \frac{1}{n} \log P(\overline{X}_n \leqslant a)$$

$$-\lambda(a) \leqslant \underline{\lim}_{n\to+\infty} \frac{1}{n} \log P(\overline{X}_n \geqslant a)$$

(ii) Supposons de plus $\int |x| \, d\mu(x)$ fini.

Alors on a les inégalités <u>valables pour tout n entier</u>

$$\frac{1}{n} \log P(\overline{X}_n \leqslant a) \leqslant -\lambda(a), \text{ pour } a \leqslant \int x \, d\mu(x)$$

$$\frac{1}{n} \log P(\overline{X}_n \geqslant a) \leqslant -\lambda(a), \text{ pour } \int x \, d\mu(x) \leqslant a.$$

<u>Preuve</u> : Esquissons l'élégante démonstration de Chernoff [10] . Les lois des v.a. $Y_n = X_n - a$ et $Z_n = -X_n$ ont resp. pour transformées de Cramer les fonctions $\lambda_y(x) = \lambda(x + a)$ et $\lambda_z(x) = \lambda(-x)$. Il suffit donc de traiter le cas $\{a = 0\}$ et de prouver seulement les inégalités concernant $P(\overline{X}_n \geqslant 0)$.

Supposons $\int x \, d\mu(x) \leqslant 0$. La définition de $\lambda$ entraîne alors

$$\lambda(0) = \sup_t \left[-\log \hat{\mu}(t)\right] = \sup_{t\geqslant 0} \left[-\log \hat{\mu}(t)\right]$$

Pour $t \geqslant 0$, on peut écrire

$$P(\overline{X}_n \geqslant 0) = P\left[e^{t(X_1+\ldots+X_n)} \geqslant 1\right]$$

$$\leqslant E\left[e^{t(X_1+\ldots+X_n)}\right] = \left[\hat{\mu}(t)\right]^n$$

Ceci donne

$$\frac{1}{n} \log P(\overline{X}_n \geqslant 0) \leqslant \inf_{t\geqslant 0} \left[\log \hat{\mu}(t)\right] = -\lambda(0)$$

et prouve donc (ii).

La démonstration de (i) utilise un lemme technique ("cas où $\mu$ est à support fini") :

2.4. <u>Lemme</u> (Chernoff [10]) : Soient $x_i$, $p_i$, $1 \leqslant i \leqslant k$ des nombres réels tels que $p_i > 0$ et $\inf_i (x_i) < 0 < \sup_i (x_i)$. Posons

$b = \inf_{t\in\mathbb{R}} \left(\sum_i p_i e^{tx_i}\right)$. Posons $F(n_1, \ldots, n_k) = n! \prod_i \dfrac{p_i^{n_i}}{n_i\,!}$

On peut alors déterminer une constante $c > 0$ et un entier $N$ tel que tout entier $n \geqslant N$ puisse s'écrire $n = n_1 + \ldots + n_k$, où les entiers positifs $n_i$ vérifient $\sum_i n_i x_i \geqslant 0$ et $F(n_1 \ldots n_k) \geqslant c\, n^{-k/2}\, b^n$.

<u>Preuve</u> : la formule de Stirling donne pour $z_i \in [1, +\infty[$

$$F(z_1, \ldots, z_k) \geqslant n^{-k/2} \prod_i \left(\frac{n\,p_i}{z_i}\right)^{z_i} = G(z_1 \ldots z_k)$$

Fixons $s \in \mathbb{R}$ tel que $b = \sum_i p_i\, e^{s\,x_i}$

Les réels $Z_i = \dfrac{n\,p_i\,e^{sx_i}}{b}$ vérifient alors

$$\sum_i Z_i = n \quad ; \quad \sum_i Z_i\, x_i = 0$$

$$G(Z_1 \ldots Z_k) = n^{-k/2}\, b^n.$$

Sans perte de généralité, supposons $x_i \leqslant x_k$ pour tout i. Posons $N_i = [Z_i]$ pour $i < k$ et $N_k = n - \sum_{i<k} N_i$.

Pour $A > 0$ fixé, et pour B variant dans un intervalle borné fixe, il existe une constante $c_1$ telle que

$$\left(\frac{v}{Av + B}\right)^{Av+B} \geqslant c_1 \left(\frac{v}{Av}\right)^{Av}$$

pour tout v vérifiant $v \geqslant 1$ et $Av+B \geqslant 1$. Ceci montre l'existence d'une constante $c_2$ telle que $\left(\frac{n\,p_i}{N_i}\right)^{N_i} \geqslant c_2 \left(\frac{n\,p_i}{Z_i}\right)^{Z_i}$ pour tout i et tout n assez grand. Donc, pour n assez grand, on aura, avec $c_3 = c_2^k$

$$\sum_i N_i = n \quad ; \quad \sum_i N_i\, x_i \geqslant 0$$

$$G\,(N_1 \ldots N_k) \geqslant c_3 \; n^{-k/2}\, b^n$$

ce qui prouve le lemme.

Revenons à la preuve de (i). Le résultat est trivial lorsque l'un des nombres $\mu\,(\,]0, +\infty[\,)$ et $\mu\,(\,]-\infty, 0[\,)$ est nul. Supposons les donc non nuls.

Considérons d'abord le cas où $\mu$ est de la forme $\mu = \sum_{i \geqslant 1} p_i\, \delta_{x_i}$, la somme étant finie ou dénombrable, avec $p_i > 0$ pour tout i. Pour k assez grand, on aura

$$\inf_{1 \leqslant i \leqslant k} x_i < 0 < \sup_{1 \leqslant i \leqslant k} x_i$$

La suite monotone de fonctions $f_k\,(t) = \sum_{1 \leqslant i \leqslant k} p_i\, e^{tx_i}$ converge vers $\hat{\mu}\,(t)$ quand $k \to +\infty$, et donc (théorème de Dini) converge uniformément vers $\hat{\mu}\,(t)$ sur tout compact contenu dans $\{\, t \mid \hat{\mu}\,(t) \text{ fini} \}$. Comme $f_k\,(t)$ tend vers $+\infty$ avec $|t|$, on en conclut que

$$\lim_{k \to +\infty} \left[\inf_t\; f_k(t)\right] = \inf_t\; \hat{\mu}\,(t) = e^{-\lambda(0)}$$

Posons $b_k = \inf_t f_k(t)$ et fixons k. Avec les notations du lemme, on a pour tout n, $P(\overline{X}_n \geqslant 0) \geqslant F(n_1 \dots n_k) \geqslant c_k\, n^{-k/2}\, b_k^n$ et donc

$$\varliminf_{n\to+\infty} \frac{1}{n} \log P(\overline{X}_n \geqslant 0) \geqslant \log b_k$$

ce qui, lorsque $k \to +\infty$, prouve (i).

Passons au cas général. Posons $Y_n^{(s)} = \frac{j}{s}$ dès que $\frac{j-1}{s} \leqslant X_n \leqslant \frac{j}{s}$

Soit $\mu_s$ la loi commune des $Y_n^{(s)}$ ; soit $\lambda_s$ la transformée de Cramer de $\mu_s$.
La construction de $\mu_s$ donne

$$\hat{\mu}_s(t) \geqslant e^{-\frac{|t|}{s}}\, \hat{\mu}(t) \quad \text{pour tout } t, \text{ et } s > 0.$$

Comme $]0, +\infty[$ et $]-\infty, 0[$ ne sont pas $\mu$-négligeables, on a $\hat{\mu}(t) \geqslant A\, e^{B|t|}$ où A et B sont des constantes positives strictes. On peut écrire

$$e^{-\lambda_s(1/s)} = \inf_t \left[e^{-t/s}\, \hat{\mu}_s(t)\right] \geqslant \inf_t \left[e^{-2|t|/s}\, \hat{\mu}(t)\right]$$

Les fonctions $g_s(t) = e^{-2|t|/s}\, \hat{\mu}(t)$ convergent en croissant vers $\hat{\mu}(t)$ quand $s \to +\infty$, et tendent vers $+\infty$ quand $|t| \to +\infty$, pourvu que $s > \frac{2}{B}$. Comme plus haut, ceci entraîne

$$\lim_{s\to+\infty} \left[\inf_t g_s(t)\right] = \inf_t \hat{\mu}(t) = e^{-\lambda(0)}$$

et par suite $\lim\limits_{s\to+\infty} - \lambda_s(1/s) \geqslant -\lambda(0)$.

L'étude du cas discret implique, pour <u>s fixé</u>,

$$\varliminf_{n\to+\infty} \frac{1}{n} \log P(\overline{Y}_n^{(s)} \geqslant 1/s) \geqslant -\lambda_s(1/s)$$

Compte-tenu de l'inclusion de $\{\overline{Y}_n^s \geqslant 1/s\}$ dans $\{\overline{X}_n \geqslant 0\}$, on en conclut que

$$\varliminf_{n\to+\infty} \frac{1}{n} \log P(\overline{X}_n \geqslant 0) \geqslant -\lambda_s(1/s)$$

Lorsque $s \to +\infty$ ceci prouve (i).

2.5. Remarque : Pour 2.3 (ii), les restrictions sur le signe de $\left[a - \int x \, d\mu\,(x)\right]$ ne sont là que pour permettre de formuler simplement le seul résultat non trivial. En effet si $\int x \, d\mu\,(x) < a$ , on a par exemple

$$\lim_n \frac{1}{n} \log P\,(\overline{X}_n \leqslant a) = \sup_n \frac{1}{n} \log P\,(\overline{X}_n \leqslant a) = 0$$

2.6. Exemples de transformées de Cramer sur $\mathbb{R}$

Le théorème 2.3 fait pressentir que la probabilité de trouver $\overline{X}_n$ au voisinage de x est "de l'ordre de" $e^{-n\lambda(x)}$ ; sous cette forme le résultat est littéralement faux, mais cet énoncé donne cependant le contenu intuitif du théorème 2.3, comme le précisera le § 3. Les grandes valeurs de $\lambda$ correspondent donc aux points où les apparitions de $\overline{X}_n$ sont peu probables.

Lorsque $\hat{\mu}\,(t)$ est finie au voisinage de 0, la fonction $\lambda$ atteint son minimum au point $m = \int_{\mathbb{R}} x \, d\mu\,(x)$ et $\lambda''(m) = \frac{1}{\sigma^2}$ . La formule 2.2 fournit facilement la forme de $\lambda$ dans les cas suivants.

(1) Quand $\mu$ est gaussienne, de moyenne m, de variance $\sigma^2$ , on a :

$$\lambda(x) = \frac{1}{2\,\sigma^2}\,(x-m)^2 \qquad \text{pour } x \in \mathbb{R}$$

(2) Quand $\mu = p\,\delta_u + (1-p)\,\delta_v$ , avec $u < v$ et $0 < p < 1$, on a, pour $u < x < v$

$$\lambda\,(x) = \frac{x-u}{v-u} \log \frac{x-u}{1-p} + \frac{v-x}{v-u} \log \frac{v-x}{p} - \log\,(v-u)$$

$$\lambda\,(u) = -\log p \qquad\qquad \lambda\,(v) = -\log\,(1-p)$$

$$\lambda\,(x) = +\infty \qquad \text{pour } x < u \text{ ou } x > v$$

(3) Quand $d\mu\,(x) = 1_{[0,+\infty[}\,(x)\; e^{-x}\,dx$ , on a

$$\lambda\,(x) = x - 1 - \log x \qquad \text{pour } x > 0$$

$$\lambda\,(x) = +\infty \qquad \text{pour } x \leqslant 0$$

Le résultat qui suit présente un intérêt descriptif ; la preuve par contre mérite peu d'attention en première lecture.

2.7. Proposition : Soit $\mu$ une probabilité sur $\mathbb{R}$ telle que $\int |x| \, d\mu\,(x)$ soit finie. Soit $\lambda$ la transformée de Cramer de $\mu$. Alors

(1) $\hat{\mu}\,(t) \equiv +\infty$ pour tout $t > 0$ si et seulement si $\lambda\,(x) \equiv 0$ pour tout $x \geqslant 0$.

(2) il existe $t > 0$ tel que $\hat{\mu}\,(t)$ soit fini si et seulement si

$\lim_{x \to +\infty} \lambda\,(x) = +\infty$

(3) $\hat{\mu}\,(t)$ est fini pour tout $t > 0$ si et seulement si $\lim_{x \to +\infty} \frac{\lambda(x)}{x} = +\infty$

(4) L'intégrale $\int_{\mathbb{R}} e^{s\lambda(x)} \, d\mu\,(x)$ est finie pour tout $s < 1$.

(5) Soit $[u, v] \subset [-\infty, +\infty]$ l'enveloppe convexe fermée du support de $\mu$ ; on a $\lambda\,(x) = +\infty$ pour $x \notin [u, v]$, tandis que $\lambda\,(x)$ est finie et continue pour $x \in \,]u, v[$ ; de plus $\lambda$ est continue à gauche en $v$ si $v$ est fini, et continue à droite en $u$ si $u$ est fini ; enfin pour $v$ fini, la relation $\lambda\,(v) = +\infty$ équivaut à $\mu\,(v) = 0$ (résultat analogue pour $u$ fini).

(6) Les résultats (1) (2) (3) (4) restent vrais si on remplace $t > 0$, $x > 0$, $x \to +\infty$ par $t < 0$, $x < 0$, $x \to -\infty$.

Preuve : Soit $m = \int x \, d\mu\,(x)$. Soit $[u, v]$ comme en (5). Puisque $\overline{X}_n \in [u, v]$ presque sûrement, le résultat 2.3 (i) appliqué avec $a < u$ et $a > v$ montre que $\lambda(a) = +\infty$ pour $a \notin [u, v]$.
Inversement si $a \geqslant m$ et $\lambda\,(a) = +\infty$, le théorème 2.3 (ii) entraîne $P\,(\overline{X}_n \geqslant a) = 0$, et donc $\mu\,(a) = 0$, ce qui impose $a \geqslant v$. Argument analogue pour $a \leqslant m$. Donc $\lambda$ est finie sur $]u\; v[$ ; de plus on vient de voir que $\lambda\,(v) = +\infty$ implique $\mu\,(v) = 0$ et le théorème 2.3 (i) prouve directement la réciproque.

Toutes les propriétés de continuité de $\lambda$ énoncées en (5) sont alors des conséquences simples du fait que $\lambda$ est convexe s.c.i. Ceci achève la preuve de (5).

Supposons $\hat{\mu}(t) \equiv +\infty$ pour tout $t > 0$. La relation 2.2 donne

$\lambda(x) = \sup\limits_{t \leqslant 0} [tx - \log \hat{\mu}(t)]$ et donc $\lambda(x) = 0$ pour $x \geqslant 0$.

Supposons qu'il existe $t > 0$ avec $\hat{\mu}(t)$ fini. La relation

$\lambda(x) \geqslant tx - \log \hat{\mu}(t)$ montre que $\lim\limits_{x \to +\infty} \lambda(x) = +\infty$.

Les implications directes en (1) et (2) sont prouvées ; l'implication inverse dans (2) résulte clairement de l'implication directe en (1) - et vice-versa ! Ainsi (1) et (2) sont vérifiées.

Prouvons (4) ; cette relation est triviale dans le cas où $\hat{\mu}(t) \equiv +\infty$ pour tout $t > 0$, à cause de (1). Supposons donc $\hat{\mu}(t)$ fini pour au moins un $t > 0$. Supposons aussi (notations (5)) que $v = +\infty$ (car d'après (5), la relation (4) est triviale si $v$ est fini).

Comme $\lambda$ est maintenant convexe continue sur $]u, +\infty[$ , sa dérivée $\lambda'$ existe sur $]u, +\infty[$ sauf peut-être sur un ensemble dénombrable. Une intégration par parties montre, grâce à la relation $\lim\limits_{x \to +\infty} \lambda(x) = +\infty$, que $\int_m^{+\infty} e^{s\lambda(x)}\, d\nu(x)$ sera finie si l'intégrale

$I = \int_m^{+\infty} e^{s\lambda(x)}\, \lambda'(x)\, \mu[x, +\infty[\, dx$ est finie. D'après 2.3 (ii) appliqué à $n = 1$, on a $\mu[x, +\infty[\, \leqslant e^{-\lambda(x)}$ pour $x \geqslant m$, et donc

$$I \leqslant \int_m^{+\infty} e^{(s-1)\lambda(x)}\, \lambda'(x)\, dx < +\infty \quad \text{pour } s < 1.$$

Une étude analogue de $\int_{-\infty}^{m}$ achève de prouver (4).

Supposons maintenant $\hat{\mu}(t)$ fini pour tout $t > 0$. D'après 2.2, on a alors pour tout $t > 0$, $x > 0$ l'inégalité $\dfrac{\lambda(x)}{x} \geqslant t - \dfrac{1}{x} \log \hat{\mu}(t)$ et donc $\underline{\lim}_{x \to +\infty} \dfrac{\lambda(x)}{x} \geqslant t$. Comme $t$ est arbitraire, on obtient

$$\lim_{x \to +\infty} \frac{\lambda(x)}{x} = +\infty .$$

Inversement, si $\dfrac{\lambda(x)}{x}$ tend vers $+\infty$ avec $x$, on aura pour tout $t > 0$ <u>fixé</u> la majoration $tx \leqslant \dfrac{1}{2}\, \lambda(x)$ pourvu que $x$ soit assez grand. Utilisant (4) on conclut que

$$\int_A^{+\infty} e^{tx} \, d\mu\,(x) \leq \int_A^{+\infty} e^{1/2\ \lambda(x)} \, d\mu\,(x) < +\infty \quad ;$$

ceci implique que $\hat{\mu}$ (t) est fini. Ainsi (3) est démontrée.

2.8. Remarque : dans la relation (4) de 2.7, la restriction $s < 1$ ne peut pas être améliorée en général. En effet (cf. 2.6 (1) et (2)) pour $\mu$ gaussienne ou exponentielle on a $\int e^{\lambda(x)} \, d\mu\,(x) = +\infty$.

Ceci reste sans doute vrai pour toute probabilité $\mu$ sur $\mathbb{R}$ à support non compact telle que $\hat{\mu}$ (t) soit finie pour au moins un $t \neq 0$. Dans ce cas, ceci caractériserait $\lambda$ comme la fonction convexe "la plus grande à l'infini" telle que $\int e^{s\lambda(x)} \, d\mu\,(x)$ soit finie pour $s < 1$ et infinie pour $s = 1$.

## 3. CAS DES V.A. A VALEURS DANS UN ESPACE VECTORIEL : LES HYPOTHESES TOPOLOGIQUES

Les hypothèses minimales qui permettent la construction formelle de la transformée de Cramer semblent à présent dégagées par Bahadur-Zabell [7] . Nous les renforçons ici pour leur donner une forme plus descriptive, donc plus pratique.

3.1. Dans tout le chapitre, nous considérons désormais des v.a. indépendantes $X_n : \Omega \to E$, à valeurs dans un espace vectoriel topologique E, localement convexe et séparé (e.v.t.l.c.s.) muni de sa $\sigma$-algèbre borélienne $\mathcal{B}(E)$. Nous noterons $\mu$ la loi commune des $X_n$, et nous supposerons que l'espace de probabilité $\Omega$ est complet.

3.2. Hypothèse technique essentielle : nous supposerons l'existence d'un convexe fermé F de E, vérifiant $\mu(F) = 1$, et l'existence d'une autre structure $\tilde{E}$ d'e.v.t.l.c.s. sur l'espace vectoriel E, topologiquement plus fine que celle de E, telle que la topologie induite par $\tilde{E}$ sur F soit polonaise (c'est-à-dire métrisable, complète, à base dénombrable).

3.3. Exemples : E est un espace de Fréchet séparable et $\mu$ est une probabilité quelconque sur $\mathcal{B}(E)$. Puisque E est polonais il suffit ici de prendre $F = E = \tilde{E}$ dans 3.2. Rappelons que ce cas inclut celui des espaces de Banach séparables, des espaces de Hilbert séparables, et le cas $E = \mathbb{R}^k$.

3.4. Exemples : E est un espace de Fréchet séparable, muni d'une topologie faible arbitraire, et $\mu$ est une probabilité quelconque sur $\mathcal{B}(E)$. Ici, 3.2 est réalisée avec F = E tandis que $\tilde{E}$ est l'espace E muni de sa topologie de Fréchet séparable. Ce cas inclut celui des Fréchet (Banach, Hilbert) séparables muni de la topologie $\sigma$ (E, E') où E' est le dual de E.

3.5. Exemple : E est l'espace $\mathcal{M}b(\Gamma)$ des mesures boréliennes bornées sur l'espace topologique polonais $\Gamma$ ; E est muni de la topologie de la convergence étroite, c'est-à-dire de la topologie associée à la dualité naturelle entre $\mathcal{M}b(\Gamma)$ et l'espace C $(\Gamma)$ des fonctions continues bornées sur $\Gamma$. La probabilité $\mu$ sur $\mathcal{B}(E)$ vérifie $\mu\,[\mathcal{M}b_1(\Gamma)] = 1$, où $\mathcal{M}b_1(\Gamma)$ est l'ensemble des probabilités boréliennes sur $\Gamma$.

Les hypothèses 3.2 sont vérifiées avec $F = \mathcal{M}b_1(\Gamma)$ et $E = \tilde{E}$. Cet exemple fournira le cadre approprié pour traiter le théorème de Sanov sur les lois empiriques (cf. § 8).

Notons ici une remarque de A. Badrikian : comme il n'existe pas de topologie polonaise plus fine que celle de E, on peut trouver des probabilités $\mu$ sur $\mathcal{B}(E)$ telles que 3.2 ne soit pas vérifiée.

Nous allons prouver deux lemmes (3.6 et 3.8) assez techniques, qui constituent l'essentiel des hypothèses de Bahadur-Zabell [7] . En première lecture, nous suggérons d'ignorer les démonstrations de ces lemmes.

3.6. <u>Lemme</u> : Soient E et F comme en 3.2. Si X, Y : $\Omega \to E$ sont deux v.a. <u>à valeurs dans F</u> , alors X + Y : $\Omega \to E$ est encore une v.a. En particulier sous les hypothèses 3.1 et 3.2, toute combinaison linéaire convexe des $X_n$ est une v.a.

<u>Preuve</u> : D'après 3.2, F est un espace lusinien, et on a donc
$\mathcal{B}(F \times F) = \mathcal{B}(F) \otimes \mathcal{B}(F)$ (cf. Badrikian [4] ). L'application
$\varphi(x, y) = x + y$ de F x F dans E est continue, et par suite est
$\{\mathcal{B}(F \times F), \mathcal{B}(E)\}$ mesurable.
L'application $(X, Y) : \Omega \to [F \times F, \mathcal{B}(F) \otimes \mathcal{B}(F)]$ est mesurable et donc $X + Y = \varphi(X, Y)$ est mesurable.

3.7. <u>Remarque</u> (A. Badrikian). Dans la proposition précédente, le résultat n'est plus nécessairement vrai si on suppose seulement que X, Y sont à valeurs dans E.

3.8. <u>Lemme</u> (d'après Bahadur-Zabell [7] ) : Soit $(E, \mu)$ un couple vérifiant 3.2. Alors pour tout ouvert convexe A de E et tout $\varepsilon > 0$, il existe un compact convexe K contenu dans A, tel que $\mu(A - K) \leqslant \varepsilon$ .

<u>Preuve</u> : Soient E, $\tilde{E}$, F et $\mu$ comme en 3.2 ; soit $i : \tilde{E} \to E$ l'application identique. Appelons $\tilde{F}$ l'espace topologique induit par $\tilde{E}$ sur l'ensemble F. Puisque i est continue injective, et puisque $\tilde{F}$ est polonais, i est aussi un isomorphisme de $\mathcal{B}(\tilde{F})$ sur $\mathcal{B}(F)$ (cf. Christensen [11] ), ce qui permet de considérer $\mu$ comme une probabilité sur $(\tilde{E}, \mathcal{B}(\tilde{E}))$. Puisque i est continue, il suffit alors clairement de prouver le lemme pour le couple $(\tilde{E}, \mu)$. Nous pouvons donc désormais supposer que F est un espace polonais.

La mesure $\mu$ est alors régulière. Si A est un ouvert convexe de E, il existe pour tout $\varepsilon > 0$ un compact G de F tel que $G \subset A \cap F$ et

$\mu$ (A - G) $\leqslant \varepsilon$ . Montrons que l'enveloppe convexe fermée de G est un compact $K \subset A$.

Il existe pour tout $x \in G$ un voisinage ouvert convexe $V_x$ de x tel que $\overline{V}_x \subset A$ , car E est un e.v.t.l.c.s. La compacité de G fournit une famille finie de points de G telle que $G \subset \overline{V}_{x_1} \cup \ldots \cup \overline{V}_{x_n}$ . Posons $G_i = G \cap \overline{V}_{x_i}$. Chaque $G_i$ est compact et inclus dans le convexe fermé polonais F.

L'enveloppe convexe fermée $\hat{G}_i$ de $G_i$ est alors encore compacte (car pré-compacte et complète). De plus l'inclusion $G_i \subset \overline{V}_{x_i}$ entraîne $\hat{G}_i \subset \overline{V}_{x_i} \subset A$. Finalement G est contenu dans $\hat{G}_1 \cup \ldots \cup \hat{G}_n$ , qui est une partie de A. Comme les $\hat{G}_i$ sont compacts convexes, l'enveloppe convexe fermée H de $\hat{G}_1 \cup \ldots \cup \hat{G}_n$ est exactement l'ensemble des $\sum_{i=1}^{n} \alpha_i v_i$ avec $v_i \in \hat{G}_i$ , et $\sum_i \alpha_i = 1$, $\alpha_i \geqslant 0$. Il est donc clair que $H \subset A$. De $G \subset H$ on conclut alors $\hat{G} \subset H$ et enfin $\hat{G} \subset A$. Le compact convexe $\hat{G}$ est contenu dans A et vérifie $\mu$ (A - $\hat{G}$) $\leqslant \varepsilon$ .

## 4. TRANSFORMEE DE CRAMER : CAS GENERAL

4.1. Notations : Soient E un espace vectoriel topologique et $\mu$ une probabilité sur E vérifiant 3.1, 3.2. Soit $(\Omega, P)$ un espace de probabilité complet et soit $X_n : \Omega \to E$ une suite de v.a. indépendantes de même loi $\mu$. On pose $\overline{X}_n = \frac{1}{n} (X_1 + \ldots + X_n)$ et on note $\mu_n$ la loi de $\overline{X}_n$.

Pour tout $A \in \mathcal{B}(E)$ on note

$$\underline{\ell} (A) = \varliminf_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A)$$

$$\overline{\ell} (A) = \varlimsup_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A)$$

de sorte que $-\infty \leqslant \underline{\ell} (A) \leqslant \overline{\ell} (A) \leqslant 0$.

Si $\underline{\ell} (A) = \overline{\ell} (A)$, on note $\ell (A) = \lim_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A)$.

Les méthodes et résultats du paragraphe 4 s'inspirent du travail de Bahadur-Zabell [7] .

4.2. Proposition (d'après [7] [30] ). Sous les hypothèses 3.1, 3.2 la limite $\ell$ (A) = $\lim\limits_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A)$ existe pour toute partie A convexe borélienne de E.

Preuve : Posons $X = \frac{1}{n} (X_1 + \dots + X_n)$, $Y = \frac{1}{m} (X_{n+1} + \dots + X_{n+m})$.

Alors on a $\overline{X}_{n+m} = \frac{n}{n+m} X + \frac{m}{n+m} Y$ , et la convexité de A entraîne $(X \in A) \cap (Y \in A) \subset (\overline{X}_{n+m} \in A)$. On en déduit $\mu_n (A)\, \mu_m (A) \leqslant \mu_{n+m} (A)$, et la fonction $f : N \to [0, +\infty]$ définie par $f(n) = -\log \mu_n (A)$ est sous additive. On sait alors que $\lim\limits_{n\to+\infty} \frac{f(n)}{n}$ existe (cf. Billingsley [8] ).

4.3. Proposition : Quels que soient A, B dans $\mathcal{B}(E)$, on a

$$\max \left[\underline{\ell} (A), \underline{\ell} (B)\right] \leqslant \underline{\ell} (A \cup B) \leqslant \overline{\ell} (A \cup B) \leqslant \max \left[\overline{\ell} (A), \overline{\ell} (B)\right]$$

En particulier si $\ell (A_i)$ existe pour i = 1 ... n , la limite $\ell (A_1 \cup \dots \cup A_n)$ existe aussi, et vaut $\max \left[\ell (A_1), \dots, \ell (A_n)\right]$.

Preuve : Notons d'abord que $\underline{\ell}$ et $\overline{\ell}$ sont des fonctions croissantes. Par suite on a $\underline{\ell} (A) \leqslant \underline{\ell} (A \cup B)$ et $\underline{\ell} (B) \leqslant \underline{\ell} (A \cup B)$, ce qui prouve la première inégalité de 4.3

Soit $c > \max \left[\overline{\ell} (A), \overline{\ell} (B)\right]$. Pour n grand, $\mu_n (A)$ et $\mu_n (B)$ sont majorés par $e^{nc}$ , et donc $\mu_n (A \cup B) \leqslant 2\, e^{nc}$ , ce qui entraîne $\overline{\ell} (A \cup B) \leqslant c$. Comme c est arbitraire, on obtient la dernière inégalité de 4.3.

4.4. Définition (cf. [7] [30]) : Sous les hypothèses 3.1, 3.2 on appelle transformée de Cramer de $\mu$ la fonction $\lambda : E \to [0, +\infty]$ définie pour $x \in E$ par

$$\lambda(x) = -\inf \{\ell(A) \mid A \text{ ouvert convexe, et } x \in A\}$$

Puisque $\ell$ est une fonction monotone d'ensemble, on peut considérer $[-\lambda(x)]$ comme la limite de $\ell(A)$ suivant le filtre des voisinages ouverts convexes de x.

Lorsque $E = \mathbb{R}$, cette définition ne coïncide pas a priori avec la définition 2.2. Nous verrons (th. 5.3) que lorsque les moments d'ordre 1 de toutes les projections (continues) unidimensionnelles de $\mu$ sont finis, la définition 2.2 de $\lambda$ s'étend au cas général, et est bien compatible avec la définition 4.4.

4.5 Proposition : Sous les hypothèses 3.1, 3.2, la transformée de Cramer $\lambda$ de $\mu$ est convexe et semi-continue inférieurement.

Preuve : pour $x \in E$ et $c < \lambda(x)$ donnés, il existe par définition un voisinage ouvert convexe A de x tel que $c < -\ell(A)$. Pour tout $y \in A$, on aura $\lambda(y) = \sup\{-\ell(B) \mid B \text{ voisinage ouvert convexe de } y\}$ et donc $\lambda(y) \geq -\ell(A) > c$. La fonction $\lambda$ vérifie bien $\lambda(x) \leq \underline{\lim}_{y \to x} \lambda(y)$, et est donc s.c.i.

La convexité d'une fonction se vérifie en restreignant cette fonction à toutes les droites de l'espace. Or pour une fonction $g : \mathbb{R} \to [0, +\infty]$ supposée s.c.i., la convexité de g équivaut à

$$g\left(\frac{x+y}{2}\right) \leq \frac{1}{2} g(x) + \frac{1}{2} g(y) \quad \text{pour tout } x, y \in \mathbb{R}$$

Il suffit donc de vérifier que $\lambda\left(\frac{x+y}{2}\right) \leq \frac{1}{2} \lambda(x) + \frac{1}{2} \lambda(y)$ pour tout $x, y \in E$.

Pour tout ouvert convexe A contenant $\frac{x+y}{2}$ , il existe des ouverts convexes B, C contenant resp. x et y tels que $(\frac{1}{2} B + \frac{1}{2} C) \subset A$.

Posons $X = \frac{1}{n} (X_1 + \dots + X_n)$ , $Y = \frac{1}{n} (X_{n+1} + \dots + X_{2n})$.

L'inclusion $\{(X \in B) \cup (Y \in C)\} \subset (\overline{X}_{2n} \in A)$ entraîne $\mu_n(B)\ \mu_n(C) \leqslant \mu_{2n}(A)$.

Un passage à la limite en n donne $\frac{1}{2}\ \ell\ (B) + \frac{1}{2}\ \ell\ (C) \leqslant \ell\ (A)$ et donc par définition de $\lambda\ (x)$ , $\lambda\ (y)$

$$- \frac{1}{2} \lambda\ (x) - \frac{1}{2} \lambda\ (y) \leqslant \ell\ (A)$$

Prenons l'inf. en A au second membre pour obtenir

$- \frac{1}{2}\ \lambda\ (x) - \frac{1}{2}\ \lambda\ (y) \leqslant - \lambda\ (\frac{x+y}{2})$.

4.6. <u>Définition</u> (cf. Donsker-Varadhan, Bahadur-Zabell) : soit $(E, \mu)$ un couple vérifiant 3.1, 3.2. Soit $\lambda$ la transformée de Cramer de $\mu$ (définition 4.4). Nous appelons <u>fonctionnelle de Cramer associée à $\mu$</u>, l'application $\Lambda$ de l'ensemble des parties de E dans $[0, +\infty]$ définie par

$$\Lambda\ (A) = \inf_{x \in A} \lambda\ (x) \qquad , \qquad A \subset E$$

La donnée de $\lambda$ est clairement équivalente à celle de $\Lambda$.

4.7. <u>Proposition</u> (d'après [7] [15]). Soit $X_n : \Omega \to E$ une suite de v.a. indépendantes de même loi $\mu$. Supposons les hypothèses 3.2 et 3.1 vérifiées. Soit $\Lambda$ la fonctionnelle de Cramer associée à $\mu$. Alors pour tout $A \in \mathcal{B}(E)$ on a, en notant $\mathring{A}$ l'intérieur de A

(1) $$- \Lambda\ (\mathring{A}) \leqslant \underline{\ell}\ (A) = \varliminf_{n \to +\infty} \frac{1}{n} \log P\ (\overline{X}_n \in A)$$

(2) si de plus $\overline{A}$ est <u>compact</u>, on a

$$\overline{\ell}\ (A) = \varlimsup_{n \to +\infty} \frac{1}{n} \log P\ (\overline{X}_n \in A) \leqslant - \Lambda\ (\overline{A})$$

<u>Preuve</u> : Si $x \in \overset{\circ}{A}$ , il existe un voisinage ouvert convexe C de x tel que $C \subset \overset{\circ}{A}$ , et on a $\underline{\ell}\ (A)\ \geqslant \underline{\ell}\ (C) = \ell\ (C) \geqslant - \lambda\ (x)$

Prenons le sup. en $x \in \overset{\circ}{A}$ pour obtenir 4.7 (1).

Puisque $\overline{\ell}$ est monotone, on a $\overline{\ell}\ (A) \leqslant \overline{\ell}\ (\overline{A})$ ce qui permet de supposer A compact pour prouver (2). Soit $a > - \Lambda\ (A)$. Pour tout $x \in A$, on aura $a > - \lambda\ (x)$, et il existe donc un ouvert convexe $C_x$ contenant x tel que $a\ > \ell\ (C_x)$. Soit $C_{x_i}$ , $i = 1 \ldots k$ , un recouvrement fini de A. D'après 4.3 on sait que

$$\overline{\ell}\ (A) \leqslant \max_{i=1\ldots k} \ell\ (C_{x_i}) < a$$

L'arbitraire du choix de a donne alors $\overline{\ell}\ (A) \leqslant - \Lambda\ (A)$.

4.8. <u>Théorème</u> (Bahadur-Zabell [7] ). Sous les hypothèses 3.1, 3.2, pour toute union finie A d'ouverts convexes de E , on a

$$\lim_{n\to+\infty} \frac{1}{n} \log P\ (\overline{X}_n \in A) = \ell\ (A) = - \Lambda\ (A)$$

où $\Lambda$ est la fonctionnelle de Cramer associée à la loi $\mu$ commune des $X_n$.

<u>Preuve</u> : Par définition de $\Lambda$, on a toujours $\Lambda\ (A \cup B) = \inf\ [\Lambda\ (A),\ \Lambda\ (B)]$. Par 4.3, il suffit donc de prouver le théorème lorsque A est ouvert convexe. De 4.7 (1) on tire $- \Lambda\ (A) \leqslant\ \ell\ (A)$. Il suffit maintenant de prouver l'inégalité inverse lorsque $\ell\ (A)$ est <u>fini</u>.

Pour $\varepsilon > 0$ donné, on peut trouver N tel que

$$\frac{1}{n} \log \mu_n\ (A) \geqslant \ell\ (A) - \varepsilon \quad , \text{ pour } n \geqslant N$$

Il existe (lemme 3.8) un compact convexe K contenu dans A tel que

$$\frac{1}{N} \log \mu_N\ (A) - \frac{1}{N} \log \mu_N\ (K) \leqslant \varepsilon$$

Donc $f\ (n) = - \log \mu_n\ (K)$ vérifie

$$f\ (N) \leqslant - (\ell\ (A) - 2\ \varepsilon)\ N$$

Comme K est convexe, f est sous additive (voir 4.2) et donc $f(pN) \leqslant -(\ell(A) - 2\varepsilon)\, Np$ pour tout entier p. Ceci entraîne

$$- \ell(K) = \lim_{p \to +\infty} \frac{f(pN)}{pN} \leqslant -\ell(A) + 2\varepsilon$$

Mais 4.7 (2) fournit l'inégalité $\Lambda(K) \leqslant -\ell(K)$. Comparons les deux dernières inégalités avec l'inégalité triviale $\Lambda(A) \leqslant \Lambda(K)$ pour obtenir $\Lambda(A) \leqslant -\ell(A) + 2\varepsilon$, et donc comme $\varepsilon$ est quelconque, $\Lambda(A) \leqslant -\ell(A)$.

4.9. Proposition (d'après [7] ). Sous les hypothèses 3.1, 3.2 la transformée de Cramer $\lambda$ de $\mu$ est donnée par

$$\lambda(x) = \mathrm{Sup}\,\{-\ell(H) \mid x \in H\ ,\ H \text{ demi-espace ouvert}\}$$

Preuve : soit $a < \lambda(x)$. L'ensemble $C = \{y \mid \lambda(y) \leqslant a\}$ convexe fermé (car $\lambda$ est s.c.i. convexe) et ne contient pas x. Il existe donc un demi espace ouvert H de E tel que $H \cap C = \emptyset$ et $x \in H$. Puisque $H \cap C = \emptyset$ on a $\Lambda(H) \geqslant a$. Puisque $-\ell(H) = \Lambda(H)$, on a $-\ell(H) \geqslant a$ ; par suite $r = \mathrm{Sup}\,\{-\ell(H) / x \in H\ ,\ H \text{ demi-espace ouvert}\}$ est supérieur ou égal à a pour tout $a < \lambda(x)$, ce qui force $r \geqslant \lambda(x)$. Comme la définition 4.4 de $\lambda(x)$ fournit directement $r \leqslant \lambda(x)$, on conclut que $r = \lambda(x)$.

4.10. La proposition 4.9 ramène le calcul de $\lambda$ à celui de $\ell(H)$ pour H demi-espace ouvert, ce qui est un problème unidimensionnel, très proche de celui résolu par Cramer-Chernoff. Ceci permettra au § 5, d'obtenir $\lambda$ en dimension quelconque, à partir de $\hat{\mu}$.

La fonctionnelle $\Lambda$ est alors en principe calculable, et l'évaluation de $\ell(A) = \lim_{n \to +\infty} \frac{1}{n} \log P(\overline{X}_n \in A)$ repose sur 4.8 et les encadrements 4.7. La restriction "$\overline{A}$ compact" pour la validité de " $\overline{\ell}(A) \leqslant -\Lambda(\overline{A})$" est trop forte pour la plupart des applications ; nous étudions aux § 6 et 7 quelques situations essentielles où cette restriction est inutile.

## 5. TRANSFORMEE DE CRAMER ET TRANSFORMEE DE LAPLACE

5.1. Transformée de Laplace : soit $\mu$ une probabilité sur les boréliens de l'espace vectoriel topologique E. Supposons 3.1 et 3.2 vérifiées. On note

$$\hat{\mu}(t) = \int_E e^{<t,x>} d\mu(x) \qquad t \in E'$$

(où E' est le dual de E), la transformée de Laplace de $\mu$. Noter que $\hat{\mu}(t) = +\infty$ n'est pas exclu a priori. La fonction $\log \hat{\mu}(t)$ est définie sur E', à valeurs dans $[0, +\infty]$ , convexe et s.c.i. (E' étant muni de la topologie faible $\sigma(E', E)$.

5.2. Hypothèses sur $(E, \mu)$ : on supposera que le couple $(E, \mu)$ vérifie 3.1, 3.2 (hypothèse essentiellement topologique) et de plus que

$$\int_E |< t, x >| \, d\mu(x) \quad \text{est fini pour tout } t \in E'$$

Le théorème suivant étend au cas général la définition adoptée en dimension 1 pour $\lambda$ (cf. 2.2).

5.3. Théorème : (d'après [14] [7 ] [15])Sous l'hypothèse 5.2, la transformée de Cramer de $\mu$, définie en 4.4, est donnée par

$$\lambda(x) = \sup_{t \in E'} [< t, x > - \log \hat{\mu}(t)] \qquad \text{pour } x \in E$$

Preuve : Considérons d'abord le cas $\mu$ centrée et $E = \mathbb{R}$

Posons provisoirement

$$\varphi(x) = \sup_{t \in \mathbb{R}} [tx - \log \hat{\mu}(t)]$$

Soit $\lambda$ la transformée de Cramer de $\mu$ au sens de 4.4, de sorte que

$$(1) \qquad -\lambda(x) = \lim_{\varepsilon \to 0+} \ell(\,]x - 2\varepsilon, x + 2\varepsilon[\,) \quad , x \in \mathbb{R}$$

Comme $\mu$ est centrée, le théorème 2.3 (ii) entraîne, pour $x \geqslant 0$,

$\ell\,(]x-2\varepsilon,x+2\varepsilon[) \leqslant \overline{\lim}_{n\to+\infty} \frac{1}{n} \log P\,(\overline{X}_n \geqslant x-2\varepsilon) \leqslant -\varphi(x-2\varepsilon)$

d'où, grâce à (1),

(2) $\qquad \lambda\,(x) \geqslant \lim_{\varepsilon\to 0+} \varphi\,(x - 2\varepsilon) \qquad\qquad$ pour $x \geqslant 0$

D'autre part, le théorème 2.3 (i) donne pour $x \geqslant 0$,

$$- \varphi(x) \leqslant \underline{\lim}_{n\to+\infty} \frac{1}{n} \log P\,(\overline{X}_n \geqslant x) \leqslant \underline{\lim} \frac{1}{n} \log P\,(\overline{X}_n > x)$$

D'après le théorème 4.8, on a

$$\lim_{n\to+\infty} \frac{1}{n} \log P\,(\overline{X}_n > x) = - \Lambda\,(]x, +\infty[) = - \inf_{y>x} \lambda\,(y)$$

et par suite

(3) $\qquad \varphi(x) \geqslant \inf_{y>x} \lambda\,(y) \qquad\qquad$ pour $\quad x \geqslant 0.$

La loi des grands nombres et (1) montrent que $\lambda(0) = 0$. Comme $\lambda$ est convexe, s.c.i. et positive, on a donc $\inf_{y>x} \lambda\,(y) = \lim_{\varepsilon\to 0+} \lambda\,(x+\varepsilon)$ <u>pour</u> <u>$x \geqslant 0$</u> ; ainsi (3) devient

(4) $\qquad \varphi\,(x) \geqslant \lim_{\varepsilon\to 0+} \lambda\,(x + \varepsilon)$

Les fonctions $\lambda$ et $\varphi$ sont convexes et s.c.i. sur $[0, +\infty[$ , nulles en 0, à valeurs dans $[0, +\infty]$ . Il existe donc des nombres a, b $\in [0, +\infty]$ tels que $\lambda$, $\varphi$ soient <u>resp</u>. continues finies sur $[0, a[$ , $[0, b[$ , soient <u>resp</u>. identiques à $+\infty$ sur $]a, +\infty[$ , $]b, +\infty[$ , et soient <u>resp</u>. continues à gauche en a pour a fini, b pour b fini.

A partir des inégalités (2) et (4), on vérifie alors successivement que $a = b$ , puis que $\lambda$ et $\varphi$ coïncident sur $[0, +\infty[ - \{a\}$ , puis qu'elles coïncident sur $[0, +\infty[$. La comparaison de $\lambda$ et $\varphi$ sur $]-\infty, 0]$ se traite de façon symétrique ; on obtient donc $\lambda \equiv \varphi$. Le cas où $\mu$ n'est pas centrée se réduit au précédent par translation, et le théorème 5.3 est prouvé pour $E = \mathbb{R}$.

Passons au <u>cas général</u> ; soit $x \in E$. Tout demi-espace ouvert contenant $x$ s'écrit $H^{+}_{t,r} = \{y \in E \mid \langle t,y\rangle - r > 0\}$ avec

$$\langle t,x\rangle - r > 0 \ , \ t \in E', \ r \in \mathbb{R}$$

ou bien

$$H^{-}_{t,r} = \{y \in E \mid \langle t,y\rangle - r < 0\} \quad \text{avec}$$

$$\langle t,x\rangle - r < 0, \ t \in E' \ , \ r \in \mathbb{R}$$

Notons $t : E \to \mathbb{R}$ la forme linéaire associée à $t \in E'$.

Posons $Y_n = t\,(X_n)$ , $L^{+}_{r} = t\,(H^{+}_{t,r}) = ]r\,,+\infty[$ et

$L^{-}_{r} = t\,(H^{-}_{t,r}) = ]-\infty\,,r[$ , $u = t\,(x)$ , $\mu_t = t\,(\mu)$,

$\ell_t\,(A) = \lim\limits_{n\to+\infty} \dfrac{1}{n} \log P\,(\overline{Y}_n \in A) = \ell\,[t^{-1}\,(A)]$ où $A$ est un ouvert convexe quelconque de $\mathbb{R}$.

D'après 4.9, on a, en notant $\lambda_t$ la transformée de Cramer de $\mu_t$ ,

$$\lambda_t\,(u) = \max\left[\{\sup_{r<u} -\ell_t\,(L^{+}_{r})\}\ ,\ \{\sup_{r>u} -\ell_t\,(L^{-}_{r})\}\right]$$

d'où, comme $\ell_t\,(L^{\pm}_{r}) = \ell\,|t^{-1}\,(L^{\pm}_{r})| = \ell\,(H^{\pm}_{t,r})$,

$$\lambda_t(u) = \max\left[\{\sup_{r<t(x)} -\ell(H^{+}_{t,r})\}\ ,\ \{\sup_{r>t(x)} -\ell\,(H^{-}_{t,r}\}\right]$$

Lorsque $t$ varie dans $E'$ , le sup. des seconds membres de cette égalité est clairement égal à $\sup\,\{-\ell\,(H) \mid H \text{ demi espace ouvert, } H \text{ contient } x\}$ , et donc à $\lambda\,(x)$ d'après 4.9. On obtient donc

$$(5) \qquad \lambda(x) = \sup_{t\in E'} \lambda_t\,(\langle t, x\rangle) = \sup_{t\in E'} \lambda_{t(\mu)}\,[t(x)]$$

L'étude du cas $E = R$ permet d'écrire

$$\lambda_t\,[t\,(x)] = \varphi_t\,[t(x)] = \sup_{s\in\mathbb{R}}\,[s\,.\,t(x) - \log\hat{\mu}_t\,(s)]$$

Comme $\hat{\mu}_t\,(s) = \hat{\mu}\,(st)$ pour $s \in \mathbb{R}$ , $t \in E'$ , on a alors

$$\lambda\,(x) = \sup_{t\in E'}\ \sup_{s\in\mathbb{R}}\ [st\,(x) - \log\hat{\mu}\,(st)]$$

ce qui prouve $\lambda\,(x) = \sup\limits_{t\in E'}\ [t\,(x) - \log\hat{\mu}\,(t)]$.

5.4. Exemples de transformées de Cramer

(1) $E = \mathbb{R}^k$ , $\mu$ est gaussienne de moyenne M , de matrice de covariance $\sum$ , supposée inversible.

Alors $\lambda (x) = \frac{1}{2} (x - M)^* \sum^{-1} (x - M)$

(2) $E = \mathbb{R}^k$ , $d\mu (x) = f(x)\, dx$ avec $f (x) \sim \frac{cte}{|x|^r}$ lorsque $|x| \to + \infty$ . Alors $\hat{\mu} (t) = + \infty$ pour tout $t \neq 0$, et donc (th. 5.3) $\lambda \equiv 0$.

(3) Lemme (cas où $\lambda \equiv 0$) : Soit $\mu$ une probabilité sur un espace vectoriel topologique E et soit $X_n$ des v.a. indépendantes de même loi $\mu$. Supposons vérifiées les hypothèses 5.2. Les conditions suivantes sont alors équivalentes :

(i) la transformée de Cramer $\lambda$ de $\mu$ est identiquement nulle

(ii) la transformée de Laplace $\hat{\mu}$ (t) de $\mu$ vaut $+ \infty$ pour tout $t \neq 0$ dans E'.

(iii) pour toute partie borélienne A de E d'intérieur non vide, on a $\lim\limits_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A) = 0$.

Preuve : L'équivalence de (i) et (ii) résulte de 5.3 (5) et 2.7 (1). Celle de (i) et (iii) est une conséquence de la proposition 4.7.

La situation $\lambda \equiv 0$ exige donc une description plus fine du comportement asymptotique de $P (\overline{X}_n \in A)$, qui peut par exemple (lorsque $\int_E x\, d\mu (x) \notin \overline{A}$) tendre vers 0 à une vitesse polynomiale en $\frac{1}{n}$ (cf. travaux de l'école russe Linnik [31] Nagaev [38] et Altri).

La théorie à la Cramer-Chernoff n'est donc vraiment pertinente que lorsque $\lambda$ n'est pas identiquement nulle, donc lorsque $\hat{\mu}$ (t) est fini sur un ensemble assez "grand" en particulier lorsque $\hat{\mu}$ (t) est fini sur un voisinage de zéro dans E , ($\lambda$ ne s'annule alors qu'en $m = \int_E x\, d\mu (x)$).

(4) Nous donnons au § 9 en fin de chapitre quelques résultats décrivant l'allure des transformées de Cramer sur $\mathbb{R}^n$. Notre premier exemple de transformée de Cramer explicite en dimension infinie apparaîtra au § 8.

(5) Lemme : Soit $\mu$ une probabilité sur un e.v.t.l.c.s. E ; soit $S_\mu$ l'enveloppe convexe fermée du support de $\mu$. Alors si $(E, \mu)$ vérifie 3.2, la transformée de Cramer $\lambda$ de $\mu$ est infinie sur $(E - S_\mu)$.

Preuve : Si $x \notin S_\mu$, il existe un voisinage ouvert convexe W de x tel que $P(\overline{X}_n \in W) = 0$ pour tout n. Par suite $\ell(W) = -\infty$ et l'inégalité $\lambda(u) \geqslant -\ell(W)$ donne $\lambda(x) = +\infty$.

## 6. EXTENSION DE LA MAJORATION DE $\overline{\lim\limits_n} \frac{1}{n} \log P(\overline{X}_n \in A)$

Il s'agit d'étendre la majoration $\overline{\ell}(A) \leqslant -\Lambda(\overline{A})$, déjà prouvée pour $\overline{A}$ compact, au cas où A est borélien quelconque. Ce résultat est crucial dans les applications. En dimension infinie, il demande une sérieuse dépense d'énergie, ainsi que des hypothèses de décroissance "rapide" à l'infini pour la queue de la mesure $\mu$. Ce paragraphe s'inspire des élégants résultats de Donsker-Varadhan [15] sur le cas où E est un Banach, et a aussi bénéficié d'un exposé de Bretagnolle [47] . L'originalité de notre présentation consiste en particulier à court-circuiter le beau théorème de Sanov (cf. th. 8.2), qui peut alors être considéré pour ce qu'il est, c'est-à-dire un corollaire de la théorie à la Cramer-Chernoff.

C'est un peu ce qu'ont tenté de faire Bahadur et Zabell [7] , mais au prix de l'utilisation, en cours de route de résultats de Donsker-Varadhan [15] , qui eux-mêmes n'étaient obtenus qu'après démonstration du théorème de Sanov !

6.1. <u>Lemme</u> : Soit $\mu$ une probabilité sur un e.v.t. E, et soient $X_n$ une suite de v.a. indépendantes, de loi $\mu$, à valeurs dans E. Supposons vérifiées les hypothèses 3.1, 3.2. Supposons de plus que pour tout $a > 0$, il existe un compact $K_a \subset E$ tel que $P(\overline{X}_n \notin K_a) \leq e^{-na}$ pour tout $n \geq N(a)$. Alors quel que soit $B \in \mathcal{B}(E)$, on a

$$\overline{\ell}(B) = \overline{\lim_{n\to+\infty}} \frac{1}{n} \log P(\overline{X}_n \in B) \leq -\Lambda(\overline{B}).$$

De plus si $\lambda$ est la transformée de Cramer de $\mu$, l'ensemble $\{x \in E \mid \lambda(x) \leq b\}$ est un convexe <u>compact</u> quel que soit b positif fini.

<u>Preuve</u> : il est clair que $\overline{\ell}(E - K_a) \leq -a$. Par suite on a

$$\overline{\ell}(B) \leq \max \{\overline{\ell}(B \cap K_a), \overline{\ell}(B \cap (E - K_a))\} \leq \max \{\overline{\ell}(\overline{B} \cap K_a), -a\}$$

Comme $\overline{B} \cap K_a$ est compact, 4.7 donne

$$\overline{\ell}(\overline{B} \cap K_a) \leq -\Lambda(\overline{B} \cap K_a) \leq -\Lambda(\overline{B})$$

Ainsi $\overline{\ell}(B) \leq \max \{-\Lambda(\overline{B}), -a\}$ pour tout $a > 0$, ce qui prouve

$$\overline{\ell}(B) \leq -\Lambda(\overline{B}).$$

Soit b positif fini. Fixons $a > b$. L'ensemble $(E - K_a)$ est ouvert, donc 4.7 entraîne

$$-\Lambda(E - K_a) \leq \underline{\lim} \frac{1}{n} \log P\{\overline{X}_n \in (E - K_a)\}$$

ce qui force (par définition de $K_a$) l'inégalité $-\Lambda(E - K_a) \leq -a$, ou encore $\Lambda(E - K_a) \geq a$.

En particulier, on a $\lambda(x) \geq a$ dès que $x \notin K_a$, ce qui entraîne l'inclusion de $L_b = \{x \in E \mid \lambda(x) \leq b\}$ dans $K_a$. Mais $L_b$ est fermé (car $\lambda$ est s.c.i.) et $K_a$ est compact. Donc $L_b$ est compact (et d'ailleurs convexe car $\lambda$ est convexe).

6.2.<u>Corollaire</u> : Soit E un e.v.t.l.c.s. et soit $\mu$ une probabilité sur $\mathcal{B}(E)$ portée par un compact convexe métrisable de E. Alors, si $\Lambda$ est la fonctionnelle de Cramer associée à $\mu$, on a pour tout $A \in \mathcal{B}(E)$

$$- \Lambda (\mathring{A}) \leqslant \varliminf_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A) \leqslant \varlimsup_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A) \leqslant - \Lambda (\overline{A})$$

(notations 3.1, 4.1). De plus $\{x \in E \mid \lambda (x) \leqslant b\}$ est compact pour tout b positif fini.

<u>Remarque</u> : Ce corollaire fournit déjà tous les ingrédients du théorème de Sanov sur les grandes déviations des lois empiriques (cf. § 8) en prenant pour E l'espace des mesures bornées sur un espace topologique compact métrisable $\Gamma$ et pour $\mu$ la loi de $\delta_X$ où X est une v.a. à valeurs dans $\Gamma$.

6.3. <u>Proposition</u> : Soit E le dual d'un espace de Banach séparable F. Supposons que $\int_E e^{s||x||} d\mu (x)$ soit fini pour tout s appartenant à un voisinage de 0 dans $\mathbb{R}$ ; $\mu$ est ici une probabilité quelconque sur $\mathcal{B}(E)$. Alors, si $\Lambda$ est la fonctionnelle de Cramer associée à $\mu$, on a (notations 3.1, 4.1) pour tout $A \in \mathcal{B}(E)$

$$- \Lambda (\mathring{A}) \leqslant \varliminf_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A) \leqslant \varlimsup_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A \leqslant - \Lambda (\overline{A})$$

où $\mathring{A}$ , $\overline{A}$ sont resp. l'intérieur et l'adhérence de A <u>pour la topologie faible $\sigma$ (E, F) sur E</u>. De plus, $\{x \in E \mid \lambda(x) \leqslant b\}$ est un compact (faible) pour tout b positif fini.

<u>Remarque préliminaire</u> : Lorsque E = F' , où F est un Banach séparable, on peut munir E de la topologie de la norme notée $\tau$ (E, F) ou de la topologie faible $\sigma$ (E, F). Comme $\tau$ est une topologie polonaise plus fine que $\sigma$, <u>la tribu borélienne</u> $\mathcal{B}(E)$ <u>est la même pour toute topologie comprise entre $\tau$ et $\sigma$</u> . La transformée de Cramer $\lambda$ et la fonctionnelle de Cramer $\Lambda$ ne dépendent que des ouverts convexes de E d'après les définitions 4.4 et 4.6 donc

$\lambda$ (et $\Lambda$) sont les mêmes pour les topologies $\tau$ et $\sigma$ . Notons enfin que le dual E' de (E, $\sigma$) est égal à F.

Preuve de 6.3

Les inégalités annoncées se déduiront de 4.7 et 6.1 après construction des compacts $K_a$ introduits en 6.1. Soit $B_a$ la boule fermée de rayon a dans E. On sait que $B_a$ est compacte pour la topologie $\sigma$. De plus, d'après le théorème 2.3 (ii)

$$P\,(\overline{X}_n \notin B_a) = P\,(||\overline{X}_n|| > a) \leqslant P\,\left\{\frac{1}{n}\,(||X_1|| + \ldots + ||X_n||) > a\right\}$$

$$\leqslant e^{-n\,\varphi(a)}$$

où $\varphi$ est la transformée de Cramer de $\nu$ = loi $(||X_n||)$.
Par hypothèse $\hat{\nu}$ (s) est fini au voisinage de 0, et donc (proposition 2.7 (2)) on a $\lim\limits_{a\to+\infty} \varphi\,(a) = +\infty$ . On peut donc appliquer le lemme 6.1 pour prouver la proposition 6.3.

6.4. Corollaire : soit $E = \mathbb{R}^k$ et $\mu$ une probabilité sur $\mathcal{B}(E)$ telle que $\hat{\mu}$ (t) soit finie pour tout t dans un voisinage de 0. Alors pour tout $A \in \mathcal{B}(E)$ , on a

$$-\,\Lambda\,(\mathring{A}) \leqslant \varliminf_{n\to+\infty} \frac{1}{n} \log P\,(\overline{X}_n \in A) \leqslant \varlimsup_{n\to+\infty} \frac{1}{n} \log P\,(\overline{X}_n \in A) \leqslant -\,\Lambda\,(\overline{A})$$

De plus $\lambda\,(x) \to +\infty$ quand x tend vers l'infini dans $\mathbb{R}^k$.

6.5. Sur l'égalité $\ell(A) = -\,\Lambda\,(A)$ : Dans les situations décrites par les corollaires 6.2 et 6.4, on voit que la relation $\Lambda\,(\mathring{A}) = \Lambda\,(\overline{A})$ (qui force d'ailleurs $\Lambda\,(\mathring{A}) = \Lambda\,(\overline{A}) = \Lambda\,(A)$) suffit à impliquer l'existence de $\lim\limits_{n\to+\infty} \frac{1}{n} \log P\,(\overline{X}_n \in A) = \ell\,(A)$ et la relation $\ell\,(A) = -\,\Lambda\,(A)$

Remarquons cependant que la réciproque est fausse. Prenons par exemple $E = \mathbb{R}$ , $\mu = p\,\delta_u + (1-p)\,\delta_v$ avec $u < v$ , $0 < p < 1$ , et $A = \,]v, +\infty[$

On a $\ell$ (A) = $-\Lambda$ (A) car A est ouvert convexe, tandis que $\Lambda$ (A) = $\Lambda$ $(\mathring{A})$ = $+\infty$ et $\Lambda$ $(\overline{A})$ = $\lambda$ (v) = $-\log$ (1-p).

En général, la classe des ensembles A tels que $\ell$ (A) existe est stable par union finie, contient les ouverts convexes et les ensembles A tels que $\Lambda$ $(\mathring{A})$ = $\Lambda$ $(\overline{A})$. Pour le cas de E = $\mathbb{R}^k$ , nous renvoyons à Azencott-Ruget [3] où des résultats plus précis sont donnés sur cette classe d'ensembles. Dans ce cours, nous nous limiterons à prouver que pour E = $\mathbb{R}^k$ , $\lim\limits_{n\to+\infty} \frac{1}{n} \log P(\overline{X}_n \in A)$ existe dès que A est ouvert (cf. proposition 9.8).

6.6. <u>Sur les changements de topologie</u> : soit E un Banach réflexif, séparable (par exemple un Hilbert séparable), de dual E'. Soit $\mu$ une probabilité sur $\mathcal{B}$(E) telle que $\int |< t, x >| \, d\mu\,(x)$ soit fini pour t $\in$ E'.
Pour toute topologie sur E comprise entre la topologie de la norme $\tau$ (E, E') et la topologie faible $\sigma$ (E, E'), les boréliens $\mathcal{B}$(E) restent les mêmes, le dual E' ne change pas, la fonction $\lambda$ et la fonctionnelle $\Lambda$ restent les mêmes (cf. théorème 5.3). Mais l'intérieur d'un même ensemble B $\in$ $\mathcal{B}$ (E) est plus grand pour $\tau$ que pour $\sigma$ tandis que son adhérence est plus petite pour $\tau$ que pour $\sigma$. <u>Les bornes</u> $-\Lambda$ $(\mathring{B})$ <u>et</u> $-\Lambda$ $(\overline{B})$ <u>s'améliorent donc lorsqu'on enrichit la topologie de $\sigma$ à $\tau$</u> . Le résultat 6.3 par exemple ne permet d'utiliser ces bornes que pour $\sigma$, au moins en ce qui concerne $-\Lambda$ $(\overline{B})$. Nous allons donc renforcer l'hypothèse sur $\mu$ pour permettre d'utiliser la topologie forte $\tau$.

7. MAJORATION DE $\lim \frac{1}{n} \log P(\overline{X}_n \in A)$ ; CAS DES ESPACES DE BANACH

7.0. Voir le début du § 6 en ce qui concerne les sources de ce paragraphe, essentiellement [7] [15] [16] . Dans tout le § 7, E est un Banach séparable, $\mu$ une probabilité sur $\mathcal{B}(E)$, $X_n : \Omega \to E$ une suite de v.a. indépendantes de loi $\mu$.

7.1. Convergence étroite : soit $\Gamma$ un espace topologique polonais (sans structure vectorielle !), et soit $\mathcal{B}(\Gamma)$ la tribu borélienne de $\Gamma$. Soient $\mathcal{M}(\Gamma)$ l'espace des mesures bornées sur $\mathcal{B}(\Gamma)$, $\mathcal{M}_+(\Gamma)$ le cône des mesures positives bornées et $\mathcal{M}_1(\Gamma)$ le convexe des probabilités sur $\mathcal{B}(\Gamma)$. On munit $\mathcal{M}(\Gamma)$ de la topologie de la convergence étroite (la topologie la moins fine rendant continue les applications $\Theta \to \int_E f \, d\Theta$ où f est une fonction continue bornée quelconque sur $\Gamma$).

Une partie C de $\mathcal{M}(\Gamma)$ est relativement compacte si et seulement si pour tout $\varepsilon > 0$ il existe un compact $K_\varepsilon$ de $\Gamma$ tel que $(\Gamma - K_\varepsilon)$ soit de mesure inférieure à $\varepsilon$ pour $\Theta^+$ et $\Theta^-$ quelle que soit $\Theta = \Theta^+ - \Theta^-$ dans C.

7.2. Barycentres : considérons le Banach séparable E. Si $\Theta \in \mathcal{M}_1(E)$, et si $\int_E ||x|| \, d\Theta(x)$ est fini, il existe un unique point $b(\Theta) \in E$ tel que

$$< t, b(\Theta) > = \int_E < t, x > d\Theta(x) \qquad \text{pour tout } t \in E'$$

et $b(\Theta)$ est appelé barycentre de $\Theta$.

(1) Si $\Theta(C) = 1$ ou C est un convexe fermé de E, alors on a $b(\Theta) \in C$.

(2) Soit $\Theta_n \in \mathcal{M}_1(E)$ une suite convergeant vers $\Theta \in \mathcal{M}_1(E)$. Si B est une boule fermée telle que $\Theta_n(B) = 1$ pour tout n, alors $b(\Theta_n)$ tend vers $b(\Theta)$ quand $n \to +\infty$.

7.3. Lemme : Soit $f : \mathbb{R}^+ \to \mathbb{R}^+$ une fonction continue positive telle que $\lim\limits_{t\to+\infty} \frac{f(t)}{t} = +\infty$. L'ensemble

$$G(f, a) = \left\{ \Theta \in \mathcal{M}_1(E) \mid \int_E f(||x||)\, d\mu(x) \leqslant a \right\}$$

est un fermé de $\mathcal{M}(E)$. L'application $\Theta \to b(\Theta)$ est définie et continue sur $G(f, a)$.

Preuve : le lemme de Fatou montre que $G(f, a)$ est fermé. D'autre part, pour $\Theta \in G(f, a)$, on a quel que soit $r > 0$,

$$||\int_{||x||\geqslant r} x\, d\Theta(x)|| \leqslant \int_{||x||\geqslant r} ||x||\, d\Theta(x)$$

$$\leqslant \varepsilon(r) \int_{||x||\geqslant r} f(||x||)\, d\Theta(x) \leqslant a\, \varepsilon(r)$$

avec $\varepsilon(r) = \sup\limits_{||x||\geqslant r} \left[ \frac{||x||}{f(||x||)} \right]$, de sorte que $\lim\limits_{r\to+\infty} \varepsilon(r) = 0$.

Il suffit d'appliquer 7.2 (2) pour conclure que $\Theta \to b(\Theta)$ est une application continue sur $G(f, a)$.

7.4. Lemme (d'après Donsker-Varadhan [15] ) : soit $\Gamma$ un espace topologique polonais. Soit $X_n$ une suite de v.a. indépendantes, à valeurs dans $\Gamma$, de même loi $\mu$. Soit $Z_n$ la loi empirique de l'échantillon $X_1 \dots X_n$, de sorte que $Z_n = \frac{1}{n} \delta_{X_1} + \dots + \frac{1}{n} \delta_{X_n}$ est une v.a. à valeurs dans $\mathcal{M}_1(\Gamma)$. Alors pour tout $a > 0$, il existe un compact $C_a$ de $\mathcal{M}_1(\Gamma)$ et un entier $N(a)$, tels que $n \geqslant N(a)$ entraîne $P(Z_n \notin C_a) \leqslant e^{-an}$.

Preuve : Soient $b_p$ et $\varepsilon_p$ deux suites de nombres positifs tendant vers 0. Pour chaque p choisissons un compact $K_p$ de $\Gamma$ tel que $\mu(\Gamma - K_p) \leqslant b_p$. Soit $G_p$ l'ensemble des $\Theta \in \mathcal{M}_1(\Gamma)$ telles que $\Theta(\Gamma - K_p) \leqslant \varepsilon_p$. Montrons que

(1) $$P(Z_n \notin G_p) \leqslant \left(\frac{b_p}{\varepsilon_p}\right)^{1/2\, n\, \varepsilon_p} \quad \text{pour tout } n, p.$$

Supprimons un instant l'indice p pour alléger les notations ;

$$P(Z_n \notin G) = P\left\{ \frac{1}{n} \left[ 1_{K^c}(X_1) + \ldots + 1_{K^c}(X_n) \right] > \varepsilon \right\} = P(\overline{U}_n > \varepsilon)$$

où $U_i = 1_{K^c}(X_i)$ est une v.a. de loi binomiale $\nu = \alpha\, \delta_1 + (1-\alpha)\, \delta_o$ ,

avec $\alpha = \mu(K^c) \leqslant b$.

D'après 2.3 (ii), on a donc, en notant $\lambda$ la transformée de Cramer de $\nu$ ,

(2) $$P(Z_n \notin G) \leqslant e^{-n\lambda(\varepsilon)}$$

D'après 2.6 (2), pour $\varepsilon \in \,]0, 1[$ , on sait que

$$\lambda(\varepsilon) = \varepsilon \log \frac{\varepsilon}{\alpha} + (1-\varepsilon) \log \frac{1-\varepsilon}{1-\alpha}$$

Prenons $b_p < \varepsilon_p$ , de sorte que $\alpha \leqslant b_p < \varepsilon_p$ , et $\varepsilon_p \leqslant \frac{1}{2}$ , ce qui entraîne $\lambda(\varepsilon_p) \geqslant \varepsilon_p \log \frac{\varepsilon_p}{b_p} + \frac{1}{2} \log \frac{1}{2}$

Imposons $\frac{1}{2}\, \varepsilon_p \log \frac{\varepsilon_p}{b_p} \geqslant \frac{1}{2} \log \frac{1}{2}$ pour obtenir

$\lambda(\varepsilon_p) \geqslant \frac{1}{2}\, \varepsilon_p \log \frac{\varepsilon_p}{b_p}$ , ce qui, grâce à (2), démontre (1).

Soit maintenant $C = \bigcap_{p \geqslant 1} G_p$ , de sorte que C est un compact de $\mathcal{M}_1(\Gamma)$.
On aura, d'après (1),

$$P(Z_n \notin C) \leqslant \sum_{p \geqslant 1} P(Z_n \notin G_p) \leqslant \sum_{p \geqslant 1} \left(\frac{b_p}{\varepsilon_p}\right)^{1/2\; n\varepsilon_p}$$

Il suffit de poser $b_p = \varepsilon_p\, u^{2p/\varepsilon_p}$ avec $0 < u < 1/2$ pour garantir

$P(Z_n \notin C) \leqslant \sum_{p \geqslant 1} u^{np} \leqslant 2\, u^n$ , ce qui prouve le lemme.

7.5. <u>Théorème</u> (d'après Donsker-Varadhan [15] ). Soit E un Banach séparable et $\mu$ une probabilité sur E telle que $\int_E e^{s\|x\|}\, d\mu(x)$ soit fini pour tout $s \in \mathbb{R}$.

Alors pour toute partie borélienne A de E on a

$$- \Lambda (\mathring{A}) \leq \underline{\lim}_{n\to+\infty} \frac{1}{n} \log P (\overline{X}_n \in A) \leq \overline{\lim}_{n\to+\infty} \frac{1}{n} \log P (X_n \in A) \leq - \Lambda (\overline{A})$$

où $\Lambda$ est la fonctionnelle de Cramer associée à $\mu$. De plus, $\{x \in E \quad | \lambda (x) \leq b\}$ est compact pour tout b positif fini.

Preuve : soit $\nu$ la loi commune des $||X_n||$. Posons $m = \int_{\mathbb{R}} x \, d\nu (x)$. Montrons l'existence d'une fonction f continue convexe sur $[0, +\infty[$ telle que

$$(1) \quad \begin{cases} f (x) = 0 \quad \text{pour} \quad 0 \leq x \leq m \\ \lim_{x\to+\infty} \frac{f(x)}{x} = + \infty \\ \int_0^{+\infty} e^{sf(x)} \, d\nu (x) \quad \text{est fini pour } s < 1 \end{cases}$$

Comme $\hat{\nu}$ est finie en tout point de $\mathbb{R}$ par hypothèse, il suffit d'après la proposition 2.7 de prendre $f(x) = \lambda_\nu (x)$ pour $x \in [m, +\infty[$, où $\lambda_\nu$ est la transformée de Cramer de $\nu$, et $f (x) = 0$ pour $x \in [0, m]$.

Gardons les notations de 7.3, 7.4. On a

$$P (Z_n \notin G (f,a)) = P \left[ \frac{1}{n} (f (||X_1||) + \ldots + f (||X_n||)) > a \right] \leq e^{-n \lambda_\Theta(a)}$$

où $\lambda_\Theta$ est la transformée de Cramer de la loi $\Theta$ de $f (||X_1||)$. D'après (1), $\hat{\Theta} (s)$ est fini au voisinage de 0 , et donc (proposition 2.7) $\lambda_\Theta (a)$ tend vers $+ \infty$ quand a tend vers $+ \infty$ . Pour tout $A > 0$, il existe donc $a_A$ tel que $P (Z_n \notin G (f, a_A)) \leq e^{-nA}$ .

Soit alors $K_A = C_A \cap G (f, a_A)$ , où $C_A$ est le compact de $\mathcal{M}_1 (E)$ défini en 7.4. D'après 7.4 et l'inégalité précédente, on a $P (Z_n \notin K_a) \leq 2 e^{-nA}$ pour $n \geq N (A)$. De plus (cf. 7.3) $K_A$ est un compact de $\mathcal{M}_1 (E)$ sur lequel l'application barycentre : $\pi \to b (\pi)$ est définie et continue. Par suite l'image $L_A$ de $K_A$ par b est un compact de E et

$P(\overline{X}_n \notin L_A) = P(b(Z_n) \notin b(K_A)) \leqslant P(Z_n \notin K_A) \leqslant 2e^{-nA}$

pour $n \geqslant N(A)$. Une application de 6.1 et 4.7 achève la preuve.

## 8. GRANDES DEVIATIONS DES LOIS EMPIRIQUES. THEOREME DE SANOV

8.1. Information de Kullback : soit $\Gamma$ un espace topologique polonais. Soient $\pi$ et $\nu$ deux probabilités sur la tribu borélienne $\mathcal{B}(\Gamma)$ de $\Gamma$. Définissons l'information de $\nu$ par rapport à $\pi$ (cf. Kullback [28] ) par

$$I_\pi(\nu) = +\infty \quad \text{si } \nu \text{ n'est pas absolument continue par rapport à } \pi$$

$$I_\pi(\nu) = \int_\Gamma f(x) \log f(x) \, d\pi(x) \ , \text{ avec } f(x) = \frac{d\nu}{d\pi}(x)$$

si $\nu$ est absolument continue par rapport à $\pi$. On définit ainsi une fonctionnelle $I_\pi : \mathcal{M}_1(\Gamma) \longrightarrow [0, +\infty]$, où $\mathcal{M}_1(\Gamma)$ est l'espace des probabilités sur $\mathcal{B}(\Gamma)$. Dans la suite $\mathcal{M}_1(\Gamma)$ est muni de la topologie de la convergence étroite (cf. 7.1). Pour les propriétés élémentaires de $I_\pi$ nous renvoyons à Kullback [28] . Notons que d'un point de vue heuristique, $I_\pi(\nu)$ est d'autant plus grand que $\nu$ est plus "éloignée" de $\pi$. En particulier, on a $I_\pi(\nu) = 0$ si et seulement si $\nu = \pi$.

8.2. Théorème (Sanov [43] Donsker-Varadhan [15] Bahadur-Zabell [7] )

Soit $\pi$ une probabilité sur un espace topologique polonais $\Gamma$ (noter que $\Gamma$ n'est pas supposé être un espace vectoriel !). Soit $X_n$ une suite de v.a. indépendantes, de même loi $\pi$, à valeurs dans $\Gamma$. Soit
$Z_n = \frac{1}{n}\delta_{X_1} + \ldots + \frac{1}{n}\delta_{X_n}$ la loi empirique de l'échantillon $X_1 \ldots X_n$.
Pour toute partie borélienne $A$ de $\mathcal{M}_1(\Gamma)$ on a

$$-I_\pi(\mathring{A}) \leqslant \varliminf_{n\to+\infty} \frac{1}{n} \log P(Z_n \in A) \leqslant \varlimsup_{n\to+\infty} \frac{1}{n} \log P(Z_n \in A) \leqslant -I_\pi(\overline{A})$$

où $I_\pi(A) = \inf_{\nu \in A} I_\pi(\nu)$.

Pour toute union finie A d'ouverts convexes de $\mathcal{M}_1$ $(\Gamma)$, on a

$$- I_\pi (A) = \lim_{n\to+\infty} \frac{1}{n} \log P (Z_n \in A)$$

Enfin, si $\mu$ est la loi de $\delta_{X_1}$ , (considérée comme probabilité sur l'espace vectoriel $E = \mathcal{M}(\Gamma)$) la transformée de Cramer $\lambda$ de $\mu$ coïncide avec $I_\pi$ sur $\mathcal{M}_1$ $(\Gamma)$ et vaut $+ \infty$ sur $(\mathcal{M}(\Gamma) - \mathcal{M}_1 (\Gamma))$. De plus $\{x \in E \mid \lambda(x) \leqslant b\}$ est un compact pour tout b positif fini.

<u>Preuve</u> : l'espace vectoriel $E = \mathcal{M}(\Gamma)$ des mesures bornées sur $\Gamma$ est muni de la convergence étroite (cf. 7.1). Les v.a. $Y_n = \delta_{X_n}$ sont indépendantes, à valeurs dans E, de même loi $\mu$, portée par le convexe fermé $F = \mathcal{M}_1$ $(\Gamma)$. De plus $Z_n = \overline{Y}_n$. Les hypothèses topologiques 3.2 sont vérifiées pour le couple (E, $\mu$) comme on l'a vu en 3.5.

D'après le lemme 7.4, il existe pour tout $a > 0$ un compact $C_a$ de F tel que $P (\overline{Y}_n \notin C_a) \leqslant e^{-na}$. Le lemme 6.1 fournit alors la majoration de $\frac{1}{n} \overline{\lim_{n\to+\infty}} \log P (Z_n \in A)$ par $- \Lambda (\overline{A})$ ; et la proposition 4.7 donne la minoration analogue par $- \Lambda (\mathring{A})$ , où $\Lambda$ est la fonctionnelle de Cramer associée à $\mu$. Pour prouver le théorème de Sanov, il suffit donc de montrer que la transformée de Cramer $\lambda$ de $\mu$ coïncide avec $I_\pi$ sur $\mathcal{M}_1$ $(\Gamma)$ et vaut $+ \infty$ sur le complémentaire de $\mathcal{M}_1$ $(\Gamma)$.

Ce dernier point se déduit de 5.5, car $\mathcal{M}_1$ $(\Gamma)$ contient l'enveloppe convexe fermée de support de $\mu$. Reste à comparer $\lambda$ et $I_\pi$ sur $\mathcal{M}_1$ $(\Gamma)$. Pour $\nu \in \mathcal{M}_1$ $(\Gamma)$, on a d'après 5.3, et puisque $E' = C (\Gamma)$ (i.e. fonctions continues bornées)

$$\lambda (\nu) = \sup_{f \in C(\Gamma)} \left[< f, \nu > - \log \hat{\mu} (f)\right]$$

Par construction, on a pour $f \in C (\Gamma)$

$$\hat{\mu} (f) = \int_{\mathcal{M}_1(\Gamma)} \exp (< f, \tau >) \, d\mu (\tau) = \int_\Gamma e^{f(x)} \, d\pi (x)$$

d'où

$$\lambda(\nu) = \sup_{f \in C(\Gamma)} \left[ \int_\Gamma f(x)\, d\nu(x) - \log \int_\Gamma e^{f(x)}\, d\pi(x) \right]$$

Supposons d'abord $I_\pi(\nu)$ fini. Alors $\nu$ est absolument continue par rapport à $\pi$ ; soit $g = \frac{d\nu}{d\pi}$ . L'inégalité de Jensen donne pour $f \in C(\Gamma)$

$$\exp \left[ \int_\Gamma (f - \log(g))\, 1_{\{g>0\}}\, g\, d\pi \right] \leqslant \int_\Gamma \exp(f - \log(g))\, 1_{\{g>0\}}\, g\, d\pi$$

d'où

$$\int_\Gamma f\, g\, d\pi - \int_\Gamma g \log(g)\, 1_{\{g>0\}}\, d\pi \leqslant \log \int_\Gamma e^f\, d\pi$$

c'est-à-dire

$$\int_\Gamma f\, d\nu - \log \int_\Gamma e^f\, d\pi \leqslant \int_\Gamma g \log(g)\, d\pi = I_\pi(\nu)$$

Comme le sup du premier membre, quand $f$ décrit $C(\Gamma)$, vaut $\lambda(\nu)$, on obtient $\lambda(\nu) \leqslant I_\pi(\nu)$, inégalité qui reste trivialement vraie si $I_\pi(\nu) = +\infty$

Inversement, supposons que $\lambda(\nu) < +\infty$ , de sorte que

$$(1) \qquad \int_\Gamma f\, d\nu - \log \int_\Gamma e^f\, d\pi \leqslant \lambda(\nu) < +\infty \text{ , pour } f \in C(\Gamma)$$

Soit $\varphi$ une fonction mesurable bornée. Sa restriction à des compacts bien choisis, de $\nu$-mesure et $\pi$-mesure arbitrairement proches de 1, est continue (th. de Lusin) ; par passage à la limite on voit donc que (1) reste valable pour $f = \varphi$ .

Appliquons alors (1) à $f = n\, 1_A$ où $A$ est une partie $\pi$-négligeable de $\Gamma$, pour obtenir $n\nu(A) \leqslant \lambda(\nu)$ pour tout $n$ , ce qui entraîne $\nu(A) = 0$. Ainsi $\nu$ est absolument continue par rapport à $\pi$. Posons $g = \frac{d\nu}{d\pi}$ .

Si $\log g$ était bornée, en appliquant (1) à $\log(g)$ on aurait

$$\int_\Gamma \log(g)\, d\nu - \log \int_\Gamma d\pi \leqslant \lambda(\nu)$$

c'est-à-dire exactement $I_\pi(\nu) \leqslant \lambda(\nu)$. Quand $\log(g)$ n'est pas bornée on approche (péniblement !) $\log(g)$ par des fonctions mesurables bornées pour obtenir le même résultat (voir Donsker-Varadhan [15]) ce qui donne finalement $I_\pi \equiv \lambda$ sur $\mathcal{M}_1(\Gamma)$.

8.3. Aspects des transformées de Cramer en dimension infinie

(1) Nous venons de calculer $\lambda = I_\pi$, la transformée de Cramer de $\mu$ = loi de $\delta_{X_n}$, probabilité sur $E = \mathcal{M}(\Gamma)$ portée par le fermé $\mathcal{M}_1(\Gamma)$. En particulier (cf. 8.1), on aura $\lambda(\nu) = +\infty$ si $\nu \in \mathcal{M}_1(\Gamma)$ et n'est pas absolument continue par rapport à $\pi$ = loi de $X_n$. L'ensemble des $\nu \in E$ tels que $\lambda(\nu) = +\infty$ est donc dense dans $E$.

Le même argument montre que si $S_\mu$ est l'enveloppe convexe fermée du support de $\mu$, l'ensemble des $\nu \in S_\mu$ telles que $\lambda(\nu) = +\infty$ est encore dense dans $S_\mu$.

Les deux propriétés précédentes seront encore vraies dans tous les exemples explicites de transformées de Cramer en dimension infinie étudiés dans la suite. Par contre elles sont fausses en dimension finie (cf. § 9).

(2) Dans tous les cas décrits par les théorèmes 6.2, 6.3, 6.4, 7.5, 8.2 la transformée de Cramer $\lambda$ de $\mu$ est telle que $\{x \in E \mid \lambda(x) \leq b\}$ soit compact pour tout b positif fini. Ce résultat, chaque fois obtenu par application du lemme 6.1 (existence de compacts $K_a$ tels que $P(\overline{X}_n \notin K_a) \leq e^{-na}$) est utile pour certaines applications.

En effet, les résultats à la Cramer-Chernoff se transportent formellement assez bien par des applications $f : E \to G$ "à peu près" continues ; on en verra plusieurs exemples dans la suite de notre exposé. Mais le passage du formel au rigoureux est fortement lié à la propriété de compacité (2) de $\lambda$. Le cas où f est linéaire est abordé dans Bahadur-Zabell [7] ;

nous renvoyons aux chapitres sur les processus gaussiens et les diffusions pour des applications où f n'est pas linéaire.

## 9. BOTANIQUE DES TRANSFORMEES DE CRAMER SUR $\mathbb{R}^k$

Il est suggéré au lecteur méticuleux de ne pas lire les démonstrations un peu techniques des résultats du § 9, qui utilisent en grande partie un texte de Rockafellar [42] sur les fonctions convexes en dualité. Le côté descriptif des résultats est par contre utile pour avoir une idée de l'allure et du calcul des transformées de Cramer sur $\mathbb{R}^k$.

9.1. Hypothèses : dans tout ce paragraphe, on suppose que $E = \mathbb{R}^k$ et que $\mu$ est une probabilité sur E telle que $\int_E |x| \, d\mu(x)$ soit fini. On suppose (sans perdre de généralité en fait) qu'aucun sous espace affine strict de $\mathbb{R}^k$ ne porte $\mu$. Dans les démonstrations, on prendra toujours $\int_E x \, d\mu(x) = 0$.

9.2. Fonctions convexes en dualité : considérons la famille $\mathcal{F}$ des fonctions $m : \mathbb{R}^k \to ]-\infty, +\infty]$ , qui sont convexes, s.c.i. et finies en au moins un point de $\mathbb{R}^k$. A toute fonction $m \in \mathcal{F}$ on peut associer $m^* \in \mathcal{F}$ par

(1) $$m^*(x) = \sup_{t \in \mathbb{R}^k} [\langle t, x \rangle - m(t)] \qquad x \in \mathbb{R}^k$$

et on a toujours

(2) $$(m^*)^* = m$$

Nous renvoyons à Rockafellar [42] pour une preuve de ces résultats (dans la terminologie de [42] , les fonctions de $\mathcal{F}$ sont appelées "proper, closed, convex functions").

Dans la littérature, la dualité $m \longleftrightarrow m^*$ est parfois appelée dualité de Fenchel, ou dualité de Young. Elle est aussi très proche de la dualité de Legendre, que nous rappelons ci-dessous.

Pour revenir au contexte qui nous concerne ici, sous l'hypothèse 9.1, la fonction m (t) = log $\hat{\mu}$ (t) est dans $\mathcal{F}$ , et d'après le théorème 5.3, la transformée de Cramer $\lambda$ de $\mu$ s'écrit

(3) $$\lambda = (\log \hat{\mu})^*$$

D'après (2), on voit que <u>la correspondance entre $\lambda$ et $\mu$ est biunivoque</u>.

Pour toute fonction $f \in \mathcal{F}$ posons

(4) $$D_f = \{t \in \mathbb{R}^k \mid f(t) \text{ soit fini}\}$$

Les hypothèses 9.1 entraînent la convexité stricte de $m = \log \hat{\mu}$ sur $D_m$. La fonction m est alors "essentially strictly convex" au sens de [42] . D'après [42] il en résulte que $\lambda = m^*$ est "essentially smooth" au sens de [42] , ce qui s'énonce explicitement comme suit.

9.3. <u>Proposition</u> (corollaire de Rockafellar [42] ) : sous l'hypothèse 9.1, l'ensemble $D_\lambda = \{x \in \mathbb{R}^k \mid \lambda(x) \text{ fini}\}$ est d'intérieur non vide ; la fonction $\lambda$ est différentiable à l'intérieur de $D_\lambda$ ; de plus si y est sur la frontière de $D_\lambda$ , on a $\lim \|\lambda'(x)\| = +\infty$ quand x tend vers y en restant dans l'intérieur de $D_\lambda$.

9.4. <u>Dualité de Legendre</u> : soit f une fonction différentiable sur un ouvert C de $\mathbb{R}^k$. Supposons que f' soit une bijection de C sur une partie D de $\mathbb{R}^k$. Appelons, avec Rockafellar [42] , <u>dual de Legendre du couple</u> (C, f) le couple (D, g) où $g : D \to \mathbb{R}$ est définie pour $x \in D$ par

$$\begin{cases} g(x) = \langle z, x \rangle - f(z) \\ f'(z) = x, \qquad z \in C \end{cases}$$

La transformation de Legendre intervient par exemple en mécanique ou en calcul des variations (cf. Courant-Hilbert [12] ).

9.5. Proposition (corollaire de Rockafellar [42] ) : soit $\mu$ une probabilité sur $\mathbb{R}^k$ vérifiant 9.1. Supposons de plus $\hat{\mu}$ (t) finie sur un ouvert non vide de $E = \mathbb{R}^k$. Alors (notation (4)), la fonction $m(t) = \log \hat{\mu}(t)$ est continuement différentiable sur $\overset{\circ}{D}_m$ , et la transformée de Legendre du couple $(\overset{\circ}{D}_m, m)$ s'écrit $(D, \lambda)$, où $\lambda$ est la transformée de Cramer de $\mu$ et $D = m'(\overset{\circ}{D}_m)$ est un ouvert contenu dans $D_\lambda$. En particulier on a, pour tout $x \in D$

$$(5) \quad \begin{cases} \lambda(x) = \langle z, x \rangle - \log \hat{\mu}(z) \\ \dfrac{\hat{\mu}'}{\hat{\mu}}(z) = x \ , \ z \text{ unique solution dans } \overset{\circ}{D}_m \end{cases}$$

et $\lambda$ est continuement différentiable sur D.

Preuve : On applique un résultat de [42] au couple $(\overset{\circ}{D}_m, m)$, résultat valable pour $m \in \mathcal{F}$ , $\overset{\circ}{D}_m \neq \emptyset$ et m continuement différentiable sur $\overset{\circ}{D}_m$.

9.6. Le cadre adéquat en dualité de Legendre : soit $\mu$ une probabilité sur $\mathbb{R}^k$ , telle que $\mu$ ne soit portée par aucun sous espace affine strict de $\mathbb{R}^k$. Supposons de plus que $\hat{\mu}$ (t) est finie sur un ouvert non vide de $\mathbb{R}^k$.

La fonction $m(t) = \log \hat{\mu}(t)$ est alors continuement différentiable sur $\overset{\circ}{D}_m$ . Supposons enfin que pour u sur la frontière de $D_m$ , on a $\lim \left\| \dfrac{\hat{\mu}'(t)}{\hat{\mu}(t)} \right\| = +\infty$ quand $t \to u$ en restant dans $\overset{\circ}{D}_m$ .

Dans ces conditions m est "essentially smooth" au sens de [42] . Remarquons que cette situation est trivialement réalisée dès que $\hat{\mu}$ (t) est finie pour tout $t \in \mathbb{R}^k$ . Il existe cependant des cas où $\hat{\mu}$ est finie sur un ouvert non vide tandis que la condition sur $\dfrac{\hat{\mu}'(t)}{\hat{\mu}(t)}$ n'est pas vérifiée ; ainsi sur $\mathbb{R}$ la mesure définie par $d\mu(x) = C \, e^{-|x|} \dfrac{1}{1+|x|^2} dx$ est un exemple de cette situation. Dans ce cas d'ailleurs la transformée de Cramer de $\mu$ n'est pas strictement convexe.

Sous l'hypothèse 9.6, on montre que (cf [42] ) la transformation de Legendre permet de passer non seulement de $(\overset{\circ}{D}_m, m)$ à $(D, \lambda)$, mais aussi de $(D, \lambda)$ à $(\overset{\circ}{D}_m, m)$. De plus on peut alors décrire $D$ et $D_\lambda$ assez explicitement à partir de $\mu$, par le résultat suivant.

9.7. Proposition : Soit $\mu$ une probabilité sur $\mathbb{R}^k$ vérifiant l'hypothèse 9.6. Soit $S_\mu$ l'enveloppe convexe fermée du support de $\mu$ et soit $D_\lambda = \{x \in \mathbb{R}^k \mid \lambda(x) \text{ fini}\}$ , où $\lambda$ est la transformée de Cramer de $\mu$. Les propriétés suivantes sont alors vérifiées :

(6) $\overset{\circ}{S}_\mu = \overset{\circ}{D}_\lambda \subset \overline{D}_\lambda = S_\mu$

(7) $\lambda$ est continuement différentiable et strictement convexe sur $\overset{\circ}{S}_\mu = \overset{\circ}{D}_\lambda$

(8) L'application $t \to \dfrac{\hat{\mu}'}{\hat{\mu}}(t)$ est un difféomorphisme de $\overset{\circ}{D}_m$ sur $\overset{\circ}{D}_\lambda$ . De plus pour tout $x \in \overset{\circ}{D}_\lambda$ , on a
$\lambda(x) = \langle z, x \rangle - \log \hat{\mu}(z)$ , où $z$ est l'unique solution dans $\overset{\circ}{D}_m$ de l'équation $\dfrac{\hat{\mu}'}{\hat{\mu}}(z) = x$

(9) Pour tout $y$ sur la frontière de $D_\lambda$ , on a
$\lim \|\lambda'(x)\| = +\infty$ quand $x \to y$ en restant dans $\overset{\circ}{D}_\lambda$ .

Preuve : D'après 9.5, le dual de Legendre du couple $(\overset{\circ}{D}_m, m)$ est $(D, \lambda)$ , où $D = m'(\overset{\circ}{D}_m)$. Grâce à l'hypothèse 9.6, on a simultanément {m "essentially smooth" au sens de [42]} et {m différentiable sur $\overset{\circ}{D}_m$ } ; d'après [42] , ces propriétés impliquent que le couple dual $(D, \lambda)$ vérifie $D = \overset{\circ}{D}_\lambda$ .

Montrons que cet ensemble n'est autre que $\overset{\circ}{S}_\mu$. On sait déjà (lemme 5.4(5)) que $D_\lambda \subset S_\mu$ , et donc que $D = \overset{\circ}{D}_\lambda$ est contenu dans $\overset{\circ}{S}_\mu$ . Supposons l'existence de $\omega \in \overset{\circ}{S}_\mu$ tel que $\omega \notin D$. Alors il existe un vecteur $t \in \mathbb{R}^k$ tel que $v = \langle t, \omega \rangle$ n'appartienne pas à $t(D) = \{\langle t, x \rangle \mid x \in D\}$ (car $D = \overset{\circ}{D}_\lambda$ est convexe).

Puisque $D = m'(\overset{\circ}{D}_m)$ , l'ensemble $t(D)$ contient tous les points de la forme $\frac{1}{\hat{\mu}(u)} \int_{\mathbb{R}^k} \langle t, x\rangle e^{\langle u,x\rangle} d\mu(x)$ où $u \in \overset{\circ}{D}_m$ . Appliquons ce résultat à tous les points $u$ de $\overset{\circ}{D}_m$ de la forme $u = s\,t$ avec $s \in \mathbb{R}$ (il est facile de voir que de tels points existent en utilisant les hypothèses sur $\mu$ et la convention $\int x\, d\mu(x) = 0$).

En posant $\nu = t(\mu)$ = probabilité sur $\mathbb{R}$ , $n(s) = \hat{\nu}(s)$ pour $s \in \mathbb{R}$, $D_n = \{s \in \mathbb{R} \mid n(s) \text{ fini}\}$ , $\Delta = n'(\overset{\circ}{D}_n)$ , on conclut que $t(D)$ contient $\Delta$.

La relation $D = \overset{\circ}{D}_\lambda$ déjà démontrée plus haut, fournit $\Delta = \overset{\circ}{D}_\gamma$ où $\gamma$ est la transformée de Cramer de $\nu$ sur $\mathbb{R}$. Puisque $v \notin t(D)$ et $\Delta \subset t(D)$, on a donc $v \notin \overset{\circ}{D}_\gamma$ .

D'autre part, il est facile de voir que $\overset{\circ}{S}_{t(\mu)} = t(\overset{\circ}{S}_\mu)$ et donc $v = \langle t, \omega\rangle$ appartient à $\overset{\circ}{S}_\nu$ . Or sur $\mathbb{R}$ , on a déjà démontré (cf. proposition 2.7 (5)) que $\overset{\circ}{S}_\nu = \overset{\circ}{D}_\gamma$ , ce qui contredit l'existence de $v \in (\overset{\circ}{S}_\nu - \overset{\circ}{D}_\gamma)$ , et donc l'existence de $\omega$. On a ainsi prouvé par l'absurde que $\overset{\circ}{S}_\mu = \overset{\circ}{D}_\lambda$ .

Pour tout convexe $C$ de $\mathbb{R}^k$ d'intérieur non vide, on sait que $C$ et $\overset{\circ}{C}$ ont même adhérence. Les relations $\overset{\circ}{S}_\mu = \overset{\circ}{D}_\lambda \subset D_\lambda \subset S_\mu$ entraînent alors $\overline{D_\lambda} = \overline{\overset{\circ}{D}_\lambda} = \overline{\overset{\circ}{S}_\mu} = S_\mu$ , ce qui achève la preuve de (6).

La propriété (7) résulte alors de (9.5). D'après les résultats de [42] sur la dualité de Legendre dans le cadre envisagé ici, l'application $t \to m'(t)$ est une bijection de $\overset{\circ}{D}_m$ sur $D = m'(\overset{\circ}{D}_m)$ , donc sur $\overset{\circ}{D}_\lambda$ . Cette bijection est un difféomorphisme (théorème des fonctions implicites), grâce à l'hypothèse 9.1 sur $\mu$. Le reste de l'assertion (8) se déduit alors de 9.5 (5) tandis que (9) résulte de 9.3.

Nous renvoyons à Azencott-Ruget [3] et Bahadur-Zabell [7] pour des compléments sur les transformées de Cramer en dimension finie. Dans [3] est abordée l'étude des propriétés de continuité de $\lambda_\mu$ (x) comme fonction du couple (x, $\mu$) et de robustesse de la loi des grandes déviations, ainsi que l'allure de $\lambda$ au bord de $S_\mu$ . Dans [7] on trouvera quelques exemples sur les pathologies du comportement de $\lambda$ au bord de $S_\mu$.

## CHAPITRE II

## APPLICATIONS AUX MESURES GAUSSIENNES ET AUX PROCESSUS GAUSSIENS

1. Transformées de Cramer des mesures gaussiennes sur les espaces de Banach :

1.1 Définition : Soit E un espace de Banach séparable.

Une probabilité $\mu$ sur la tribu borélienne $\mathcal{B}(E)$ de E est dite gaussienne si pour toute forme linéaire continue $t : E \to R$, $t(\mu)$ est une loi gaussienne sur $\mathbb{R}$. Si de plus $t(\mu)$ est centrée pour tout $t \in E'$, nous disons que $\mu$ est centrée.

1.2 Théorème (Skorokhod [48] Fernique [22]) Si $\mu$ est une probabilité gaussienne sur un Banach séparable, alors $\int_E e^{s||x||} \, d\mu(x)$ est fini

pour tout $s \in R$. En particulier l'intégrale $\int_E ||x||^2 d\mu(x)$ est finie.

Preuve : Nous renvoyons à Kuo [29]

1.3 Covariance et espace de Hilbert associé :

Notons E' le dual de E, et $<t,x>$ le produit scalaire de $t \in E'$ par $x \in E$. Définissons , lorsque $\mu$ est centrée, la covariance $K : E' \times E' \to \mathbb{R}$ de $\mu$ par

$$K(s,t) = \int <s,x> <t,x> \, d\mu(x) \qquad \text{par } s,t \in E' \;;$$

comme $|K(s,t)| \leq c^2 \, ||s|| \; ||t||$ avec $c^2 = \int ||x||^2 d\mu(x)$,

l'application K est bilinéaire continue sur $E' \times E'$ ; de plus $K(t,t) \geq 0$ pour tout $t \in E'$.

Classiquement (cf. Kuo [29], Bahadur-Zabell [ 7]), on plonge E' dans un espace de Hilbert H de la façon suivante :

à tout $t \in E'$ on associe la v.a. gaussienne $Zt : E \to \mathbb{R}$, définie sur l'espace de probabilité $(E, \mathcal{B}(E), \mu)$ par $Zt(x) = <t,x>$. Soit H la fermeture dans $L_2(E, \mathcal{B}(E), \mu)$ de l'espace vectoriel $\{Zt,\ t \in E'\}$.

L'application $Z : E' \to H$ est linéaire continue et d'image dense. De plus si on note $[v,w]$ le produit scalaire dans $L_2(E, \mathcal{B}(E),\mu)$ et donc dans H, on a

$$K(s,t) = [Zs,Zt] \qquad \text{pour } s,t \in E'$$

Remarquons que si $v \in L_2(E,\mathcal{B}(E),\mu)$ on a

$$\int_E ||x|| \; |v(x)|d\mu(x) \leq c||v||_{L_2(E,\mu)} \text{ avec } c^2 = \int_E ||x||^2 d\mu(x).$$

<u>La formule</u> $Sv = \int_E x\, v(x)d\mu(x)$ <u>définit donc un opérateur linéaire continu $S : L_2(E,\mathcal{B}(E),\mu) \to E$</u>. Par construction on a

$$\int_E Zs(x)\; v(x)\; d\mu(x) = \int_E \langle s,x\rangle\; v(x)\; d\mu(x) = \langle s, \int_E x\; v(x)d\mu(x)\rangle$$

et donc

$$[Zt,v] = \langle t,Sv\rangle \qquad \text{pour } t \in E',\; v \in L_2(E,\mu)$$

ce qui prouve que la restriction de S à H est injective.

1.4 Ainsi étant donnée une probabilité gaussienne $\mu$ sur un Banach séparable E, on peut lui associer un espace de Hilbert séparable H (où le produit scalaire est noté $[.,.]_H$), une injection linéaire continue $S : H \to E$, une application linéaire continue (d'image dense) $Z : E' \to H$ telles que

$$[Zt,v]_H = \langle t,Sv\rangle \qquad \text{pour } t \in E',\; v \in H$$

$$[Zt,Zs]_H = K(t,s) \qquad \text{pour } t,s \in E'$$

où $K(t,s)$ est la covariance de $\mu$.

1.5 <u>Proposition</u> (d'après Bahadur-Zabell [ 7]) Soit $\mu$ une probabilité gaussienne centrée sur l'espace de Banach séparable E. Soient H un espace de Hilbert et $S : H \to E$ une injection linéaire continue associées à $\mu$ comme en 1.4. Alors la transformée de Cramer $\lambda$ de $\mu$ est donnée par

$$\begin{cases} \lambda(x) = \frac{1}{2} \|S^{-1}x\|_H^2 & \text{si } x \in SH \\ \lambda(x) = +\infty & \text{si } x \in (E - SH). \end{cases}$$

De plus S est un operateur compact.

Preuve : La transformée de Laplace $\hat{\mu}(t)$ s'écrit

$$\hat{\mu}(t) = \int_E e^{<t,x>} d\mu(x) = E\left[\exp(Zt)\right]$$

où Zt est la v.a. gaussienne de $L_2(E,\mu)$ définie en 1.3. On a donc $\log\hat{\mu}(t) = \frac{1}{2} \operatorname{var}(Zt) = \frac{1}{2} K(t,t)$ où K est la covariance de $\mu$. Le Th.5.3 Ch.I, fournit alors

(1) $\lambda(x) = \sup_{t \in E'} \left[ <t,x> - \frac{1}{2} K(t,t)\right]$ pour $x \in E$ ce qui devient en tenant compte de 1.4

(2) $\lambda(x) = \sup_{t \in E'} \left[ <t,x> - \frac{1}{2} \|Zt\|_H^2 \right]$

Soit x fixé dans E, tel que <u>$\lambda(x)$ soit fini</u>. Pour tout t dans le noyau ker(Z) de Z et tout réel a, on a d'après (2), $\lambda(x) \geq a <t,x>$ ce qui force $<t,x> = 0$. Soit im(Z) l'image de Z. La remarque qui vient d'être faite prouve l'existence d'une application linéaire $g : \text{im}(Z) \to R$ bien définie (mais pas nécessairement continue à priori) telle que $g(Zt) = <t,x>$ pour tout $t \in E'$.

Toujours d'après (2), on a

$$|g(v)| \leq \lambda(x) + \frac{1}{2} \|v\|_H^2 \quad , \quad \text{pour } v \in \text{im}(Z)$$

de sorte que g est bornée sur l'intersection de im(Z) avec la boule unité de H. Comme im(Z) est dense dans H, g se prolonge de façon unique en une forme linéaire continue sur H. Par suite, il existe

$w \in H$ tel que $g(v) = [v,w]_H$ pour tout $v \in H$.

En particulier 1.4 entraîne

$$<t,x> = g(Zt) = [Zt,w]_H = <t,Sw> , \quad \text{pour tout } t \in E'$$

ce qui prouve $x = Sw$ et donc $x \in SH$.

Inversement soit $x \in SH$ et écrivons $x = Sw$ avec $w \in H$. La formule (2) et 1.4 donnent alors

$$\lambda(x) = \sup_{t \in E'} \{ [Zt, w]_H - \frac{1}{2} ||Zt||_H^2 \}$$

d'où, puisque im(Z) est dense dans H

$$\lambda(x) = \sup_{v \in H} \{ [v, w]_H - \frac{1}{2} ||v||_H^2 \}$$

Pour $||v||_H = a$ fixé, le sup. de $[v, w]_H$ est atteint et vaut $a\,||w||_H$. Comme le sup. de $(a||w||_H - \frac{1}{2} a^2)$ pour $a \geqslant 0$ vaut $\frac{1}{2}||w||_H^2$, on obtient $\lambda(x) = \frac{1}{2}||w||_H^2 = \frac{1}{2}\,||S^{-1}x||_H^2$ et en particulier $\lambda(x)$ est fini, ce qui achève la détermination de $\lambda$.

Le calcul de $\lambda$ montre aussi que l'image par S de la boule unité de H est $\{ x \in E \mid \lambda(x) \leqslant 1 \}$ ; ce dernier ensemble est compact (cf. th. 7.5 ch. I) et S est donc un opérateur compact (on retrouve ici un résultat classique sur les mesures gaussiennes voir Kuo [29]).

1.6 Théorème : Soit $\mu$ une probabilité gaussienne centrée sur un Banach séparable E. Soit $X : \Omega \to E$ une variable aléatoire de loi $\mu$. Soit $\Lambda$ la fonctionnelle de Cramer associée à $\mu$ (def. ch.I). Pour toute partie borélienne A de E, on a

$$- \Lambda(\overset{\circ}{A}) \leqslant \varliminf_{\varepsilon \to 0} \varepsilon^2 \log P(\varepsilon X \in A) \leqslant \varlimsup_{\varepsilon \to 0} \varepsilon^2 \log P(\varepsilon X \in A) \leqslant - \Lambda(\overline{A})$$

Note bibliographique : Des résultats de ce type ont été obtenus, sous des formes moins générales, par Ventcell et Freidlin [52] [23] dans le cas où E est un espace de Hilbert ainsi que lorsque E est l'espace des fonctions continues sur [0,1] . Comme on le verra plus loin, la prop. 1.6 contient un résultat sur la queue des mesures gaussiennes, démontré à l'aide de théorèmes à la Cramer-Chernoff par Donsker-Varadhan [15].

<u>Preuve</u> : Soit $\varepsilon \in ]0,1[$ . Soit $n(\varepsilon)$ l'entier positif défini par l'inégalité $\frac{1}{n(\varepsilon)+1} \leqslant \varepsilon^2 < \frac{1}{n(\varepsilon)}$ . On peut alors écrire :

(1) $\varepsilon = \dfrac{b(\varepsilon)}{\sqrt{n(\varepsilon)}}$ avec $1 - \varepsilon^2 \leqslant b^2(\varepsilon) \leqslant 1$.

Soient $X_p : \Omega \to E$ une suite de v.a. indépendantes de loi $\mu$. Comme $\mu$ est gaussienne centrée, la moyenne normalisée $\sqrt{p}\ \overline{X}_p$ a même loi que $X$ pour tout $p$.

Par conséquent on a

(2) $P(\varepsilon X \in A) = P(\varepsilon\sqrt{n(\varepsilon)}\ \overline{X}_{n(\varepsilon)} \in A) = P(\overline{X}_{n(\varepsilon)} \in \frac{1}{b(\varepsilon)}\ A)$

Posons $\Gamma = \{ x \in E \mid x = ay,\ 2 \geqslant a \geqslant 1,\ y \in \overline{A} \}$. Par définition de $\Lambda$ on peut écrire, en remarquant (prop. 1.6) que la transformée de Cramer $\lambda$ de $\mu$ vérifie $\lambda(ax) = a^2\lambda(x)$ par tout $a \in \mathbb{R}$, $x \in E$,

(3) $\Lambda(\Gamma) = \inf\limits_{x \in \Gamma} \lambda(x) = \inf\limits_{y \in \overline{A}}\ \inf\limits_{a \in [1,2]} a^2\lambda(y) = \inf\limits_{y \in \overline{A}} \lambda(y) = \Lambda(\overline{A})$

D'autre part $\frac{1}{b(\varepsilon)}A$ est contenu dans l'ensemble <u>fermé</u> $\Gamma$ pour $\varepsilon$ assez petit. Comme $\lim\limits_{\varepsilon \to 0} n(\varepsilon) = +\infty$, le th. 7.5 ch. I entraîne

$$\overline{\lim_{\varepsilon \to 0}}\ \frac{1}{n(\varepsilon)} \log P(\overline{X}_{n(\varepsilon)} \in \frac{1}{b(\varepsilon)}\ A) \leqslant \overline{\lim_{n \to +\infty}}\ \frac{1}{n} \log P(\overline{X}_n \in \Gamma) \leqslant -\Lambda(\Gamma)$$

c'est à dire d'après (1) (2) et (3)

$$\overline{\lim_{\varepsilon \to 0}}\ \varepsilon^2 \log P(\varepsilon X \in A) \leqslant -\Lambda(\overline{A}).$$

Soit $x \in \overset{\circ}{A}$ . Il existe un voisinage ouvert convexe $V$ de $x$ tel que pour $\varepsilon$ assez petit on ait $V \subset \frac{1}{b(\varepsilon)}\ A$

D'après le th. 7.5 ch. I, on a donc

$$\underline{\lim_{\varepsilon \to 0}}\ \varepsilon^2 \log P(\varepsilon X \in A) \geqslant \underline{\lim_{\varepsilon \to 0}}\ \frac{b^2(\varepsilon)}{n(\varepsilon)} \log P(\overline{X}_{n(\varepsilon)} \in V) = -\Lambda(V) \geqslant -\lambda(x)$$

Prenons le sup. du dernier terme pour $x \in \overset{o}{A}$ pour obtenir

$\lim\limits_{\varepsilon \to 0} \varepsilon^2 \log P(\varepsilon X \in A) \geqslant - \Lambda(\overset{o}{A})$ , ce qui termine la preuve de 1.6.

1.7 <u>Corollaire</u> (d'après Donsker-Varadhan [15]) : Soit $\mu$ une probabilité gaussienne centrée sur un Banach séparable E. Soit $\lambda$ sa transformée de Cramer et K sa covariance.

Les nombres $a = \inf \{\lambda(x) \mid x \in E, \|x\| = 1\}$ et $b = \sup \{ k(t,t) \mid t \in E', \|t\| = 1 \}$ sont alors finis, strictement positifs et on a

$$\lim_{R \to +\infty} \frac{1}{R^2} \log \mu \{ x \in E \mid \|x\| \geqslant R \} = -a = -\frac{1}{2b}$$

<u>Preuve</u> : Soit $B = \{ x \in E \mid \|x\| \geqslant 1 \}$. La relation $\lambda(\alpha x) = \alpha^2 \lambda(x)$ pour $\alpha \in R$, $x \in E$ donne

$$\Lambda(B) = \inf_{x \in B} \lambda(x) = \inf_{\|y\|=1} \inf_{\alpha \geqslant 1} \alpha^2 \lambda(y) = a$$

$$\Lambda(\overset{o}{B}) = \inf_{\|y\|=1} \inf_{\alpha > 1} \alpha^2 \lambda(y) = a$$

Comme B est fermé, le th. 1.6 montre que (en notant X une v.a. à valeurs dans E, de loi $\mu$)

$$\lim_{R \to +\infty} \frac{1}{R^2} \log \mu \{x \in E \mid \|x\| \geqslant R \} = \lim_{R \to +\infty} \frac{1}{R^2} \log P(\tfrac{1}{R} X \in B) = -a$$

Reste à comparer a et 1/2b.

La formule 1.5 (1) donne pour $x \in E$, $t \in E'$, $r \in \mathbb{R}^+$

$$\lambda(x) \geqslant r \langle t, x\rangle - \frac{1}{2} r^2 K(t,t)$$

Soit x fixé dans E, avec $\|x\| = 1$. Choisissons $t_o \in E'$ tel que $\|t_o\| = 1$ et $\langle t_o, x\rangle = 1$, ce qui est possible grâce au th. de Hahn-Banach.

On obtient

$$\lambda(x) \geqslant \sup_{r\in\mathbb{R}^+} \left[ r - \frac{1}{2} r^2 K(t_o,t_o)\right] = \frac{1}{2K(t_o,t_o)} \geqslant \frac{1}{2b}$$

pour tout x tel que $||x|| = 1$ , c'est à dire $a \geqslant \frac{1}{2b}$ .

Soit maintenant $t \in E'$ tel que $||t|| = 1$. La forme linéaire $t : E \to R$ vérifie

$$t^{-1} \left[1, +\infty\right[ \subset B = \{x \in E \mid ||x|| \geqslant 1 \}$$

Soit $X_n : \Omega \to E$ une suite de v.a. de loi $\mu$, indépendantes.
Les moyennes $\overline{X}_n$ et $\overline{t(X_n)}$ des echantillons $X_1 \dots X_n$ et $t(X_1)\dots t(X_n)$ vérifient $P(\overline{t(X_n)} \in [1, +\infty[) \leqslant P(\overline{X}_n \in B)$. Passons au logarithme et faisons tendre n vers $+\infty$ pour obtenir, en notant $\lambda_{t(\mu)}$ la transformée de Cramer de $t(\mu)$,

$$-\frac{1}{2K(t,t)} = -\lambda_{t(\mu)}(1) = -\Lambda_{t(\mu)}\left[1, +\infty\right[ \leqslant -\Lambda(B) = -a$$

L'inégalité $\frac{1}{2K(t,t)} \geqslant a$ pour tout $t \in E'$ de norme 1 entraîne $\frac{1}{2b} \geqslant a$, et finalement $\frac{1}{2b} = a$.

Comme K est bilinéaire continue (cf. 1.3), le nombre b est fini et donc a est strictement positif. Le calcul de $\lambda$ montre que a est fini et donc que b est strictement positif.

Nous verrons dans les paragraphes suivants des situations où on peut calculer plus explicitement la fonction $\lambda$ et le nombre a.

## 2. Cas particulier des espaces de Hilbert :

2.1 Notations et hypothèses : Soit E un espace de Hilbert séparable et $\mu$ une probabilité gaussienne (centrée) sur $\mathcal{B}$ (E). On a E' = E, et on sait (cf. Kuo [29] Badrikian [ 4]) que la covariance K(s,t) s'écrit

$$K(s,t) = <s,\Gamma t> = <\Gamma s,t> \qquad s,t \in E$$

où $\Gamma$: $E \to E$ est un opérateur linéaire continu, self-adjoint, positif, et à trace finie. Ceci équivant à dire qu'il existe une base orthonormale $e_n$ de E et des nombres positifs $\rho_n$ tels que $\sum_{n=1}^{+\infty} \rho_n$ soit fini, vérifiant

$$\Gamma x = \sum_{n=1}^{+\infty} <x,e_n> \rho_n e_n , \qquad x \in E$$

On notera $\sqrt{\Gamma}$ l'unique racine carrée self adjointe de $\Gamma$, définie par

$$\sqrt{\Gamma}\ x = \sum_{n=1}^{+\infty} <x,e_n> \sqrt{\rho_n}\ e_n , \qquad x \in E$$

On notera H l'adhérence de $\Gamma(E)$, N le noyau de $\Gamma$. Il est évident que H est l'orthogonal de N dans E, et que N, H sont aussi le noyau et l'adherence de l'image de $\sqrt{\Gamma}$. En particulier la réstriction de $\sqrt{\Gamma}$ au sous espace fermé H est injective, et on a $\sqrt{\Gamma}(E) = \sqrt{\Gamma}(H)$.

2.2 Proposition : Soit $\mu$ une probabilité gaussienne (centrée) sur l'espace de Hilbert séparable E. Soit $\Gamma : E \to E$ l'opérateur de covariance associé à $\mu$ et soit H l'orthogonal dans E du noyau de $\Gamma$. Soit S la restriction de $\sqrt{\Gamma}$ à H, de sorte que $S : H \to E$ est injectif et envoie H sur $\sqrt{\Gamma}(E)$. Alors la transformée de Cramer $\lambda$ de $\mu$ est donnée par

$$\lambda(x) = \frac{1}{2} ||S^{-1}x||^2 \qquad \text{pour } x \in \sqrt{\Gamma}(E)$$

$$\lambda(x) = +\infty \qquad \text{pour } x \notin \sqrt{\Gamma}(E)$$

Preuve : Pour tout $s \in E' = E$ posons $Zs = \sqrt{\Gamma}s$. Puisque $\sqrt{\Gamma}$ laisse H invariant et vérifie $\overline{\sqrt{\Gamma}(E)} = H$ on voit que $Z : E' \to H$ est un opérateur linéaire continu d'image dense, tandis que $S : H \to E$ est injectif.

Les relations 1.4 du paragraphe 1 entre K, S et Z sont trivialement vraies. Il ne reste plus qu'à appliquer la prop. 1.5, en notant que $S(H) = \sqrt{\Gamma}(E)$.

## 3. Application aux processus gaussiens :

3.1 Hypothèses et notations : Soit $X_s : \Omega \to \mathbb{R}$, $s \in [0,1]$ un processus aléatoire gaussien, centré. Nous supposerons que le processus $(X_s)$ est presque sûrement à trajectoires continues.

La covariance $\rho$ du processus

$$\rho(s,u) = \int_\Omega X_s(\omega)\, X_u(\omega)\, dP(\omega) \quad , \; s,u \in [0,1]$$

est alors continue en (s,u).

Rappelons d'ailleurs (cf. Fernique [22]) que si $\rho$ est assez régulière (par exemple holderienne d'ordre $\alpha > 0$, ce que s'écrit

$$|\rho(s,u) - \rho(s',u')| \leqslant \text{cte}\left[\,|s-s'|^\alpha + |u-u'|^\alpha\right],$$

pour $s, s', u, u' \in [0,1]$ ), alors le processus $(X_s)$ admet une version à trajectoires presque sûrement continues.

Nous pouvons toujours, dès que le processus $(X_s)$ admet une version à trajectories p.s. continues, en déduire une version du processus qui soit à trajectoires sûrement continues. On obtient ainsi une application mesurable $X : \Omega \to C\,[0,1]$ , où $C[0,1]$ est l'espace des fonctions continues sur $[0,1]$ muni de sa tribu borélienne et où $X(\omega)$ est la fonction qui à $s \in [0,1]$ associe $X_s(\omega)$.

Soit $\mu$ l'image de la probabilité P (donnée sur $\Omega$) par l'application X. Alors $\mu$ est une probabilité gaussienne centrée sur $(E, \mathcal{B}(E))$, où E est $C[0,1]$ muni de sa structure d'espace de Banach séparable usuelle.

En effet E' est l'espace des mesures de Radon bornées sur $[0,1]$ ; pour tout $\pi \in E'$, la v.a. $Z\pi : E \to R$ définie par

$$Z\pi(f) = \langle \pi, f\rangle = \int_{[0,1]} f(s)d\pi(s), \qquad f \in E$$

a même loi (lorsque E est muni de la probabilité $\mu$) que la v.a. $Y = \Omega \to R$ définie par $Y(\omega) = \int_{[0,1]} X_s(\omega)\, d\pi(s)$

Par suite, la mesure image de $\mu$ par toute forme linéaire continue est gaussienne centrée, et $\mu$ est bien gaussienne centrée.

La <u>covariance K de</u> $\mu$ est donnée par

$$K(\pi,\eta) = \mathrm{cov}\left[\int_{[0,1]} X_s d\pi(s), \int_{[0,1]} X_s d\eta(s)\right] = \int_{[0,1]\times[0,1]} \rho(s,t)d\pi(s)d\eta(t)$$

3.2 <u>Proposition</u> (d'après Donsker-Varadhan [ 15]) :

Soit $X_s : \Omega \to R$ un procesus gaussien centré à trajectoires presque sûrement continues, de covariance $\rho(s,t)$, $s,t \in [0,1]$. Posons

$b = \sup_{s\in[0,1]} \rho(s,s)$. Alors on a

$$\lim_{r\to\infty} \frac{1}{r^2} \log P\{ \sup_{s\in[0,1]} |X_s| \geq r \} = -\frac{1}{2b}$$

<u>Preuve</u> : D'après le cor.1.7, il suffit de vérifier que, E' désignant l'espace des mesures bornées sur $[0\ 1]$.

$$\sup\{K(\eta,\eta) \mid \eta \in E' \quad ||\eta|| = 1\} = b$$

Prenons $\eta \in E'$, telle que $||\eta|| = 1$. On peut ecrire

$$\eta = \eta^+ - \eta^-, \quad \tilde{\eta} = \eta^+ + \eta^-, \text{ et } ||\tilde{\eta}|| = ||\eta|| = 1$$

d'où en notant $J = [0,1]$

$$K(\eta,\eta) = \int_{J\times J} \rho(s,t)d\eta(s)d\eta(t) \leq \int_{J\times J} |\rho(s,t)|d\tilde{\eta}(s)d\tilde{\eta}(t)$$

La fonction $\rho$ étant définie positive, on a $\rho(s,t)^2 \leqslant \rho(s,s)\rho(t,t)$, et donc $\sup_{s,t\in J} |\rho(s,t)| = b$, ce qui démontre $K(\eta,\eta) \leqslant b$. D'autre part, par compacité, il existe $s \in J$ tel que $\rho(s,s) = b$, ce qui donne $K(\delta_s,\delta_s) = b$ et achève la preuve.

3.3 Remarque : Si le processus gaussien $(X_s)$ est à trajectoires presque sûrement dans $L_p[0,1]$, avec $p \geqslant 1$, le même résultat 1.7 montre que

$$\lim_{r\to+\infty} \frac{1}{r^2} \log P\{(\int_0^1 |X_s|^p ds)^{1/p} \geqslant r\}$$

existe, mais le calcul de cette limite à partir de $\rho$ n'est pas possible "explicitement" en général.

3.4 Proposition : (résultats analogues dans Ventcell [52] et Freidlin[23]): Soit $X_s : \Omega \to \mathbb{R}$ un processus gaussien centré continu sur l'intervalle $[0,1]$, de covariance $\rho(s,u)$. Définissons un opérateur $R$ de $L_2[0\ 1]$ dans $L_2[0,1]$ pour $Rf(s) = \int_{[0,1]} \rho(s,u)f(u)du$. Alors $R$ est à valeurs dans $C[0\ 1]$, self-adjoint, positif, compact, et à trace finie. Soit $H$ l'orthogonal dans $L_2[0,1]$ du noyau de $R$. Soit $S$ la restriction à $H$ de l'opérateur $\sqrt{R}$, de sorte que $S : H \to L_2[0,1]$ est injectif et envoie $H$ sur $\sqrt{R}(L_2[0\ 1])$.

Soit $\mu$ la probabilité gaussienne sur $C\,[0,1]$ associée au processus $(X_s)$ comme en 2.1. La transformée de Cramer $\lambda$ de $\mu$ est donnée par les formules suivantes (où $f \in C[0,1]$) :

$$\lambda(f) = \frac{1}{2}\, ||S^{-1}f||^2_{L_2[0,1]} \qquad \text{si } f \in \sqrt{R}(L_2[0,1])$$

$$\lambda(f) = +\infty \qquad \text{si } f \notin \sqrt{R}(L_2[0,1])$$

La démonstration va utiliser le lemme suivant :

3.5 Lemme : Soient $C$ et $L$ deux espaces de Banach séparables ; soit $\mu$ une probabilité gaussienne centrée sur $C$. Supposons qu'il existe une injection continue $i$ de $C$ dans $L$. Alors les transformées de Cramer des mesures $\mu$ et $i(\mu)$ sont liées par la relation

$$\lambda_\mu = \lambda_{i(\mu)} \circ i$$

Preuve : Soit $i^* : L^* \to C^*$ l'application duale de i. Posons $\nu = i(\mu)$. La définition 1.3 des covariances montre immédiatement que

$$K_\nu(s,t) = K_\mu(i^*s,\, i^*t) \quad \text{pour } s,t \in L^*$$

Comme i est injective, $i^*(L^*)$ est dense dans $C^*$, et la formule 1.5 (1) donne, pour $x \in C$

$$\lambda_\mu(x) = \sup_{t \in C^*} \left[<t,x> - \frac{1}{2} K_\mu(t,t)\right]$$

$$= \sup_{s \in L^*} \left[<i^*(s),x> - \frac{1}{2} K_\mu(i^*(s),\, i^*(s))\right]$$

$$= \sup_{s \in L^*} \left[<s,\, i(x)> - \frac{1}{2} K_\nu(s,s)\right]$$

c'est à dire $\lambda_\mu(x) = \lambda_\nu(i(x))$.

Preuve de la prop. 3.4 : Prenons $C = C[0,1]$ et $L = L_2[0,1]$; appliquons le résultat précédent en appelant i l'injection naturelle de C dans L. Pour $f \in C[0,1]$ on aura donc $\lambda(f) = \lambda_\nu(f)$ où $\nu$ est l'image de $\mu$ par i, et où par abus de langage on identifie f et i(f).

Il est clair (cf. 3.1) que la covariance $K_\nu$ de $\nu$ est donnée par $K_\nu(f,g) = <Rf,g>_{L_2[0,1]}$. Le calcul de $\lambda_\nu$ pour l'espace de Hilbert L a été fait dans la prop. 2.2. Comme $\lambda$ est la "restriction" de $\lambda_\nu$ à C, ceci achève la preuve de 3.4.

.6 Proposition : (cas du mouvement Brownien) : Soit $\beta_t : \Omega \to R$ un mouvement brownien indexé par $t \in [0,1]$, tel que $\beta_o = 0$ presque sûrement. Soit $C_o$ l'espace de Banach des functions continues sur $[0,1]$ nulles en 0. Soit $\mu$ la loi des trajectoires du mouvement brownien $(\beta_t)$ sur $C_o$. La transformée de Cramer $\lambda$ de $\mu$ est donnée, pour $f \in C_o$, par la formule

$$\lambda(f) = \frac{1}{2} \int_o^1 f'(t)^2 dt \qquad \text{si f est absolument continue,}$$

de dérivée (au sens de Lebesgue) $f'$, tandis que $\lambda(f) = +\infty$ si $f$ n'est pas absolument continue.

<u>Preuve</u> : Nous allons appliquer 3.4 à l'espace $C_o$ (au lieu de l'espace $C[0,1]$, ce qui ne change évidemment rien à la validité de 3.4). On a ici $\rho(s,t) = s\wedge t$, et l'opérateur $R : L \to L$ (où $L = L_2[0,1]$) est donné pour $g \in L$, par la formule

$$Rg(s) = \int_o^1 (s\wedge t)g(t)dt = \int_o^s tg(t)dt + s\int_s^1 g(t)dt$$

Il est clair que si $f = Rg$, $f$ est nulle en 0, continue, admet une dérivée continue $f'(s) = \int_s^1 g(t)dt$ nulle en 1, et admet une dérivée seconde $f'' = -g$ de carré intégrable. On en déduit que $R$ est injectif, que $R(L)$ est l'ensemble des $f \in C_o$, deux fois dérivables, telles que $f''$ soit dans $L$ et $f'(1) = 0$ ; enfin pour $f \in R(L)$ on a $R^{-1}f = -f''$.

Prenons $f \in R(L)$ ; puisque $f = -Rf''$ on a

$$||(\sqrt{R})^{-1}f||_L^2 = ||\sqrt{R}f''||_L^2 = \langle Rf'',f''\rangle_L = -\langle f,f''\rangle_L = ||f'||_L^2$$

où la dernière égalité s'obtient en intégrant par parties, car $f(0) = f'(1) = 0$.

D'après 3.4 on a donc

$$\lambda(f) = \frac{1}{2}||(\sqrt{R})^{-1}f||_L^2 = \frac{1}{2}||f'||_L^2 \quad \text{pour } f \in R(L).$$

Il s'agit d'étendre ce résultat au cas $f \in \sqrt{R}(L)$. Soit donc $f = \sqrt{R}\,g$ avec $g \in L$, $f \in C_o$. Puisque $R$ est injectif, $\sqrt{R}$ l'est aussi, et par suite $\sqrt{R}(L)$ est dense dans $L$. Il existe alors une suite $h_n \in L$ telle que $\sqrt{R}h_n = g_n$ tende vers $g$ dans $L$, ce qui entraîne la convergence de $f_n = \sqrt{R}g_n = Rh_n$ vers $f$ dans $C_o$. Puisque $f_n \in R(L)$, on peut écrire

$$||g_n||_L = ||(\sqrt{R})^{-1}f_n||_L = ||f_n'||_L$$

et comme $g_n$ tend vers $g$ on voit que $f_n'$ est une suite bornée dans $L$. On peut donc en extraire une sous-suite (encore notée $f_n'$) qui converge faiblement vers $k \in L$.

Par suite $f_n(t) = \langle f'_n, 1_{[0,t]}\rangle_L$ converge pour chaque $t \in [0,1]$ vers $\bar{k}(t) = \langle k, 1_{[0,t]}\rangle = \int_0^t k(s)ds$. Comme $f_n$ tend vers f dans $C_o$ on voit que $f = \bar{k}$, et donc $k = f'$ existe et appartient à L. De plus puisque $f'_n$ converge faiblement vers f' dans L, on a

$$||f'||_L \leqslant \underline{\lim}_{n\to\infty} ||f'_n||_L = \lim_{n\to\infty} ||g_n||_L = ||g||_L = ||(\sqrt{R})^{-1}f||_L$$

c'est à dire en tenant compte de 3.4,

$$\frac{1}{2}||f'||_L^2 \leqslant \lambda(f)$$

inégalité que nous venons de démontrer pour f dans $C_o \cap \sqrt{R}(L)$. De plus on vient de voir que f dans $C_o \cap \sqrt{R}(L)$ entraîne l'existence de f' et la propriété $f' \in L$.

Inversement soit $f \in C_o$ telle que f' existe et appartienne à L. Approchons f' dans $L_2[0,1]$ par des fonctions de classe 1 pour obtenir une suite $v_n \in C_o$ telle que $v'_n$, $v''_n$ existent, $v''_n \in L$, $v'_n(1) = 0$, vérifiant de plus $\lim_{n\to\infty} v'_n = f'$ dans L et $\lim_{n\to\infty} v_n = f$ dans $C_o$.

La semi-continuité inférieure de $\lambda$ donne

$$\lambda(f) \leqslant \underline{\lim}_{n\to\infty} \lambda(v_n)$$

D'autre part, puisque $v_n \in R(L)$ on a

$$\lambda(v_n) = \frac{1}{2}||v'_n||_L^2$$

et finalement $\lambda(f) \leqslant \lim_{n\to\infty} \frac{1}{2} ||v'_n||_L^2 = \frac{1}{2}||f'||_L^2$

En particulier $\lambda$ est fini, ce qui d'après 3.4 implique $f \in \sqrt{R}(L)$.

Ceci achève l'identification de $C_o \cap \sqrt{R}(L)$ avec l'ensemble des $f \in C_o$ telles que f' existe et appartienne à L, et la preuve de $\lambda(f) = \frac{1}{2}||f'||_L^2$ pour f dans $C_o \cap \sqrt{R}(L)$, tandis que $\lambda$ vaut $+\infty$ hors de ce sous-espace.

# CHAPITRE III

# PETITES PERTURBATIONS DE SYSTEMES DYNAMIQUES

## 1. PETITES PERTURBATIONS PAR UN BRUIT GAUSSIEN

1.1. Le modèle : nous reprenons ici un modèle de petites perturbations gaussiennes étudié par Ventcell [52] et Freidlin [23] , en introduisant quelques améliorations techniques.

Soit $b : \mathbb{R}^n \to \mathbb{R}^n$ un champ de vecteurs lipschitzien (localement). On lui associe le système dynamique

$$(S) \qquad y'_t = b\,(y_t)$$

où $y_t$ est solution maximale de (S).

Soit $y \to \sigma\,(y)$ une application lipschitzienne (localement) de $\mathbb{R}^n$ dans l'espace des matrices (n, k). Pour toute fonction $f : \mathbb{R} \to \mathbb{R}^k$ continue, pour tout $\varepsilon \geqslant 0$ , pour tout $x \in \mathbb{R}^n$ , l'équation différentielle

$$(D^\varepsilon) \qquad y'_t = b\,(y_t) + \varepsilon\,\sigma\,(y_t)\,f_t$$

admet une unique solution maximale telle que $y_o = x$ ; si $]a, b[$ est l'intervalle de définition de cette solution maximale, on a $0 < b \leqslant +\infty$ et b ne peut être fini que si $\lim_{t\to b} ||y_t|| = +\infty$ . Nous appellerons b le temps d'explosion de la solution $y_t$.

Considérons un processus gaussien $X_t : \Omega \to \mathbb{R}^k$ à trajectoires continues, un nombre $\varepsilon > 0$ et le système

$$(S^\varepsilon) \qquad \begin{cases} y'_t = b\,(y_t) + \varepsilon\,\sigma\,(y_t)\,X_t \\ y_o = x \qquad \text{avec} \qquad x \in \mathbb{R}^n \text{ donné} \end{cases}$$

où pour chaque $\omega \in \Omega$ , $Y_t$ $(\omega)$ est défini comme l'unique solution maximale de $D^\varepsilon$ lorsqu'on remplace $f_t$ par $X_t$ $(\omega)$ et lorsqu'on impose $Y_o$ $(\omega) = x$.

La continuité des solutions de $D^\varepsilon$ en tant que fonctionnelles de f montre facilement que $Y_t$ est mesurable par rapport à la tribu du passé de $(X_s)$ au temps t. Notons en passant que $(Y_t)$ n'est pas gaussien en général !

Il s'agit de préciser le comportement des trajectoires de $(Y_t)$ lorsque $\varepsilon \to 0$, plus précisément d'évaluer la probabilité pour que la trajectoire de $(Y_t)$ , $t \in [0, 1]$ appartienne à un ensemble A de trajectoires "aberrantes" du point de vue du système (S). Ces probabilités vont tendre vers 0 avec $\varepsilon$, à des vitesses en $\exp\left(-\frac{C\ (A)}{\varepsilon^2}\right)$, et il s'agit d'évaluer C (A).

Remarquons que si $b \equiv 0$, $k = n$, $x = 0$, et $\sigma \equiv$ Identité, on a $Y_t = \varepsilon Z_t$ , où $Z_t = \int_o^t X_s\, ds$ est un processus gaussien à trajectoires continues, et le problème considéré a été résolu complètement au Chapitre II.

<u>Détail technique traditionnel</u> : lorsque t est supérieur ou égal au temps d'explosion $\tau$ $(\omega)$ de la trajectoire $Y_t$ $(\omega)$ du système $S^\varepsilon$ , on pose $Y_t$ $(\omega) = \delta$ , point à l'infini de $\mathbb{R}^n$.

Notons l'interprétation <u>heuristique</u> suivante : les "accroissements" $\Delta Y_t = Y_{t+\Delta t} - Y_t$ "vérifient" $\Delta Y_t = b\ (Y_t)\ \Delta t + \varepsilon\ \sigma\ (Y_t)\ \Delta Z_t$ . Autrement dit, l'accroissement déterministe $b\ (Y_t)\ \Delta t$ prévu par le système (S) est ici perturbé par la variable aléatoire $\varepsilon\ \sigma\ (Y_t)\ \Delta Z_t$ ("le bruit") dont la loi conditionnée par le passé au temps t est gaussienne centrée de matrice de covariance proportionnelle à $\varepsilon^2$.

Rappelons que les temps d'explosion des solutions de S , $D^\varepsilon$ , $S^\varepsilon$ sont certainement infinis si b et $\sigma$ sont à croissance sous linéaire, c'est-à-dire si $||b(y)|| + ||\sigma(y)|| \leqslant$ (constante) $(1 + ||y||)$ mais nous ne ferons pas cette hypothèse.

1.2. L'espace des trajectoires explosives : soit U un espace polonais localement compact, et soit $U \cup \delta$ son compactifié d'Alexandroff. Nous dirons qu'une fonction $f : [0, 1] \to U \cup \delta$ appartient à l'ensemble $\mathcal{E}(U)$ si f est continue, et si les relations $t_o \in [0, 1]$, $f(t_o) = \delta$ impliquent $f(u) = \delta$ pour $t_o \leqslant u \leqslant 1$.

Nous définirons le temps d'explosion $\tau(f)$ par $\tau(f) = \inf \{t \in [0, 1] \mid f(t) = \delta\}$, avec la convention $\tau(f) = +\infty$ si $f(t) \neq \delta$ pour tout $t \in [0, 1]$.

Nous dirons qu'une suite $f_n \in \mathcal{E}(U)$ converge vers $f \in \mathcal{E}(U)$ si $f_n$ converge vers f uniformément sur tout compact de $[0, \tau(f)[$ ; il est alors clair que $\tau(f) \leqslant \underline{\lim}_{n\to\infty} \tau(f_n)$. $\mathcal{E}(U)$ est ainsi muni d'une topologie à base dénombrable. Cet espace topologique $\mathcal{E}(U)$ appelé l'espace des trajectoires explosives, définies sur $[0, 1]$, à valeurs dans $U \cup \delta$.

Par construction, l'espace C (U) des fonctions continues de $[0, 1]$ dans U muni de la topologie de la convergence uniforme coïncide avec le sous-espace topologique $\{f \in \mathcal{E}(U) \mid \tau(f) = +\infty\}$ de $\mathcal{E}(U)$.

Pour $x \in U$ nous noterons $\mathcal{E}_x(U)$ le sous espace des $f \in \mathcal{E}(U)$ telles que $f(0) = x$. On définit de même $C_x(U) \subset C(U)$.

Notons que si $Y_t$ est le processus aléatoire issu de $x \in \mathbb{R}^n$ est solution du système $S^\varepsilon$, défini en 1.1, l'application qui à $\omega \in \Omega$ associe la trajectoire $t \to Y_t(\omega)$, $t \in [0, 1]$ est une application mesurable Y de $\Omega$ dans l'espace $\mathcal{E}_x(\mathbb{R}^n)$.

Justifions maintenant le choix de la topologie de $\mathscr{C}(\mathbb{R}^n)$.
Soient b et $\sigma$ des champs localement lipschitziens de vecteurs et de matrices définis sur $\mathbb{R}^n$. Soit x un point fixé de $\mathbb{R}^n$. Pour toute fonction continue $f : [0, 1] \to \mathbb{R}^k$ notons $B_x(f)$ la restriction à $[0, 1]$ de l'unique solution maximale g de l'équation différentielle $g'_t = b(g_t) + \sigma(g_t) f_t$ telle que $g(0) = x$ ; bien entendu si le temps d'explosion de g est plus petit que 1, on pose $g_t = \delta$ pour $t \geqslant \tau(g)$.

<u>Alors l'application</u> $B_x$ <u>de</u> $C(\mathbb{R}^k)$ <u>dans</u> $\mathscr{C}_x(\mathbb{R}^n)$ <u>est continue</u>. Pour le voir, il suffit d'utiliser des résultats classiques sur les équations différentielles ordinaires ; une démonstration directe se construit facilement en reprenant la méthode de Poincaré (remplacement par l'équation intégrale associée, résolue par itération).

1.3. <u>Lemme</u> : soient E, F deux espaces topologiques. Soit $\tilde{\lambda} : E \to [0, +\infty]$ une fonction s.c.i. telle que pour tout $a \geqslant 0$, a fini, l'ensemble $\tilde{C}_a = \{x \in E \mid \tilde{\lambda}(x) \leqslant a\}$ soit compact. Soit B une application de $E_o = \{x \in E \mid \tilde{\lambda}(x) \text{ fini}\}$ dans F, telle que la restriction de B à $\tilde{C}_a$ soit continue, pour tout $a \geqslant 0$, a fini. Définissons $\lambda : F \to [0, +\infty]$ par $\lambda(y) = \inf\{\tilde{\lambda}(x) \mid x \in E_o \text{ et } Bx = y\}$. Alors $\lambda$ est s.c.i. et $C_a = \{y \in F \mid \lambda(y) \leqslant a\}$ est compact pour tout $a \geqslant 0$, a fini.

<u>Preuve</u> : si $\lambda$ n'était pas s.c.i., il existerait une suite $y_n$ tendant vers y dans F telle que $a = \lim_n \lambda(y_n) < \lambda(y)$. Par définition de $\lambda$ il existe alors une suite $x_n \in E_o$ telle que $B x_n = y_n$ et vérifiant $\lim_{n\to+\infty} \tilde{\lambda}(x_n) = a$. Comme la suite $\tilde{\lambda}(x_n)$ est bornée, la suite $x_n$ est relativement compacte. Il existe donc $x \in E$ et une suite extraite de la suite $(x_n)$ - notée encore $x_n$ - telle que $x_n$ tende vers x. Comme $\tilde{\lambda}$ est s.c.i., on aura $\tilde{\lambda}(x) \leqslant \lim_n \tilde{\lambda}(x_n) = a$, et donc $x \in E_o$. La continuité de B sur l'ensemble

$\tilde{C}_b$ avec b convenable, b > a, entraîne alors

$$y = \lim_n y_n = \lim_n B\, x_n = B\, x$$

Par définition de $\lambda$ on en déduit $\lambda\ (y) \leqslant \tilde{\lambda}\ (x) \leqslant a$ , ce qui contredit l'hypothèse initiale sur a. Par suite $\lambda$ est bien s.c.i.

Pour $a \geqslant 0$, l'ensemble $C_a$ est par construction inclus dans $B\ (\tilde{C}_{a+1})$, qui est compact grâce aux hypothèses sur B et $\tilde{C}_{a+1}$. Comme $\lambda$ est s.c.i., l'ensemble $C_a$ est fermé par construction. Par suite, $C_a$ est compact.

1.4. Proposition : soit $X : \Omega \to E$ une variable aléatoire à valeurs dans un espace de Banach séparable E, de loi gaussienne centrée $\mu$. Soit $\lambda$ la transformée de Cramer de $\mu$. Soit F un espace topologique quelconque à base dénombrable, et soit $B : E \to F$ une application continue. Pour tout $\varepsilon > 0$ posons $Y^\varepsilon = B\ (\varepsilon\ X)$.

Définissons les fonctions

$$\lambda_B\ (y) = \inf \{\lambda\ (x) \mid x \in E\ ,\ B\ (x) = y\} \qquad \text{pour } y \in F$$

$$\Lambda_B\ (A) = \inf_{y \in A} \lambda_B\ (y) \qquad \text{pour } A \subset F$$

Alors on a pour tout borélien A de F

$$- \Lambda_B\ (\mathring{A}) \leqslant \varliminf_{\varepsilon \to 0} \varepsilon^2 \log P\ (Y^\varepsilon \in A) \leqslant \varlimsup_{\varepsilon \to 0} \varepsilon^2 \log P\ (Y^\varepsilon \in A) \leqslant - \Lambda_B\ (\bar{A})$$

De plus la fonction $\lambda_B : F \to [0, +\infty]$ est s.c.i. et pour tout $a \geqslant 0$ l'ensemble des $y \in F$ tels que $\lambda_B\ (y) \leqslant a$ est compact.

Preuve : puisque $P\ (Y^\varepsilon \in A) = P \{ \varepsilon\ X \in B^{-1}\ (A)\}$ , la proposition II.1.6 montre que lorsque $\varepsilon$ tend vers 0, les limites inf. et sup. de $\varepsilon^2 \log P\ (Y^\varepsilon \in A)$ sont dans l'intervalle $J = [- \Lambda\ (\mathring{C}), - \Lambda\ (\bar{C})]$ où $C = B^{-1}\ (A)$ et $\Lambda$ est la fonctionnelle de Cramer de $\mu$. La continuité de B donne $\bar{C} \subset B^{-1}\ (\bar{A})$ et $\mathring{C} \supset B^{-1}\ (\mathring{A})$ de sorte que J est contenu dans

$[- \Lambda (B^{-1} (\mathring{A})) , - \Lambda (B^{-1} (\bar{A}))]$ . Ceci prouve l'inégalité annoncée en 1.4 car $\Lambda_B = \Lambda \circ B^{-1}$ . On termine la preuve grâce au lemme 1.3 et à la proposition I.7.5.

1.5. <u>Théorème</u> : soient b un champ de vecteurs et $\sigma$ un champ de matrices (n, k) sur $\mathbb{R}^n$ ; on suppose b et $\sigma$ localement lipschitziens. Soit $X_t : \Omega \to \mathbb{R}^k$ un processus gaussien continu et soit $\lambda : C(\mathbb{R}^k) \to [0, +\infty]$ la transformée de Cramer de la loi des trajectoires de $X_t$ dans $C(\mathbb{R}^k)$. Soit $x \in \mathbb{R}^n$ et soit $Y_t^\varepsilon$ le processus solution du système dynamique perturbé $S^\varepsilon$ (cf. 1.1) issu de x au temps 0. Notons $\mathcal{E}_x(\mathbb{R}^n)$ l'espace des trajectoires explosives issues de x (cf. 1.2) et $Y^\varepsilon : \Omega \to \mathcal{E}_x(\mathbb{R}^n)$ la variable aléatoire définie par les trajectoires du processus $(Y_t^\varepsilon)$ sur $[0, 1]$ (cf. 1.2).

Pour $g \in \mathcal{E}_x(\mathbb{R}^n)$ considérons l'ensemble (éventuellement vide) $B_x^{-1}(g)$ des $f \in C(\mathbb{R}^k)$ telles que g soit sur $[0, 1]$ la solution maximale de l'équation différentielle $y_t' = b(y_t) + \varepsilon \sigma (y_t) f_t$ , issue de x au temps 0. Définissons

$$\tilde{\lambda}(g) = \inf \{ \lambda(f) \mid f \in B^{-1}(g) \} \qquad \text{pour } g \in \mathcal{E}_x(\mathbb{R}^n)$$

$$\text{et } \tilde{\Lambda}(A) = \inf_{g \in A} \tilde{\lambda}(g) = \Lambda \circ B^{-1}(A) \qquad \text{pour } A \subset \mathcal{E}_x(\mathbb{R}^n).$$

Alors pour toute partie borélienne A de $\mathcal{E}_x(\mathbb{R}^n)$ on a

$$- \tilde{\Lambda}(\mathring{A}) \leqslant \varliminf_{\varepsilon \to 0} \varepsilon^2 \log P(Y^\varepsilon \in A) \leqslant \varlimsup_{\varepsilon \to 0} \varepsilon^2 \log P(Y^\varepsilon \in A) \leqslant - \tilde{\Lambda}(\bar{A})$$

<u>Preuve</u> : ayant défini $B_x : C(\mathbb{R}^k) \to \mathcal{E}_x(\mathbb{R}^n)$ comme en 1.2 , on a évidemment $Y^\varepsilon = B_x(\varepsilon X)$. Comme $B_x$ est continue (cf. 1.2) le théorème est un corollaire direct de la proposition 1.4.

1.6. <u>Remarque</u> : supposons que k = n et que pour tout $x \in \mathbb{R}^n$ , $\sigma(x)$ soit une matrice inversible. Supposons de plus b et $\sigma$ à croissance sous-linéaire de sorte que tous les temps d'explosion sont infinis. On a alors $\tilde{\lambda}(g) = +\infty$ si g n'est pas de classe 1 sur $[0, 1]$ et $\tilde{\lambda}(g) = \lambda(f)$ où f est donnée par $f_t = \sigma(g_t)^{-1}(g'_t - b(g_t)) \qquad t \in [0, 1]$ .

De plus, on peut alors énoncer le résultat en remplaçant $\mathcal{C}_x(\mathbb{R}^n)$ par $C_x(\mathbb{R}^n)$.

Sous cette forme, le théorème a été prouvé par Ventcell et Freidlin [52] [23] .

Il est clair qu'on peut facilement étendre les hypothèses précédentes pour étudier des situations voisines. Par exemple, on peut supposer $(X_t)$ à trajectoires $L_2$ , et munir (dans les cas de non explosion) l'espace des trajectoires de $(Y^\varepsilon_t)$ de la norme $(||g'||^2_{L^2} + ||g||^2_{L^2})^{1/2}$ , ou bien supposer $(X_t)$ à trajectoires continues et munir les trajectoires de $(Y^\varepsilon_t)$ de la norme $(||\varphi'||_\infty + ||\varphi||_\infty)$ etc ...

## 2. CAS DES EQUATIONS DIFFERENTIELLES STOCHASTIQUES

2.1. Remarquons que le problème précédent concerne un processus $Y^\varepsilon_t$ dont les trajectoires vérifient "l'équation différentielle stochastique" $dy_t = b(y_t)\,dt + \varepsilon\,\sigma(y_t)\,d\,Z_t$ , où $Z_t$ est un processus gaussien à trajectoires continuement différentiables. Le cas des équations différentielles stochastiques ordinaire correspondrait à $Z_t = \beta_t$ où $\beta_t$ est un mouvement brownien.

Comme $\beta_t$ n'est pas différentiable, il n'est pas possible de considérer $Y^\varepsilon_t$ comme une fonctionnelle continue de "$\varepsilon\ \beta'_t$ " . Par contre il est possible de considérer $Y^\varepsilon_t$ comme une fonctionnelle $B(\varepsilon\ \beta)$ où $\beta$ désigne la trajectoire du Brownien dans $\mathbb{R}^k$ . En général, il n'est pas question d'espérer obtenir une telle $B : C(\mathbb{R}^k) \to C(\mathbb{R}^k)$ qui ait plus de régularité que la mesurabilité. Néanmoins le résultat formel reste correct !

En effet, le calcul formel de la fonctionnelle "de Cramer" $\tilde{\lambda}$ associée à l'équation stochastique $dy_t = b(y_t)\,dt + \varepsilon\,\sigma(y_t)\,d\beta_t$ donne "d'après les propositions II.3.6 et III.1.4, III.1.5"

$$\tilde{\lambda}(g) = \lambda\left[B^{-1}(g)\right] = \frac{1}{2}\left\|\frac{d}{dt}B^{-1}(g)\right\|_{L^2}^2$$

où la relation $f = B^{-1}(g)$ équivaut à

$$g_t = \int_0^t \left[b(g_s) + \sigma(g_s)\,f'_s\right] ds$$

ce qui "entraîne" finalement

$$\tilde{\lambda}(g) = \frac{1}{2}\int_0^1 \left\|\sigma^{-1}(g_s)\left[g'_s - b(g_s)\right]\right\|^2 ds$$

Nous nous sommes placés, pour esquisser ce calcul heuristique, dans le cas où $\sigma$ est inversible et le temps d'explosion infini. Dans ce cadre, le résultat obtenu est correct et a été essentiellement découvert par Ventall-Freidlin [53] . Notre méthode de démonstration est très différente de la leur et permettra d'ailleurs de traiter le cas où $\sigma$ n'est pas inversible, et de ne pas exclure les possibilités d'explosion du processus.

Enfin pour pouvoir aborder le cas des diffusions sur des variétés non compactes, nous serons amenés à généraliser légèrement le problème de perturbations dont la forme 1.1 n'est pas invariante par changement de coordonnées.

L'étape cruciale de notre approche est le théorème 2.3 qui prouve que pour $\varepsilon$ petit l'application B ressemble beaucoup à une application continue.

## 2.2. Le modèle de système dynamique perturbé

On considère sur un ouvert U de $\mathbb{R}^n$ , un champ $x \to \sigma(x)$ de matrice $(n, k)$ , une famille de champs de vecteurs $x \to b_\varepsilon(x)$ indexés par $\varepsilon > 0$ et un champ de vecteurs $x \to b(x)$.

On suppose que :

$(H_1)$ $\sigma$ est de classe 1 sur U

$(H_2)$ tous les champs $b_\varepsilon$ , $\varepsilon > 0$ et le champ b sont localement lipschitziens sur U

$(H_3)$ $\lim\limits_{\varepsilon \to} b_\varepsilon (x) = b (x)$ uniformément sur tout compact de U

On notera $y_t^\varepsilon$ la solution de l'équation différentielle stochastique sur U

$(E^\varepsilon)$ $dy_t = b_\varepsilon (y_t) dt + \varepsilon \sigma (y_t) d\beta_t$

où $\beta_t : \Omega \to \mathbb{R}^k$ est un mouvement brownien k-dimensionnel.

Appelons $\mathcal{E}(U)$ , $\mathcal{E}_x (U)$ les espaces de trajectoires explosives définies sur $[0, 1]$ à valeurs dans $U \cup \delta$ . On sait que sous les hypothèses $H_1$, $H_2$ , il existe pour chaque $x \in U$ une unique solution $y_t^\varepsilon : \Omega \to U \cup \delta$ définie sur $[0, 1]$ , et à trajectoires appartenant à $\mathcal{E}_x (U)$. (cf. par exemple Azencott [2] en ce qui concerne les temps d'explosion). C'est cette solution naturelle que nous considérerons toujours dans la suite. Le temps d'explosion de la trajectoire $y^\varepsilon$ sera noté $\tau (y^\varepsilon)$ , tandis que $y^\varepsilon : \Omega \to \mathcal{E}_x (U)$ est la variable aléatoire associant à $\omega \in \Omega$ la trajectoire $t \to y_t^\varepsilon (\omega)$.

Le système $(E^\varepsilon)$ peut être considéré comme une petite perturbation du système dynamique $y_t' = b (y_t)$. On notera que <u>le modèle</u> $(E^\varepsilon)$ <u>est "invariant" par changement de coordonnées</u>, c'est-à-dire que si $\varphi : U \to V$ est un difféomorphisme de U sur un ouvert V de $\mathbb{R}^n$ , le processus $z_t^\varepsilon = \varphi (y_t^\varepsilon)$ vérifiera une équation du type $(E^\varepsilon)$ associée à des coefficients $\tilde{b}_\varepsilon$ , $\tilde{b}$ , $\tilde{\sigma}$ vérifiant encore $(H_1 \; H_2 \; H_3)$.

Nous utiliserons les espaces de fonctions $C(U)$ , $C_x(U)$ , $\mathcal{E}(U)$ , $\mathcal{E}_x(U)$ définis en 1.2.

Nous définirons $\mathcal{E}^o(U)$ comme l'ensemble des $g \in \mathcal{E}(U)$ telles que, pour tout $T \leqslant 1$ , $T < \tau(g)$ , la dérivée au sens de Lebesgue $g'$ de $g$ existe et soit de carré intégrable sur $[0, T]$ . De même, nous poserons $C^o(U) = C(U) \cap \mathcal{E}^o(U)$ ; nous munirons $C^o(U)$ et $\mathcal{E}^o(U)$ des topologies induites par $C(U)$ , $\mathcal{E}(U)$.

Pour $g, h \in \mathcal{E}(U)$ , $T < \tau(g)$ et $0 \leqslant S \leqslant T \leqslant 1$ , nous noterons

$$d_{S,T}(g, h) = \sup_{S \leqslant t \leqslant T} |g - h|$$

Pour $g \in \mathcal{E}(U)$ , $r > 0$ , $T < \tau(g)$ et $0 \leqslant S \leqslant T \leqslant 1$ , nous noterons $\Gamma_{S,T}(g, r)$ le tube fermé d'axe $g$ , de rayon $r$.

$$\Gamma_{S,T}(g, r) = \{h \in \mathcal{E}(U) \mid d_{S,T}(g, h) \leqslant r\}$$

Pour $f \in C^o(U)$ , l'équation différentielle

$$\text{(D)} \qquad g'_t = b(g_t) + \sigma(g_t) f'_t$$

pour presque tout $t \in [0, 1 \wedge \tau(g)[$ , où $g \in \mathcal{E}^o(U)$ jouera un rôle essentiel dans la suite.

2.3. Proposition (hypothèses et notations 2.2) : pour tout $x \in U$ et $f \in C^o(U)$ , l'équation différentielle (D) admet une unique solution $g \in \mathcal{E}^o(U)$ telle que $g(0) = x$ . Nous poserons alors $g = B_x(f)$. L'application $B : U \times C^o(U) \to \mathcal{E}^o(U)$ n'est en général pas continue. Cependant, si on pose $C_a = \{f \in C^o(U) \mid \int_0^1 |f'_t|^2 \, dt \leqslant a\}$ , la restriction de $B$ à $U \times C_a$ est continue, pour tout $a \geqslant 0$.

De plus pour toute paire $K$ , $L$ de compacts de $U$ vérifiant $K \subset \overset{\circ}{L}$ , pour tout $a > 0$ , on peut trouver des nombres $T > 0$, $C > 0$ ayant les propriétés suivantes :

(i) quels que soient $v \in K$ , $f \in C_a$ , la fonction $h = B_v(f)$ vérifie $\tau(h) > T$ et $h([0, T]) \subset L$

(ii) quels que soient $v, w \in K$ , $f \in C_a$ , et $S \leqslant T$ , on a $d_{0,S}[B_v(f), B_w(f)] \leqslant e^{CS}|v - w|$

Preuve : on remplace (D) par une équation intégrale que l'on résout par itération. Tous les arguments sont des modifications mineures des preuves classiques.

2.4. Théorème : (hypothèses et notations 2.2) Sur un ouvert U de $\mathbb{R}^n$ , considérons le système dynamique perturbé $(E^\varepsilon)$, vérifiant les hypothèses $(H_1\ H_2\ H_3)$, de trajectoires $y^\varepsilon \in \mathcal{C}(U)$. Soit (D) l'équation différentielle ordinaire associée à $(E^\varepsilon)$ et soit $B : U \times C^o(U) \longrightarrow \mathcal{C}^o(U)$ l'application (déterministe) associée à (D) en 2.3. Soit $\beta : \Omega \to C(\mathbb{R}^k)$ la trajectoire (aléatoire) du Brownien qui intervient dans l'écriture de $(E^\varepsilon)$.

Fixons un compact K de U et des nombres $a > 0$ , $T \in ]0, 1]$. Alors à toute paire de nombres $\rho > 0$ , $R > 0$ , on peut associer des nombres $\varepsilon_o > 0$ , $\alpha > 0$ , $r > 0$ ayant la propriété suivante :

Quels que soient $x \in K$ , $f \in C^o(\mathbb{R}^k)$ , $g = B_x(f)$ vérifiant $\int_0^T |f_t'|^2\, dt \leqslant a$ et $g[0, T] \subset K$ , les relations $\varepsilon \leqslant \varepsilon_o$ et $|y_o^\varepsilon - x| \leqslant r$ (P - p.s.) entraînent

$$P\left[d_{0,T}(\varepsilon\beta, f) \leqslant \alpha \text{ et } d_{0,T}(y^\varepsilon, g) > \rho\right] \leqslant \exp\left(-\frac{R}{\varepsilon^2}\right)$$

2.5. Remarque : l'énoncé 2.4 garantit que, pour $\varepsilon$ et $\alpha$ petits, si l'on néglige un évènement de probabilité inférieure à $\exp(-\frac{R}{\varepsilon^2})$ , la présence de $\varepsilon\beta$ dans le tube $A = \Gamma_{0,T}(f, \alpha)$ d'axe f et de rayon $\alpha$, implique la présence de $y^\varepsilon$ dans le tube d'axe $g = B_x(f)$ et de rayon $\rho$ , ceci à condition que $y_o^\varepsilon$ soit assez proche de $g_o = x$ et que $T < \tau(g)$ = temps d'explosion de g.

La proposition 2.4 décrit donc une propriété de "continuité" de l'application mesurable qui à $\varepsilon\beta$ associe $y^\varepsilon$, lorsque $x = y_o^\varepsilon$ est fixé. L'énoncé est alourdi par la nécessité (voir chapitres IV et V) de préciser les conditions d'uniformité de cette "continuité" , et se simplifie si on se borne à considérer x fixé, $f \in C^o(\mathbb{R}^k)$ fixé , et $T < \tau\ [B_x(f)]$.

Notons que d'après la proposition II.3.6, "on a"

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log P(\varepsilon\beta \in A) = -\Lambda(A)$$

où $\Lambda(A) = \inf_{\varphi\in A} \lambda(\varphi)$ et $\lambda(\varphi) = \frac{1}{2}\int_0^T ||\varphi_t'||^2\, dt$

La norme $||v||$ est ici la norme euclidienne usuelle. La probabilité $P\ [\varepsilon\beta \in A]$ a donc aussi un comportement en $\exp(-\frac{cte}{\varepsilon^2})$ ; mais pour R arbitrairement grand les probabilités en $\exp(-\frac{R}{\varepsilon^2})$ sont négligeables devant $P\ [\varepsilon\beta\in A]$ . Ceci fait comprendre le corollaire suivant, qui donne une réponse précise à une question posée par B. Roynette pendant les "Saint-Flour sessions".

2.6. <u>Corollaire</u> (mêmes hypothèses et notations que 2.4) : fixons x, f et donc $B_x(f) = g$.

Alors, pour tout $\rho > 0$ et $T < \tau(g)$ , il existe $\alpha_o > 0$ tel que la relation $\alpha \leqslant \alpha_o$ entraîne

$$\lim_{\varepsilon\to 0} P\ [y^\varepsilon \in \Gamma_{0,T}(g,\rho) \mid \varepsilon\beta\in\Gamma_{0,T}(f,\alpha)] = 1$$

Plus précisément, pour tout $\rho > 0$ , $T < \tau(g)$ , $R > 0$ , il existe $\alpha_o > 0$ et une fonction $\varepsilon_o(\alpha) > 0$ tels que

$$1 \geqslant P\ [y^\varepsilon \in \Gamma_{0,T}(g,\rho) \mid \varepsilon\beta\in\Gamma_{0,T}(f,\alpha)| \geqslant 1 - \exp(-\frac{R}{\varepsilon^2})$$

pour $\alpha \leqslant \alpha_o$ et $\varepsilon \leqslant \varepsilon_o(\alpha)$.

Preuve : pour alléger l'écriture, notons $\Gamma_{0,T}\ (g,\ \rho) = \Gamma_g$ et $\Gamma_{0,T}\ (f,\ \alpha) = \Gamma_f$ . Soit $p = P\left[y^\varepsilon \in \Gamma_g \mid \varepsilon\,\beta \in \Gamma_f\right]$ . Alors

$$(1) \qquad p = 1 - \frac{P\left\{(\varepsilon\,\beta \in \Gamma_f) \cap (y^\varepsilon \notin \Gamma_g)\right\}}{P\ (\varepsilon\,\beta \in \Gamma_f)}$$

Soit $a = \frac{1}{2} \int_0^T |f'_t|^2\, dt$. D'après II.3.6, on a

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log P\ [\varepsilon\,\beta \in \Gamma_f] = -\Lambda\,[\Gamma_f] \geqslant -a$$

et donc pour $\varepsilon \leqslant \varepsilon_1\ (\alpha)$

$$(2) \qquad P\ [\varepsilon\,\beta \in \Gamma_f] \geqslant \exp\left(-\frac{2a}{\varepsilon^2}\right)$$

Donnons-nous $R > 0$ arbitraire, et posons $R_1 = R + 2\,a$. Le théorème 2.4 (appliqué à $\rho$ , $T$, $R_1$) fournit $\varepsilon_o\ (R + 2a)$ et $\alpha_o\ (R + 2a)$ tels que la majoration 2.4 s'applique dès que $\varepsilon \leqslant \varepsilon_o$ , $\alpha \leqslant \alpha_o$ . D'après (1), (2), on a alors, pourvu que $\alpha \leqslant \alpha_o$ , $\varepsilon \leqslant \varepsilon_o \wedge \varepsilon_1$

$$p \geqslant 1 - \frac{\exp\left(-\frac{R + 2a}{\varepsilon^2}\right)}{\exp\left(-\frac{2a}{\varepsilon^2}\right)} = 1 - \exp\left(-\frac{R}{\varepsilon^2}\right)$$

2.7. Preuve du théorème 2.4 : donnons-nous deux compacts $K$ , $L$ de $U$ tels que $K \subset \overset{\circ}{L}$ . Notons $K^r$ la réunion des boules fermées de centre $v \in K$ et de rayon $r > 0$. Fixons $r_o > 0$ tel que $K^{3r_o} \subset \overset{\circ}{L}$ , $a > 0$, et $T \in\, ]0,\ 1]$. Posons $C_a = \{f \in C^o\ (\mathbb{R}^k) \mid \int_0^T |f'_t|^2\ dt \leqslant a\}$ .

D'après la proposition 2.3, il existe $S_o \in\, ]0,\ T]$ tel que pour tout $f \in C_a$ et tout $v \in K^{r_o}$ , la fonction $h = B_v\ (f)$ (notations 2.3) soit définie sur $[0,\ S_o]$ et vérifie $h\ ([0,\ S_o]) \subset K^{2r_o}$ . De plus (cf. 2.3) pour $S_o$ assez petit, il existe une constante $C > 0$ telle que :

$$(2) \qquad d_{0,S}\ [B_v\ (f)\ ,\ B_w\ (f)] \leqslant e^{CS}\ |v - w|$$

dès que $S \leqslant S_o$ , $v$ , $w \in K^{r_o}$ , $f \in C_a$.

Soient $\rho \in ]0, r_o]$ et $v \in K^{r_o}$ . Posons $g = B_v(f)$. Les définitions données en 2.2, 2.3 montrent que $\varphi_t = y_t^\varepsilon - g_t$ vérifie pour $t \leqslant S_o$ , $t \leqslant \tau\ (y^\varepsilon)$

$$\varphi_t = \int_o^t a_s\, d\beta_s + z_t + \int_o^t h_s\, ds + (y_o^\varepsilon - v)$$

avec les notations

$$a_s = \varepsilon\ [\sigma\ (y_s^\varepsilon) - \sigma\ (g_s)]$$

$$h_s = b_\varepsilon\ (y_s^\varepsilon) - b\ (g_s)$$

$$z_t = \int_o^t \sigma\ (g_s)\ [\varepsilon\ d\ \beta_s - f'_s\ ds]$$

La fonction $m_s = \frac{d}{ds}\ [\sigma(y_s)]$ est dans $L_2\ [0, S_o]$ et une intégration par parties donne

$$z_t = \sigma\ (g_t)\ [\varepsilon\ \beta_t - f_t] - \sigma\ (g_o)\ [\varepsilon\ \beta_o - f_o] - \int_o^t m_s\ (\varepsilon\ \beta_s - f_s)\ ds$$

Pour $\alpha > 0$ , $\rho \in ]0, r_o]$ introduisons les temps de sortie

$$\eta = \inf \{s \in [0, 1]\ |\ |\varepsilon\ \beta_s - f_s| \geqslant \alpha\}$$

$$\Theta = \inf \{s \in [0, 1]\ |\ |y_s^\varepsilon - g_s| \geqslant \rho\}$$

Pour $t \leqslant \Theta \wedge S_o$ , $y_t^\varepsilon$ et $g_t$ restent dans $K^{3r_o} \subset L$ , et il existe donc une constante majorant sûrement

$$|\sigma\ (g_t)|,\ \frac{|\sigma\ (y_t^\varepsilon) - \sigma\ (g_t)|}{|y_t^\varepsilon - g_t|}\ ,\ \frac{|b\ (y_t^\varepsilon) - b\ (g_t)|}{|y_t^\varepsilon - g_t|}\ ,\ \text{et } \int_o^t |m_s|^2\ ds.$$

La convergence de $b_\varepsilon$ vers $b$ est uniforme sur les compacts (hypothèse $H_3$), de sorte que pour tout $\gamma > 0$ , il existe $\varepsilon_o\ (\gamma) > 0$ tel que

$$|b_\varepsilon\ (y_t^\varepsilon) - b\ (y_t^\varepsilon)| \leqslant \gamma \qquad \text{pour} \qquad \varepsilon \leqslant \varepsilon_o\ ,\ t \leqslant \Theta\ \wedge\ S_o$$

Imposons donc les conditions

$$(2) \qquad v \in K^{r_o} \ ; \ f \in C_a \ ; \quad \varepsilon \leqslant \varepsilon_o = \varepsilon_o\,(\gamma) \ ; \quad \rho \leqslant r_o \ ;$$

$$|y_o^\varepsilon - v| \ \leqslant \frac{\rho}{R} \quad (P\text{-p.s.})$$

On obtient ainsi les majorations suivantes (où C est une constante adéquate) :

$$(3) \qquad |a_t| \leqslant \ \varepsilon\, C_\rho \qquad\qquad (\text{pour } t \leqslant \Theta \ \wedge S_o) \ ;$$

$$|\int_o^t h_s\, ds| \leqslant \ \int_o^t |b_\varepsilon\,(y_s^\varepsilon) - b\,(y_s^\varepsilon)| \ \ ds + \int_o^t |b\,(y_s^\varepsilon) - b\,(g_s)| \ ds$$

$$\leqslant \ \gamma\, t + C\,\rho\, t \quad (\text{pour } t \leqslant \Theta \wedge S_o) \ ;$$

$$|z_t| \leqslant C^\alpha \qquad\qquad (\text{pour } t \leqslant \eta \quad \text{et} \quad t \leqslant \Theta \wedge S_o).$$

Posons

$$w_t = \int_o^t a_s\, d\beta_s = \varphi_t - z_t - \int_o^t h_s\, ds + (v - y_o^\varepsilon)$$

de sorte que

$$(4) \qquad |w_t| \geqslant |\varphi_t| \ - |z_t| - |\int_o^t h_s\, ds| - \frac{1}{5}\,\rho$$

Soit $S \leqslant S_o$. L'évènement $A = \{\Theta \leqslant \ S \leqslant \eta\}$ implique $|\varphi_\Theta| = \rho$ , et donc d'après (3), (4) implique

$$|w_\Theta| \ \geqslant \ \rho \ - C\alpha \ - \quad \gamma S - C\,S\,\rho - \frac{1}{5}\,\rho$$

Imposons les conditions (2) et

$$(5) \qquad C\,\alpha \leqslant \frac{1}{5}\,\rho \ \ ; \ S \leqslant S_o \ ; \ C\,S \leqslant \frac{1}{5} \ ; \ \ \gamma\,S_o = \frac{1}{5}\,\rho \quad \text{pour}$$

conclure que $A \subset \{|w_\Theta| \ \geqslant \frac{1}{5}\,\rho\}$ . En notant $w_\Theta^i$ les coordonnées de $w_\Theta$ , $i = 1 \ldots n$ , on aura donc

$$(6) \qquad P\,(A) \leqslant \sum_{i=1}^{n} \ P \ \{|w_\Theta^i| \ \geqslant \frac{1}{5}\,\rho\} \ .$$

Pour chaque $i = 1 \ldots n$, $w^i_t$ est une martingale locale dont le processus croissant $d_t$ est évidemment majoré par $\int_0^t |a_s|^2\, ds$. Donc (cf. Priouret [40]) $Z_t = \exp \left[M\, w^i_t - \frac{M^2}{2}\, d_t\right]$ est une martingale locale pour tout $M \in \mathbb{R}$. Par suite $Z_{t \wedge \Theta}$ est une martingale bornée pour $t \in [0, T]$, et $E\,(Z_{S \wedge \Theta}) = E\,(Z_o) = 1$. On peut alors écrire

$$P\,(w^i_{S \wedge \Theta} \geqslant \frac{1}{5}\,\rho) = P\,(M\, w^i_{S \wedge \Theta} - \frac{M^2}{2}\, d_{S \wedge \Theta} \geqslant \frac{M}{5}\,\rho - \frac{M^2}{2}\, d_{S \wedge \Theta})$$

tandis que grâce à (5)

$$d_{S \wedge \Theta} \leqslant \int_0^{S \wedge \Theta} |a_s|^2\, ds \leqslant \varepsilon^2\, S\, C^2\, \rho^2$$

d'où finalement (calcul classique ; voir [40] [49])

$$P\,(w^i_{S \wedge \Theta} \geqslant \frac{1}{5}\,\rho) \leqslant P\,\{Z_{S \wedge \Theta} \geqslant \exp\,(\frac{M}{5}\,\rho - \frac{M^2}{2}\,\varepsilon^2\, S\, C^2\, \rho^2)\}$$

$$\leqslant E\,(Z_{S \wedge \Theta})\, \exp\,(- \frac{M}{5}\,\rho + \frac{M^2}{2}\,\varepsilon^2\, S\, C^2\, \rho^2)$$

Choisissons $M = \frac{1}{5}\; \frac{\rho}{\varepsilon^2\, S\, C^2\, \rho^2}$ pour conclure que

$$P\,(w^i_{S \wedge \Theta} \geqslant \frac{1}{5}\,\rho) \leqslant \exp\,(- \frac{1}{50\, S\, C^2\, \varepsilon^2})$$

On a une majoration analogue pour $P\,(w^i_{S \wedge \Theta} \leqslant - \frac{1}{5}\,\rho)$ ce qui donne finalement, grâce à (6)

$$P\,(A) \leqslant 2\, n\, \exp\,(- \frac{1}{50\, S\, C^2\, \varepsilon^2})$$

quand les conditions (2) et (5) sont satisfaites.

Ce résultat se reformule de la façon suivante ; l'ensemble des conditions (7) ci-dessous

$$(7) \quad \left\{ \begin{array}{l} v \in K^{r_o}\ ;\ f \in C_a\ ;\ g = B_v\,(f)\ ;\ |y^\varepsilon_o - v| \leqslant \frac{1}{5}\,\rho\ \ (P\text{-p.s.}) \\ \rho \leqslant r_o;\ \varepsilon \leqslant \varepsilon_1\,(\rho) = \varepsilon_o\,(\frac{\rho}{5\, S_o})\ ;\ \alpha \leqslant \alpha_o\,(\rho) = \frac{\rho}{5\, C}\ ; \\ S \leqslant S_1 = S_o \wedge \frac{1}{5\, C}\ ; \end{array} \right.$$

impliquent l'inégalité (cf. notations 2.2)

(8) $P\left\{d_{0,S}(\varepsilon\beta, f) \leqslant \alpha \text{ et } d_{0,S}(y^{\varepsilon}, g) \geqslant \rho\right\} \leqslant 2\, n \exp\left(-\frac{1}{50\, S\, C^2\, \varepsilon^2}\right)$

Un simple changement de notations et la propriété de Markov permettent de translater ce résultat par $t \in [0, 1]$ ; notons $\mathcal{F}_t$ la $\sigma$-algèbre des évènements (browniens) antérieurs au temps t ; alors les conditions (9) ci-dessous

(9) $f \in C_a$ ; $\rho \leqslant r_o$ ; $\varepsilon \leqslant \varepsilon_1(\rho)$ ; $\alpha \leqslant \alpha_o(\rho)$ ; $S \leqslant S_1$ ; $t + S \leqslant T$

impliquent que sur l'ensemble $\{\omega \in \Omega \mid y_t^{\varepsilon}(\omega) \in K^{r_o}\}$ on a l'inégalité P-presque sûre

(10) $P\left[d_{t,t+S}(\varepsilon\beta, f) \leqslant \alpha \text{ et } d_{t,t+S}(y^{\varepsilon}, G^t) \geqslant \rho \mid \mathcal{F}_t\right] \leqslant 2\, n \exp\left(-\frac{1}{50\, S\, C^2\, \varepsilon^2}\right)$

où $G^t$ est la trajectoire (aléatoire) définie sur $[t, t+S]$ comme l'unique solution de l'équation différentielle $\psi'_s = b(\psi_s) + \sigma(\psi_s)\, f'_s$, issue du point $y_t^{\varepsilon}$ au temps t.

Donnons-nous maintenant $T \in ]0, 1]$ , $a > 0$, et deux compacts K, L de U tels que $K \subset \overset{\circ}{L}$ . A partir de T, a, K, L on peut déterminer $r_o$ , $S_1$ et les fonctions $\varepsilon_1(\rho)$ , $\alpha_o(\rho)$ comme ci-dessus. Soit f dans $C^o(\mathbb{R}^k)$ et N un entier arbitraire. Posons $g = B_x(f)$ et $S = \frac{T}{N}$ .

Supposons vérifiées les conditions

(11) $f \in C_a$ ; $g([0, T]) \subset K$ ; $\varepsilon \leqslant \varepsilon_1(\rho)$ ;
$\alpha \leqslant \alpha_o(\rho)$ ; $S \leqslant S_1$ ;

Définissons des évènements $A_k$ , $k = 0 \ldots (N-1)$ , par

$A_k = \{d_{kS,(k+1)S}(y^{\varepsilon}, G^{kS}) \leqslant \rho \ ; (k+1)\, S < \tau(y^{\varepsilon})\}$ , $k = 1 \ldots N-1$

$A_o = \{d_{0,S}(y^{\varepsilon}, g) \leqslant \rho\}$

Montrons par récurrence sur k que sous la condition (14) ci-dessous on aura, pour $k \in [1, N]$

$$(12) \qquad A_0 \cap \ldots \cap A_{k-1} \subset \left\{ d_{0,kS}\,(y^{\varepsilon}, g) \leqslant \rho \; \frac{e^{kCS} - 1}{e^{CS} - 1} \right\}$$

La propriété est trivialement vraie pour $k = 1$. Supposons la vraie pour un entier $k \in [1, N-1]$.

Soit $w \in A_0 \cap \ldots \cap A_k$ ; l'hypothèse de récurrence donne

$$(13) \quad |y^{\varepsilon}_{kS} - g_{kS}| \leqslant \rho \; \frac{e^{kCS} - 1}{e^{CS} - 1} \leqslant \rho \; \frac{e^{CT} - 1}{e^{CT/N} - 1} \leqslant \rho \, N \; \frac{e^{CT} - 1}{CT}$$

et donc si on impose

$$(14) \qquad \rho \; N \; \frac{e^{CT} - 1}{CT} \leqslant r_o$$

on aura $y^{\varepsilon}_{kS} \in K^{r_o}$. Par suite $G^{kS}$ est à valeurs dans U sur $[kS, (k+1)\,S]$, et la propriété de continuité (1) donne (par changement de notations)

$$d_{kS,(k+1)S} \; (G^{kS}, g) \leqslant e^{CS} \; |y^{\varepsilon}_{kS} - g_{kS}|$$

de sorte que, puisque $w \in A_k$

$$d_{kS,(k+1)S} \; (y^{\varepsilon}, g) \leqslant \rho \; + e^{CS} \; |y^{\varepsilon}_{kS} - g_{kS}|$$

Rapprochons cette inégalité de (12) (13) pour conclure que

$$d_{0,(k+1)S} \; (y^{\varepsilon}, g) \leqslant \rho \; + e^{CS} \; \rho \; \frac{e^{kCS} - 1}{e^{CS} - 1} = \rho \; \frac{e^{(k+1)\,CS} - 1}{e^{CS} - 1}$$

ce qui achève la preuve de (12).

Donnons nous $r > 0$ arbitraire et imposons la condition

$$(15) \qquad \rho \; N \; \frac{e^{CT} - 1}{CT} < r \wedge r_o$$

Alors (1 2) donne pour $k = N$

$$A_0 \cap \ldots \cap A_{N-1} \subset \{d_{0,T}(y^\varepsilon, g) < r\}$$

et par suite

$$(16) \qquad \{d_{0,T}(y^\varepsilon, g) \geqslant r\} \subset (A_o^c \cup \ldots \cup A_{N-1}^c) = \bigcup_{0 \leqslant k \leqslant N-1} F_k$$

avec $F_o = A_o^c$ , $F_k = A_0 \cap \ldots \cap A_{k-1} \cap A_k^c$ , $k = 1 \ldots N-1$.

D'après (12) (15), pour $k \in [1, N-1]$ on a $F_k \subset \{y_{kS}^\varepsilon \in K^{r_o}\} \cap A_k^c$ et donc, grâce à la définition de $A_k$ , $F_k$ est inclus dans $\{y_{kS}^\varepsilon \in K^{r_o}\} \cap \{d_{kS,(k+1)S}(y^\varepsilon, G^{kS}) > \rho\}$ .

Imposons les conditions (11) (15) et

$$(17) \qquad |y_o^\varepsilon - g_o| \leqslant \frac{1}{5}\rho \qquad P - p.s.$$

pour obtenir directement $F_o \subset \{|y_o^\varepsilon - g_o| \leqslant \frac{1}{5}\rho\} \cap \{d_{0,S}(y^\varepsilon, g) > \rho\}$

Il suffit d'appliquer les inégalités (8) et (10) pour conclure que, pour $k = 0, 1, \ldots N-1$, on a

$$(18) \qquad P\left[F_k \cap \{d_{0,T}(\varepsilon\beta, f) \leqslant \alpha\}\right] \leqslant 2\,n \exp\left(-\frac{1}{50\,S\,C^2\,\varepsilon^2}\right)$$

Posons $Q = \{d_{0,T}(\varepsilon\beta, f) \leqslant \alpha$ et $d_{0,T}(y^\varepsilon, g) \geqslant r\}$ .

De (16) (18) on déduit, puisque $S = \frac{T}{N}$

$$P(Q) \leqslant 2\,n\,N \quad \exp\left(-\frac{N}{50\,T\,C^2\,\varepsilon^2}\right)$$

Comme $s\,e^{-s} \leqslant e^{-s/2}$ pour tout $s \geqslant 0$ , ceci devient

$$P(Q) \leqslant 100\,n\,T\,C^2\,\varepsilon^2 \exp\left(-\frac{N}{100\,T\,C^2\,\varepsilon^2}\right)$$

Donnons-nous $R > 0$ arbitraire ; imposons les conditions (11) (15) (17) et

$$(19) \qquad \frac{N}{100\,T\,C^2} \geqslant R \quad ; \quad 100\,n\,T\,C^2\,\varepsilon^2 \leqslant 1$$

pour obtenir facilement $P(Q) \leqslant \exp(-\frac{R}{\varepsilon^2})$.

Etudions l'ensemble de conditions (11) (15) (17) (19) (avec $S = \frac{T}{N}$). A partir de $T \in ]0, 1]$, $a > 0$ et de deux compacts arbitraires $K$, $L$ tels que $g[0, T] \subset K$ et $K \subset \overset{\circ}{L}$, $L \subset U$ on a déterminé $r_o$, $S_1$ et les fonctions $\varepsilon_1(\rho)$, $\alpha_o(\rho)$.

Fixons un entier $N_o(R)$ tel que

$$N_o \geqslant \max\left(100\, T\, C^2 R\, ,\ \frac{T}{S_1}\right)$$

Déterminons ensuite $\rho_o(r, R)$ par

$$\rho_o = \min\left[\frac{1}{2\, N_o}\ \frac{C\, T}{e^{CT} - 1}\ (r \wedge r_o),\ r_o\right]$$

Posons enfin

$$\varepsilon_2(r, R) = \min\left[\varepsilon_1(\rho_o), \frac{1}{\sqrt{100\, n\, T\, C^2}}\right]\ ;\ \alpha_1(r, R) = \alpha_o(\rho_o)$$

Les conditions suivantes, où $g = B_x(f)$

$$(20) \qquad \varepsilon \leqslant \varepsilon_2(r, R)\ ;\ \alpha \leqslant \alpha_1(r, R)\ ;\ f \in C_a\ ;\ g[0, T] \subset K$$

$$|y_o^\varepsilon - x| \leqslant \frac{1}{5}\ \rho_o(r, R) \qquad (P - p.s.)$$

impliquent alors (11) (15) (17) (19), et donc garantissent $P(Q) \leqslant \exp(-\frac{R}{\varepsilon^2})$, ce qui prouve le théorème 2.4.

## 2.8. Dualité entre formes quadratiques dégénérées :

Soit $Q : \mathbb{R}^n \longrightarrow [0, +\infty[$ une forme quadratique positive au sens large. Appelons forme quadratique duale $Q^* : \mathbb{R}^n \to [0, +\infty]$ la fonction définie par

$$(1) \qquad \frac{1}{2}\, Q^*(v) = \sup_{t \in \mathbb{R}^n}\left[< t, v > - \frac{1}{2}\, Q(t)\right] \qquad ,\ v \in \mathbb{R}^n$$

Notons que si Q est de noyau non nul, $Q^*$ peut atteindre la valeur

On a en fait

(2) $\{Q^* (v) \text{ fini}\}$ équivaut à $\{v \text{ est orthogonal au noyau de } Q\}$

Ecrivons Q sous la forme

(3) $Q(v) = ||\sigma^* v||^2 = \langle v , \sum v \rangle$ pour tout $v \in \mathbb{R}^n$

où $\sigma : \mathbb{R}^k \to \mathbb{R}^n$ et $\sum = \sigma \sigma^* : \mathbb{R}^n \to \mathbb{R}^n$ sont des opérateurs linéaires. On a alors :

(4) $\{v \in \mathbb{R}^n \mid Q^* (v) \text{ fini}\} = \sigma(\mathbb{R}^k) = \sum(\mathbb{R}^n)$

(5) $Q^* (v) = \inf \{||w||^2 \mid w \in \mathbb{R}^k \text{ et } \sigma w = v\}$

(6) Les propriétés (i) ... (v) sont équivalentes :

(i) Q définie positive

(ii) $\sigma$ surjective

(iii) $\sum = \sigma \sigma^*$ inversible

(iv) $Q^*$ est partout finie sur $\mathbb{R}^n$

(v) $Q^* (v) = \langle v , \sum^{-1} v \rangle$ pour tout $v \in \mathbb{R}^n$

(7) pour $n = k$ et $\sigma$ inversible, on a $Q^* (v) = ||\sigma^{-1} v||^2$ quel que soit $v \in \mathbb{R}^n$.

Les preuves des propriétés (2) (4) (5) (6) (7) sont élémentaires. Précisons que la notion de dualité (1) dépend du choix d'un produit scalaire $\langle\ ,\ \rangle$ hilbertien sur $\mathbb{R}^n$ ; sauf mention précise du contraire, ce produit scalaire sera le produit scalaire euclidien usuel.

## 2.9. Fonctionnelle de Cramer associée à une famille de petites diffusions

Considérons le modèle 2.2 de petites perturbations du système $y'_t = b\ (y_t)$, c'est-à-dire la famille d'équations différentielles stochastiques

$$(E^\varepsilon) \quad d\ y_t^\varepsilon = b_\varepsilon\ (y_t^\varepsilon) + \varepsilon\ \sigma\ (y_t^\varepsilon)\ \ d\ \beta_t$$

sur un ouvert U de $\mathbb{R}^n$, avec $\beta_t$ brownien k-dimensionnel, $\sigma$ champ de matrices (n, k), $b_\varepsilon$ et b champs de vecteurs tels que $\lim_{\varepsilon\to 0} b_\varepsilon = b$. Supposons vérifiées les hypothèses $(H_1, H_2, H_3)$.

A la matrice $\sigma\ (x)$, $x \in U$, associons la forme quadratique $Q_x$ sur $\mathbb{R}^n$, définie par

(1) $$Q_x\ (v) = ||\sigma^*(x)\ v||^2 = < v, \sigma(x)\ \ \sigma\ (x)^*\ \ v\ > ,\ v \in \mathbb{R}^n$$

puis notons $Q_x^* : \mathbb{R}^n \to [0, +\infty]$ la forme quadratique duale de $Q_x$ (cf. 2.8).

Les enveloppes supérieures de fonctions continues étant s.c.i., les formules 2.8 (1), 2.9 (1) montrent que

(2) la fonction $(x, v) \to Q_x^*\ (v)$ est s.c.i. sur $U \times \mathbb{R}^n$

Notons le cas particulier important (cf. 2.8 (7))

(3) pour n = k et $\sigma\ (x)$ inversible on a $Q_x^*\ (v) = ||\sigma^{-1}(x)\ v||^2$

Soit $\mathcal{C}(U)$, $\mathcal{C}^o\ (U) \subset \mathcal{C}(U)$ les espaces des trajectoires explosives à valeurs dans $U \cup \delta$, introduits en 1.2 et 2.2.

Etant donné le système dynamique perturbé $(E^\varepsilon)$ nous définirons sa "transformée de Cramer" $\lambda : \mathcal{C}(U) \to [0, +\infty]$ comme suit :

(4) $$\begin{cases} \lambda\ (g) = +\infty \text{ pour } g \in \mathcal{C}(U)\ ,\ g \notin \mathcal{C}^o\ (U) \\ \lambda\ (g) = \frac{1}{2} \int_0^{\tau(g)\wedge 1} Q_{g_t}^*\ [g'_t - b\ (g_t)]\ \ dt \quad \text{pour } g \in \mathcal{C}^o\ (U) \end{cases}$$

D'après (2) l'intégrale est bien définie, mais peut bien sûr atteindre la valeur $+\infty$ .

Nous définirons la "fonctionnelle de Cramer" $\Lambda$ du système $(E^{\varepsilon})$ par

(5) $$\Lambda(A) = \inf_{g \in A} \lambda(g) \qquad \text{pour tout } A \subset \mathcal{C}(U)$$

Dans le cas particulier où le temps d'explosion $\tau(g)$ est supérieur ou égal à 1 et où $\sigma(x)$ est inversible pour tout $x \in U$ , la formule intégrale (4) devient

(6) $$\lambda(g) = \frac{1}{2} \int_0^1 ||\sigma(g_t)^{-1} [g'_t - b(g_t)]||^2 \, dt$$

On peut dire que $\lambda(g)$ "mesure l'écart entre g et la solution de $y'_t = b(y_t)$ issue du même point" , en utilisant comme métrique sur l'espace tangent en x à U celle qui est définie par la forme quadratique $\frac{1}{2} Q^*_x$ .
Nous reviendrons sur cette formulation dans le cadre plus adapté des variétés (cf. ch. V ).

2.10. Proposition : [Notations 2.2, 2.3] Sur un ouvert U de $\mathbb{R}^n$ , considérons le modèle de petites perturbations $(E^{\varepsilon})$, vérifiant les hypothèses $(H_1\, H_2\, H_3)$. Soit $\lambda : \mathcal{C}(U) \to [0, +\infty]$ la transformée de Cramer de $(E^{\varepsilon})$ [définition 2.9 (4)]. Alors $\lambda$ est s.c.i. et pour tout compact $\Gamma \subset U$ , tout $a \geq 0$, l'ensemble $\Sigma = \{g \in \mathcal{C}(U) \mid g_0 \in \Gamma \text{ et } \lambda(g) \leq a\}$ est compact. De plus pour tout $g \in \mathcal{C}_x(U)$, on a

$$\lambda(g) = \inf \{\tilde{\lambda}(f) \mid f \in C^0(\mathbb{R}^k) \text{ et } B_x(f) = g\}$$

où $\tilde{\lambda}(f) = \frac{1}{2} \int_0^1 |f'_t|^2 \, dt$ est la transformée de Cramer du Brownien (cf. II.3.6). Dans cette formule, l'inf. est atteint si $\lambda(g)$ est fini.

Preuve : considérons le borélien $\Gamma$ de $U \times \mathbb{R}^n$ défini par

$$\Gamma = \{(x, v) \in U \times \mathbb{R}^n \mid \sigma(x)^{-1}[v] \text{ non vide}\}.$$

Pour $(x, v) \in \Gamma$ posons

$$K(x, v) = \{w \in \sigma(x)^{-1}[v] \mid |w| = \inf_{u \in \sigma(x)^{-1}[v]} |u|\}$$

L'application $K : \Gamma \to$ {compacts de $\mathbb{R}^k$} est en fait une "famille mesurable de compacts non vides" au sens de Kuratowski (voir aussi Rockafellar [41] ), car $\{(u, v) \mid K(x, v) \cap S \neq \emptyset\}$ est borélien pour tout fermé $S$ de $\mathbb{R}^k$. Par suite, il existe (cf [41] [11] ) une fonction borélienne $\chi : \Gamma \to \mathbb{R}^k$ telle que $\chi(x, v) \in K(x, v)$ pour tout $(x, v) \in \Gamma$ .

Soit $g \in \mathcal{E}(U)$ tel que $\lambda(g)$ soit fini. D'après 2.9 (4) on a alors $\{Q^*_{g_t}[g'_t - b(g_t)]$ fini$\}$ pour presque tout $t \in [0, \tau(g) \wedge 1]$ . Mais 2.8 (5) et 2.9 (1) donnent $Q^*_x(v) = \inf \{|w|^2 \mid w \in \mathbb{R}^k, \sigma(x) w = v\}$ ,

de sorte que pour presque tout $t \in [0, \tau(g) \wedge 1]$, on a

$(g_t, g'_t - b(g_t)) \in \Gamma$ et

$$Q^*_{g_t}[g'_t - b(g_t)] = |\chi[g_t, g'_t - b(g_t)]|^2$$

d'où, d'après 2.9 (4),

$$(1) \qquad \lambda(g) = \frac{1}{2} \int_0^{\tau(g)\wedge 1} |\chi[g_t, g'_t - b(g_t)]|^2 \, dt < +\infty$$

Il existe donc $f \in C^0(\mathbb{R}^k)$ telle que $\lambda(g) = \frac{1}{2} \int_0^1 |f'(t)|^2 \, dt$. Il suffit en effet de poser $f'_t = \chi[g_t, g'_t - b(g_t)]$ pour presque tout $t \in [0, \tau(g) \wedge 1]$ et $f'_t = 0$ pour $t \in ]\tau(g) \wedge 1, 1]$.

Soit $\tilde{\lambda} : C(\mathbb{R}^k) \to [0, +\infty]$ la transformée de Cramer du mouvement Brownien. On vient de prouver que si $g \in \mathcal{E}_x(U)$ vérifie $\{\lambda(g)$ fini$\}$ , alors il existe $f \in C^0(\mathbb{R}^k)$ telle que $\tilde{\lambda}(f) = \lambda(g)$ et telle que $g = B_x(f)$.

Soit $h \in C^0(\mathbb{R}^k)$ telle que $B_x(h) = g$. On a alors pour presque tout $t < \tau(g)$ , $|f'_t|^2 = Q^*_{g_t}$

$$|f'_t|^2 = Q^*_{g_t} [g'_t - b(g_t)] = Q^*_{g_t} [\sigma(g_t) f'_t]$$

$$\sigma(g_t) f'_t = \sigma(g_t) h'_t$$

de sorte que, grâce à 2.8 (5), 2.9 (1), on a

$$|f'_t|^2 \leqslant |h'_t|^2 \qquad (\text{p.p. } t) \qquad t < \tau(g)$$

Par suite, comme $f'_t = 0$ sur $]\tau(g) \wedge 1, 1]$

$$\tilde{\lambda}(f) = \frac{1}{2} \int_0^1 |f'_t|^2 \, dt = \frac{1}{2} \int_0^{\tau(g)\wedge 1} |f'_t|^2 \, dt \leqslant \frac{1}{2} \int_0^1 |h'_t|^2 \, dt = \tilde{\lambda}(h)$$

Nous avons ainsi prouvé (puisque $\lambda(g) = \tilde{\lambda}(f)$) que

(1) $$\lambda(g) = \inf \{ \tilde{\lambda}(f) \mid f \in C^o(\mathbb{R}^k) \text{ et } B_x(f) = g \}$$

D'après la proposition 2.3, si $C_a$ est l'ensemble des $f \in C(\mathbb{R}^k)$ telles que $\tilde{\lambda}(f) \leqslant a$ et si $\Gamma$ est un compact de $U$, la restriction de $B$ à $\Gamma \times C_a$ est une application continue dans $\mathcal{E}(U)$. Il suffit d'appliquer le lemme 1.3 pour conclure, grâce à (1), que $\lambda$ est s.c.i. et que $\sum = B[\Gamma \times C_a]$ est compact.

2.11. Remarque : si $\sigma$, $b$ sont tels que pour tout $f \in C^o(U)$ le temps d'explosion de $g = B_x(f)$ est strictement supérieur à 1, des arguments analogues aux précédents prouvent que pour la topologie de $C_x(U)$, l'ensemble $K_a \cap C_x(U)$ est compact dans $C_x(U)$, et la restriction de $\lambda$ à $C_x(U)$ est s.c.i.

2.12. Note bibliographique : l'article de base sur le système $(E^\varepsilon)$ est du à Ventcell-Freidlin [53], qui étudient essentiellement le cas où $\sigma(x)$ est inversible (formule 2.9.(6)), bien que pour appliquer leurs résultats à la lettre, il faille se placer sur une variété compacte, ce qui revient à introduire des conditions parasites d'uniforme ellipticité et de bornitude sur les coefficients des diffusions considérées. L'amélioration technique que présente notre méthode n'est pas mince, car elle englobe le cas des diffusions à générateur infinitésimal hypo-elliptiques et des variétés quelconques (voir Chapitre V, plus bas). Un coup d'oeil aux tentatives de Gaveau [25] pour obtenir des résultats de ce type permettra d'apprécier l'efficacité de notre approche.

Le théorème suivant, qui généralise les résultats de Ventcell- Freidlin [53] , et en fournit une démonstration nouvelle, justifie la terminologie "transformée de Cramer" de $(E^{\varepsilon})$.

2.13. <u>Théorème</u> : soit U un ouvert de $\mathbb{R}^n$. Sur U considérons le système dynamique perturbé

$$(E^{\varepsilon}) \qquad d\, y_t^{\varepsilon} = b_{\varepsilon}\, (y_t^{\varepsilon}) + \varepsilon\, \sigma\, (y_t^{\varepsilon})\, d\, \beta_t$$

où les champs de vecteurs $b_{\varepsilon}$ et $b = \lim\limits_{\varepsilon\to 0} b_{\varepsilon}$ , et le champ $\sigma$ de matrices (n, k) vérifient les hypothèses $(H_1\ H_2\ H_3)$ de 2.2. Soit $x \in U$ tel que $y_o^{\varepsilon} = x$ (P-p.s.). Soit $\mathcal{E}_x$ (U) l'espace topologique des trajectoires explosives à valeurs dans $U \cup \delta$ , définies sur $[0,\ 1]$ , issues de x (cf. définition 1.2). Soit $\Lambda$ : {parties de $\mathcal{E}$ (U)} $\to$ $[0,\ +\infty]$ la fonctionnelle de Cramer du système $(E^{\varepsilon})_{\varepsilon>0}$ , définie par les formules 2.9 (4) et 2.9 (5). Soit $y^{\varepsilon} : \Omega \longrightarrow \mathcal{E}_x$ (U) la trajectoire(aléatoire) du système $(E^{\varepsilon})$. Alors pour toute partie A de $\mathcal{E}_x$ (U) on a

$$-\Lambda\,(\overset{\circ}{A}) \leqslant \varliminf_{\varepsilon\to 0}\ \varepsilon^2 \log P\,(y^{\varepsilon} \in A) \leqslant \varlimsup_{\varepsilon\to 0}\ \varepsilon^2\ \log P\,(y^{\varepsilon} \in A) \leqslant -\Lambda\,(\overline{A})$$

De plus, si pour les trajectoires issues de x , les temps d'explosion $\tau\,(y^{\varepsilon})$ et $\tau$ des systèmes $(E^{\varepsilon})$ et $y_t' = b\,(y_t)$ vérifient $\tau > 1$ et $\tau\,(y^{\varepsilon}) > 1$ (P - p.s.), alors le résultat ci-dessus reste vrai lorsqu'on remplace l'espace topologique $\mathcal{E}_x$ (U) par l'ouvert $C_x$ (U) de l'espace de Banach $C_x\,(\mathbb{R}^n)$ [cf. 1.2. pour les notations].

<u>Preuve</u> : soit A une partie borélienne de $\mathcal{E}_x$ (U). Soit $g \in \overset{\circ}{A}$ tel que $\lambda\,(g)$ soit fini, où $\lambda$ est la transformée de Cramer de $(E^{\varepsilon})$. D'après 2.10,

il existe $f \in C^o(U)$ telle que $\tilde{\lambda}(f) = \lambda(g)$ et $B_x(f) = g$ (notations 2.10). Par définition de la topologie de $\mathcal{C}_x(U)$ (cf. 1.2) il existe $T < \tau(g)$, $T \leqslant 1$, et $\rho > 0$ tel que $\{h \in \mathcal{C}_x(U) \mid d_{0,T}(h, g) \leqslant \rho\}$ soit inclus dans l'ouvert $\overset{\circ}{A}$. Donnons-nous $R > 0$, tel que $R > \lambda(g) = \tilde{\lambda}(f)$.

D'après le théorème 2.4, il existe $\alpha > 0$ et $\varepsilon_o > 0$ tels que pour $\varepsilon \leqslant \varepsilon_o$ on ait, si $y_o^\varepsilon = x$ (P-p.s.),

$$P\,(d_{0,T}(\varepsilon\beta, f) \leqslant \alpha \text{ et } d_{0,T}(y^\varepsilon, g) > \rho) \leqslant \exp\left(-\frac{R}{\varepsilon^2}\right),$$

ce qui donne

$$P\,(y^\varepsilon \in A) \geqslant P\,\{d_{0,T}(y^\varepsilon, g) \leqslant \rho\} \geqslant P\,\{d_{0,T}(\varepsilon\beta, f) \leqslant \alpha\} - \exp\left(-\frac{R}{\varepsilon^2}\right)$$

Mais d'autre part (proposition II.3.6), on a

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log P\,\{d_{0,T}(\varepsilon\beta, f) \leqslant \alpha\} = -\inf\,\{\tilde{\lambda}(\varphi) \mid d_{0,T}(\varphi, f) \leqslant \alpha\} \geqslant -\tilde{\lambda}(f)$$

et $-\tilde{\lambda}(f) > -R$, ce qui entraîne finalement

$$\underline{\lim}_{\varepsilon\to 0}\, \varepsilon^2 \log P\,(y^\varepsilon \in A) \geqslant -\tilde{\lambda}(f) = -\lambda(g).$$

Prenons le sup. en $g \in \overset{\circ}{A}$ pour obtenir par définition de $\Lambda$ (cf. 2.9 (5))

$$\underline{\lim}_{\varepsilon\to 0}\, \varepsilon^2 \log P\,(y^\varepsilon \in A) \geqslant -\Lambda(\overset{\circ}{A}).$$

Soit maintenant $a$ fini tel que $a < \Lambda(\overline{A})$.

Posons $K_a = \{g \in \mathcal{C}_x(U) \mid \lambda(g) \leqslant a\}$

$C_a = \{f \in C^o(\mathbb{R}^k) \mid \tilde{\lambda}(f) \leqslant a\}$

D'après la relation entre $\lambda$ et $\tilde{\lambda}$ explicitée par la proposition 2.10, on a $K_a = B_x(C_a)$.

Le choix de $a$ garantit $K_a \cap \overline{A} = \emptyset$. Donc pour tout $g \in K_a$, il existe un voisinage ouvert $V_g$ de $g$ dans $\mathcal{C}_x(U)$ tel que $V_g \cap A = \emptyset$. Fixons un nombre $R > a$.

Par définition de la topologie de $\mathcal{E}_x$ (U), il existe (cf. 1.2) pour chaque $V_g$ des nombres $\rho_g > 0$ $T_g \in ]0, 1]$ , $T_g < \tau(g)$ ,

tels que $G_g = \{h \in \mathcal{E}_x(U) \mid d_{0,T_g}(h, g) \leq \rho_g\}$

soit contenu dans $V_g$ . Soit $f_g \in C_a$ telle que $g = B_x(f_g)$.

D'après le théorème 2.4, il existe $\varepsilon_g > 0$ , $\alpha_g > 0$ tels que

si $F_g = \{f \in C^0(\mathbb{R}^k) \mid d_{0,T_g}(f, f_g) < \alpha_g\}$

alors $P\{(\varepsilon\beta \in F_g) \text{ et } (y^\varepsilon \notin G_g)\} \leq \exp(-\frac{R}{\varepsilon^2})$ , ce qui

entraîne $P\{(\varepsilon\beta \in F_g) \text{ et } (y^\varepsilon \notin V_g)\} \leq \exp(-\frac{R}{\varepsilon^2})$.

Les $(F_g)_{g\in K_a}$ forment un recouvrement ouvert du compact $C_a$. On peut en extraire un recouvrement fini $F_{g_1}, \dots, F_{g_k}$ . Posons $F = \bigcup_{i=1}^{k} F_{g_i}$ .

L'évènement $\{\varepsilon\beta \in F\} \cap (y^\varepsilon \in A)$ est égal à

$\bigcup_i [\{\varepsilon\beta \in F_{g_i}\} \cap \{y^\varepsilon \in A\}]$ , et par suite est inclus dans

$\bigcup_i [\{\varepsilon\beta \in F_{g_i}\} \cap \{y^\varepsilon \notin V_{g_i}\}]$ . On aura donc

$$P\{(\varepsilon\beta \in F) \cap (y^\varepsilon \in A)\} \leq k \exp(-\frac{R}{\varepsilon^2})$$

pourvu que $\varepsilon \leq \varepsilon_{g_1} \wedge \dots \wedge \varepsilon_{g_k}$ . Ceci entraîne

$$P(y^\varepsilon \in A) \leq P\{(\varepsilon\beta \in F) \cap (y^\varepsilon \in A)\} + P(\varepsilon\beta \in F^c)$$

$$\leq k \exp(-\frac{R}{\varepsilon^2}) + P(\varepsilon\beta \in F^c)$$

Comme F est ouvert, on a (proposition II.3.6),

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log P(\varepsilon\beta \in F^c) = -\tilde{\Lambda}(F^c)$$

et donc $\overline{\ell}(A) = \overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P(y^\varepsilon \in A) \leq (-R) \vee [-\tilde{\Lambda}(F^c)]$

Puisque $R > a$ et $F^c \cap C_a = \emptyset$ , le second membre est majoré par (- a). La relation $\overline{\ell}(A) \leq -a$ pour tout $a < \Lambda(\overline{A})$, a fini, entraîne alors

$\overline{\ell}(A) \leq -\Lambda(\overline{A})$.

Ceci achève la preuve du résultat général. Si les temps d'explosions sont tous plus grands que 1, il suffit de relire la preuve ci-dessus pour constater que le résultat reste valable dans $C_x(U)$.

2.14. Quelques résultats voisins du théorème 2.13 : au chapitre V, nous étendrons la validité du théorème 2.13 au cas où U est une variété différentiable. D'autre part (voir Chapitre IV, § 1) on peut dans une certaine mesure, préciser l'uniformité des $\underline{\lim}_{\varepsilon\to 0}$ et $\overline{\lim}_{\varepsilon\to 0}$ de $\varepsilon^2 \log P(y^\varepsilon \in A \mid y_o^\varepsilon = x)$ quand x et A varient.

# CHAPITRE IV

## APPLICATION : GOULOTS DE SORTIE, STABILISATION DES EQUILIBRES PERTURBES

### 1. DEUX ESTIMATIONS ASYMPTOTIQUE UNIFORMES :

1.1 Hypothèses et notations : Sur un ouvert U de $\mathbb{R}^n$ nous considérons le processus $(y_t^\varepsilon)$ solution du modèle de petites perturbations $(E^\varepsilon)$ vérifiant les hypothèses $(H_1\ H_2\ H_3)$(cf. III 2.2 et III 2.3 dont nous reprenons systèmatiquement les notations). Nous poserons $P_x(y^\varepsilon \in .) = P(y^\varepsilon \in .\,|y_0^\varepsilon = x)$. Soit $Q_x^*$, $x \in U$ le champ de formes quadratiques duales associé à $(E^\varepsilon)$(cf. III 2.9).

Pour $S < T$, nous définisons les espaces de trajectoires $\mathcal{E}_{S,T}(U)$, $\mathcal{E}^o_{S,T}(U)$, $C_{S,T}(U)$, $C^o_{S,T}(U)$ comme en III 1.2, III 2.2, mais en remplaçant l'intervalle $[0,1]$ par $[S,T]$. Pour $g \in \mathcal{E}^o_{S,T}(U)$ nous poserons

$$\lambda_{S,T}(g) = \frac{1}{2}\int_S^{T\wedge\tau(g)} Q^*_{g_t}(g_t')dt$$

où $\tau(g) > S$ est le temps d'explosion de g ; si $g \in \mathcal{E}_{S,T}(U) - \mathcal{E}^o_{S,T}(U)$, nous poserons $\lambda_{S,T}(g) = +\infty$.

Pour $f \in C_{S,T}(\mathbb{R}^k)$ nous poserons $\tilde{\lambda}_{S,T}(f) = \frac{1}{2}\int_S^T |f_t'|^2 dt$ si $f \in C^o_{S,T}(\mathbb{R}^k)$ et $\tilde{\lambda}_{S,T}(f) = +\infty$ si $f \notin C^o_{S,T}(\mathbb{R}^k)$. Enfin pour $A \subset \mathcal{E}_{S,T}(U)$, $B \subset C_{S,T}(\mathbb{R}^k)$, nous poserons

$$\Lambda_{S,T}(A) = \inf_{g \in A} \lambda_{S,T}(A) \qquad \tilde{\Lambda}_{S,T}(B) = \inf_{f \in B} \tilde{\lambda}_{S,T}(B).$$

Les propositions 1.2 et 1.4 ci-dessous sont modelées sur les deux résultats de base de Ventsel-Freidlin [53]. Nous les étendons ici au cas des diffusions à générateur differentiel hypoelliptique et à temps d'explosion. Nous suggérons au lecteur d'ignorer leurs preuves, qui reprennent avec une précision très technique des arguments utilisés pour prouver le th.III 2.13. Ces résultats seront cruciaux pour étendre le th. III 2.13 aux variétés.

1.2 Proposition : Considérons le processus $y^{\varepsilon}$ solution du système perturbé $(E^{\varepsilon})$ sur l'ouvert U de $\mathbb{R}^n$ (voir 1.1 pour les hypothèses et notations). Soit $\lambda_{0,T}$ la transformée de Cramer des $(E^{\varepsilon})$ sur $[0,T]$.

Donnons nous un compact K de U et des nombres strictement positifs T, a, $\rho$, $\eta$. Alors il existe $\varepsilon_o > 0$ et $r > 0$ tels que les conditions

(1) $x \in K$; $g \in \mathcal{E}_{0,T}(U)$ ; $|g_o - x| \leqslant r$ ; $g[0,T] \subset K$ ;

$\lambda_{0,T}(g) \leqslant a$ ; $\varepsilon \leqslant \varepsilon_o$ ;

entraînent

(2) $-\lambda_{0,T}(g) - \eta \leqslant \varepsilon^2 \log P_x\{d_{0,T}(y^{\varepsilon},g) \leqslant \rho\}$

Preuve : D'après la prop. III 2.10, pour $g \in \mathcal{E}_{0,T}(U)$ tel que $\lambda_{0,T}(g)$ soit fini il existe $f_g \in C^o_{0,T}(\mathbb{R}^k)$ telle que

(3) $\tilde{\lambda}_{0,T}(f_g) = \lambda_{0,T}(g)$ et $g = B_{g_o}(f_g)$.

Donnons nous K, T, a, $\rho$, $\eta$ et $R = a + 2$. Le th.III 2.4 fournit alors $\varepsilon_o$, r, $\alpha$ positifs stricts tels que

(4) $P_x\{d_{0,T}(\varepsilon\beta, f_g) \leqslant \alpha$ et $d_{0,T}(y^{\varepsilon},g) > \rho\} \leqslant \exp\left(-\frac{R}{\varepsilon^2}\right)$ pourvu que x, g,$\varepsilon$ vérifient (1).

Posons $C_a = \{f \in C_{0,T}(\mathbb{R}^k) \mid \tilde{\lambda}_{0,T}(f) \leqslant a\}$

Soient $f_1 \ldots f_N \in C_a$ telles que les $[0,T]$ - tubes ouverts $V_i$ d'axe $f_i$, de rayon $\frac{\alpha}{2}$ recouvrent $C_a$. A partir de $\eta, T, V_1 \ldots V_N$, le th.III 2.13 fournit $\varepsilon_1 > 0$ tel que $\varepsilon \leqslant \varepsilon_1$ entraîne

(5) $-\tilde{\Lambda}_{0,T}(V_i) - \eta \leqslant \varepsilon^2 \log P(\varepsilon\beta \in V_i)$ ; $i = 1 \dots N$.

Pour tout $g$ tel que $\lambda_{0,T}(g) \leqslant a$, on a $f_g \in C_a$, et donc $f_g \in V_{i_g}$ pour un entier $i_g$ convenable. Ceci entraîne

(6) $\tilde{\Lambda}_{0,T}(V_{i_g}) \leqslant \tilde{\lambda}_{0,T}(f_g) = \lambda_{0,T}(g)$

$\{\varepsilon\beta \in V_{i_g}\} \subset \{d_{0,T}(\varepsilon\beta, f_g) \leqslant \alpha\}$

De (5) (6) on conclut que $\varepsilon \leqslant \varepsilon_1$ et $\lambda_{0,T}(g) \leqslant a$ entraînent

(7) $-\lambda_{0,T}(g) - \eta \leqslant \varepsilon^2 \log P\{d_{0,T}(\varepsilon\beta, f_g) \leqslant \alpha\}$

De (7) (4) et (3) on déduit que si $\varepsilon \leqslant \varepsilon_o \wedge \varepsilon_1$, et si $x,g$ vérifient (1) on aura

(8) $P_x\{d_{0,T}(y^\varepsilon, g) \leqslant \rho\} \geqslant P_x\{d_{0,T}(y^\varepsilon, g) \leqslant \rho$ et $d_{0,T}(\varepsilon\beta, f_g) \leqslant \alpha\}$

$$\geqslant \exp -\frac{1}{\varepsilon^2}\left[\lambda_{0,T}(g) + \eta\right] - \exp\left(-\frac{R}{\varepsilon^2}\right)$$

Le second membre de (8) s'écrit, en posant $\lambda = \lambda_{0,T}(g) \leqslant a$,

$$\left[\exp \frac{\eta}{\varepsilon^2} - \exp \frac{-R+\lambda+2\eta}{\varepsilon^2}\right] \exp -\frac{1}{\varepsilon^2}(\lambda + 2\eta)$$

et est donc minoré (puisque $R = a + 2$) par $\exp -\frac{1}{\varepsilon^2}(\lambda + 2\eta)$ pourvu que l'on assure $\eta \leqslant 1$ et $\varepsilon^2 \leqslant \eta$, ce qui prouve la prop. 1.2.

1.3 Lemme : (hypothèses et notations de 1.1) Pour tout compact $K$ de $U$, le temps

$$\theta_K = \inf\{\tau(g) \mid g \in \mathcal{E}_{0,\infty}(U),\ g_o \in K,\ g'_t = b(g_t) \text{ pour tout } t \in [0,\tau(g)[$$

est positif strict. Pour tout temps fini $S$ tel que $0 \leqslant S < \theta_K$, le nombre

$$\alpha_{K,S} = \inf\{\lambda_{0,S}(g) \mid g \in \mathcal{E}_{0,S}(U),\ g_o \in K,\ \tau(g) \leqslant S\}$$

est positif strict. En particulier, pour tout $A < \alpha_{K,S}$, les conditions

$$\{g \in \mathcal{E}_{0,S}(U),\ g_o \in K,\ \lambda_{0,S}(g) \leqslant A\} \text{ impliquent } \{S < \tau(g)\}.$$

Preuve : Pour $x \in U$ posons $f(x) = \tau(g)$ où $g$ est la solution maximale de $g'_t = b(g_t)$ issue de $g_o = x$. D'après la prop. III 2.3, $f$ est évidemment s.c.i. sur $U$, et positive stricte. Sa borne inférieure sur le compact $K$ est atteinte, et donc positive stricte, ce qui prouve $\theta_K > 0$.

Comme $\tau$ est s.c.i. sur $\mathcal{E}_{0,S}(U)$, l'ensemble G des g telles que $g_o \in K$, $\tau(g) \leqslant S$ est fermé dans $\mathcal{E}_{0,S}(U)$. Les propriétés de s.c.i. et compacité de $\lambda_{0,S}$ montrent que $\lambda_{0,S}$ atteint sa borne inférieure $\alpha_{K,S}$ sur G, en un point $g \in G$. Si on avait $\lambda_{0,S}(g) = 0$, g serait solution de $g_t' = b(g_t)$ sur $[0,\tau(g)[$ et par suite on devrait avoir $\tau(g) \geqslant \theta_K > S$ ce qui contredit $g \in G$. Donc $\alpha_{K,S} = \lambda_{0,S}(g) > 0$.

1.4 Proposition : Considérons le processus $y^\varepsilon$ solution du système perturbé $(E^\varepsilon)$ sur l'ouvert U de $\mathbb{R}^n$ (voir 1.1 pour les hypothèses). Soit $\lambda_{0,T}$ la transformée de Cramer de $(E^\varepsilon)$ sur $[0,T]$. Donnons nous un compact K de U et des nombres T, A positifs tels que les conditions $\{g \in \mathcal{E}_{0,T}(U),\ g_o \in K,\ \lambda_{0,T}(g) \leqslant A\}$ impliquent $T < \tau(g)$ (cf. lemme 1.3).

Alors pour tout $\rho > 0$, $\eta > 0$, il existe $\varepsilon_o > 0$ tel que les conditions

(1) $\varepsilon \leqslant \varepsilon_o$ ; $a \leqslant A$ ; $x \in K$

entraînent

(2) $P_x\{d_{0,T}(y^\varepsilon,g) \geqslant \rho$ pour tout $g \in \mathcal{E}_{0,T}(U)$ vérifiant $\lambda_{0,T}(g) \leqslant a$, $g_o = x\} \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{\eta}{\varepsilon^2}\right)$

Preuve : Pour $L \subset U$ posons

$$\Gamma(A,L) = \{g \in \mathcal{E}_{0,T}(U) \mid \lambda_{0,T}(g) \leqslant A,\ g_o \in L\}$$

Par hypothèse $g \in \Gamma(A,K)$ force $T < \tau(g)$. Si l'union des $g([0,T])$ pour $g \in \Gamma(A,K)$ n'était pas relativement compacte, il existerait $g_n \in \Gamma(A,K)$ et $t_n \in [0,T]$ tels que $g_n(t_n)$ sorte de tout compact. Par extraction d'une sous suite, on peut supposer que

$$g_n \to g \in \Gamma(A,K),\ \text{d'où } \tau(g) > T ;$$

mais alors (cf. III 1.2) $g_n$ convergerait uniformément vers g sur $[0,T]$, ce qui contredirait l'existence des $t_n$. Ainsi il existe un compact M tel que $g[0,T] \subset M$ pour tout $g \in \Gamma(A,K)$.

Fixons $a \leqslant A$ et $R = a + 3$. Le th.III 2.4 montre l'existence de $\varepsilon_o$, $\alpha$ positifs strict tels que les conditions

(3) $\varepsilon \leqslant \varepsilon_o$ ; $x \in K$ ; $g = B_x(f)$ ; $\lambda_{0,T}(g) \leqslant A$

entraînent

(4) $P_x\{d_{0,T}(\varepsilon\beta,f) \leqslant \alpha$ et $d_{0,T}(y^\varepsilon,g) \geqslant \rho\} \leqslant \exp\left(-\frac{R}{\varepsilon^2}\right)$

Considérons le compact $C_a = \{f \in C_{0,T}(\mathbb{R}^k) \,|\, \tilde{\lambda}_{0,T}(f) \leqslant a\}$

Les $[0,T]$ tubes ouverts $F_f$ d'axes $f$ et de rayon $\alpha$ forment un recouvrement ouvert de $C_a$. On en extrait un recouvrement fini $F_1, \ldots, F_N$ de $C_a$, avec $F_i = F_{f_i}$, et $f_i \in C_a$. En particulier si $g_i = B_x(f_i)$, on a $\lambda_{0,T}(g_i) \leqslant a$, d'après la prop. 2.10. Posons $F = \bigcup_{1 \leqslant i \leqslant N} F_i$

$$M(x,a) = \{h \in \mathcal{E}_{0,T}(U) \,|\, d_{0,T}(h,g) \geqslant \rho \text{ pour tout } g \text{ tel que } g_o = x$$
$$\text{et } \lambda_{0,T}(g) \leqslant a\}$$

Alors comme dans III 2.13 on a en notant $M = M(x,a)$

$$\{y^\varepsilon \in M\} \subset \{\varepsilon\beta \notin F\} \cup \left[\{\varepsilon\beta \in F\} \cap \{y^\varepsilon \in M\}\right]$$
$$\subset \{\varepsilon\beta \notin F\} \cup \bigcup_{1 \leqslant i \leqslant N} \left[\{\varepsilon\beta \in F_i\} \cap \{d_{0,T}(y^\varepsilon,g_i) \geqslant \rho\}\right]$$

De (3) (4) on conclut alors que pour $\varepsilon \leqslant \varepsilon_o$, $x \in K$ on a

(5) $P_x(y^\varepsilon \in M) \leqslant P(\varepsilon\beta \notin F) + N \exp\left(-\frac{R}{\varepsilon^2}\right)$

Comme $F^c$ est un fermé disjoint de $C_c$ on aura

$$\lim_{\varepsilon \to 0} \varepsilon^2 \log P(\varepsilon\beta \notin F) = -\tilde{\lambda}_{0,T}(F^c) \leqslant -a$$

Etant donné $\eta \in \,]0,1]$ il existe donc $\varepsilon_1 > 0$ tel que $\varepsilon \leqslant \varepsilon_1$ entraîne $P(\varepsilon\beta \notin F) \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{\eta}{\varepsilon^2}\right)$. Finalement, grâce à (5) et à $R = a+3$, on voit que $\varepsilon \leqslant \varepsilon_o \wedge \varepsilon_1$, $x \in K$ impliquent

(6) $P_x(y^\varepsilon \in M) \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{2\eta}{\varepsilon^2}\right)\left[\exp\left(-\frac{\eta}{\varepsilon^2}\right) + N \exp\left(-\frac{1}{\varepsilon^2}\right)\right]$

d'où $\varepsilon_2 > 0$ tel que $x \in K$, $\varepsilon \leqslant \varepsilon_2$ entraînent

(7) $P_x\{y^\varepsilon \in M(x,a)\} \leqslant \exp(-\frac{a}{\varepsilon^2} + \frac{2\eta}{\varepsilon^2})$

Pour le moment $\varepsilon_2$ dépend de $a \leqslant A$. Appliquons (7) aux valeurs $a = k\eta$, $k = 1 \dots \left[\frac{A}{\eta}\right]$, ce qui nous fournit $\varepsilon_3 > 0$ tel que $\varepsilon \leqslant \varepsilon_3$ et $x \in K$ garantissent pour $k = 1 \dots . \left[\frac{A}{\eta}\right]$

(8) $P_x\{y^\varepsilon \in M(x,k\eta)\} \leqslant \exp(-\frac{k\eta}{\varepsilon^2} + \frac{2\eta}{\varepsilon^2})$

Soit alors $a$ quelconque dans $[0,A]$ ; choisissons $k$ tel que $k\eta \leqslant a \leqslant (k+1)\eta$. On a alors

$M(x,a) \subset M(x,k\eta)$, donc pour $\varepsilon \leqslant \varepsilon_3$, $x \in K$, $a \leqslant A$, (8) implique

$P_x\{y^\varepsilon \in M(x,a)\} \leqslant \exp(-\frac{k\eta}{\varepsilon^2} + \frac{2\eta}{\varepsilon^2}) \leqslant \exp(-\frac{a}{\varepsilon^2} + \frac{3\eta}{\varepsilon^2})$

ce qui prouve la majoration (2) annoncée.

Remarque : Lorsque les temps d'explosion des solutions de $z'_t = b(z_t) + \sigma(z_t)\, f'_t$ sont toujours plus grands que $T$ dès que $f \in C^o_{0,T}(R^k)$, les énoncés 1.2 et 1.3 se simplifient légèrement ; ce cas correspond à la situation considérée par Ventsel-Freidlin qui prenaient $\sigma$ inversible, $U$ variété compacte et donc avaient des temps d'explosion toujours infinis.

1.5 Corollaire : (hypothèses et notations de 1.1) Considérons le processus $y^\varepsilon$ solution du système perturbé $(E^\varepsilon)$ sur l'ouvert $U$ de $R^n$. Soit $\Lambda_{0,T}$ la fonctionnelle de Cramer de $(E^\varepsilon)$ sur $[0,T]$. Donnons nous un compact $K$ de $U$, et des nombres $\dot{J} > 0$, $T > 0$ tels que les conditions

$\{g \in \mathcal{E}_{0,T}(U),\ g_o \in K,\ \lambda_{0,T}(g) \leqslant \dot{J}$ entraînent $T < \tau(g)$

(cf. Lemme 1.3).

Alors pour tout $\eta > 0$ il existe $\varepsilon_o > 0$ ayant la propriété suivante :

Quel que soit la famille $F_x$, $x \in K$, de boréliens de $\mathcal{E}_{0,T}(U)$, quels que soient $x \in K$ et $\varepsilon \leq \varepsilon_0$, on a, en notant $F = \overline{\bigcup_{x\in K} F_x}$

$$P_x(y^\varepsilon \in F_x) \leq \exp\left[-\frac{1}{\varepsilon^2}\min(\dot{J}, \Lambda_{0,T}(F)) + \frac{\eta}{\varepsilon^2}\right]$$

<u>Preuve</u> : Fixons $a = \min\left[\dot{J}, \Lambda_{0,T}(F)\right] - \eta$. L'ensemble $\Gamma(a,K)$ défini plus haut en 1.4 est alors compact et disjoint du fermé F. Comme $a \leq \dot{J}$, on a par hypothèse sur $\dot{J}$, $T < \tau(g)$ dès que $g \in \Gamma(a,K)$. Soit $O(g,r)$ le $[0,T]$ - tube ouvert d'axe g et de rayon r. Pour $g \in \Gamma(a,K)$, il existe $r_g$ tel que $O(g,2r_g)$ soit disjoint de F. On recouvre $\Gamma(a,K)$ par $O_i$ $i = 1 \ldots N$ où $O_i = O(g_i, r_{g_i})$.

Soit $\rho = \inf(r_{g_i}, i = 1 \ldots N)$. Alors pour tout $f \in F$, $h \in \Gamma(a,K)$ on aura $d_{0,T}(f,h) \geq \rho$. Ceci fournit les inclusions

$$(y^\varepsilon \in F_x) \subset (y^\varepsilon \in F) \subset \{d_{0,T}(y^\varepsilon, g) \geq \rho \text{ pour tout } g \in \Gamma(a,K)\}$$
$$\subset \{d_{0,T}(y^\varepsilon, g) \geq \rho \text{ pour tout } g \in \Gamma(a,x)\}$$

ce qui prouve le corollaire, grâce à la prop. 1.4.

## 2. <u>TEMPS ET GOULOT DE SORTIE DU VOISINAGE D'UN EQUILIBRE STABLE (D'APRES VENTSEL-FREIDLIN)</u> :

2.1 <u>Le problème</u> : Sur un ouvert V de $\mathbb{R}^n$ on considère le système dynamique $y'_t = b(y_t)$, où b est un champ lipschitzien (localement) sur V. On se donne <u>un point d'équilibre $0 \in V$</u> pour le système (i.e. $b(0) = 0$), que l'on suppose <u>stable</u> au sens suivant :

il existe <u>un ouvert U contenant 0</u>, tel que <u>$\overline{U}$ soit un compact de V</u>, et tel que <u>toute trajectoire du système issue de $x \in U$ tende vers 0 quand $t \to +\infty$</u>, <u>sans sortir de U</u>.

Nous supposons que U est à bord régulier et que pour tout $x \in \partial U$, le champ $b(x)$ est strictement "rentrant" (i.e. $\langle b(x), n(x)\rangle > 0$ où $n(x)$ est la normale à $\partial U$ orientée vers l'intérieur de U).

Ce système est remplacé par un système perturbé $(E^\varepsilon)$

$$dy_t^\varepsilon = b(y_t^\varepsilon)dt + \varepsilon\,\sigma(y_t^\varepsilon)d\beta_t$$

où $\sigma(y)$ est inversible $(n,n)$, de classe un et $\varepsilon$ est petit ; $\beta$ est un Brownien n-dimensionnel.

Pour chaque $\varepsilon > 0$ et $x \in U$, le temps de sortie $\lambda_U^\varepsilon$ de U pour $y^\varepsilon$ est $P_x$ - p.s. fini. En effet la fonction $P_x(\lambda_U = +\infty)$ est $\Delta^\varepsilon$-harmonique dans U, nulle sur $\partial U$, avec $\Delta^\varepsilon$ générateur différentiel du processus $y_t^\varepsilon$. Elle est donc nulle partout sur U.

Il s'agit de déterminer en quels points de U se fait la sortie de U pour $(y_t^\varepsilon)$, et quel chemin suit $y_t^\varepsilon$ pour sortir de U, quand $\varepsilon \to 0$.

Convention : Nous pouvons toujours modifier les champs b et $\sigma$ en dehors d'un voisinage W compact de $\overline{U}$ de telle sorte que b et $\sigma$ soient définis sur $\mathbb{R}^n$ entier, aient les mêmes propriétés de régularité locales, et soient bornées sur $\mathbb{R}^n$. Ceci ne change pas le processus induit sur les voisinages ouverts de $\overline{U}$ contenus dans W. On peut donc sans perte de généralité supposer que $V = \mathbb{R}^n$, que le temps d'explosion du processus $y_t^\varepsilon$ est infini $P_x$ - p.s., et que si $\tilde{\lambda}_{0,T}(f) < +\infty$, les solutions maximales de $z_t' = b(z_t) + \sigma(z_t)f_t'$ n'explosent pas sur $[0,T]$.

2.2 Définitions : Appelons orbite du système initial $(E^o)$ l'image dans $\mathbb{R}^n$ de n'importe quelle solution maximale de $y_t' = b(y_t)$. Appelons ensemble $\omega$-limite A toute partie A de $\mathbb{R}^n$ telle que pour chaque $x \in A$, il existe une solution $y_t$ de $(E^o)$ vérifiant $\lim\limits_{+\infty} y_{t_n} = x$ pour une certaine suite $t_n \in \mathbb{R}$.

2.3 Lemme (cf. [53]) : Soit K une partie compacte de $\mathbb{R}^n$, ne contenant pas d'ensemble $\omega$-limite. Il existe alors $a > 0$, $b > 0$ tels que pour tout $T \geqslant 0$, les conditions $g \in C_{0,T}(\mathbb{R}^n)$ et $g[0T] \subset K$ impliquent $\lambda_{0,T}(g) \geqslant aT - b$

Preuve : Soit $y_t$ la solution de $(E^o)$ vérifiant $y_o = v$, où $v \in K$. Si le temps de sortie de K pour $y_t$, noté $\tau_K(v)$ était infini, K contiendrait un ensemble $\omega$-limite ; par suite $\tau_K(v)$ est fini. D'autre part, les solutions de $(E^o)$ étant fonctions continues de la donnée initiale v, uniformément sur les compacts de $[0, +\infty[$, on voit facilement que $v \to \tau_K(v)$ est une fonction s.c.s. Le maximum $\alpha$ de cette fonction sur le compact K est donc atteint, et par suite fini.

L'ensemble $A = \{g \in C_{0,2\alpha}(\mathbb{R}^n) \mid g[0,2\alpha] \subset K\}$ est fermé dans $C_{0,2\alpha}(\mathbb{R}^n)$. Supposons que $\lambda_{0,2\alpha}$ ne soit pas identiquement infinie sur A. L'ensemble $A \cap \{g \in C_{0,2\alpha}(\mathbb{R}^n) \mid \lambda_{0,2\alpha}(g) \leqslant u\}$ est alors non vide pour u assez grand, et compact (remarque III 2.11). Comme (cf. III 2.11) $\lambda_{0,2\alpha}$ est s.c.i. sur $C_{0,T}(\mathbb{R}^n)$, on voit que $\lambda_{0,2\alpha}$ atteint son minimum sur A en un point $\gamma \in A$.

Si $\lambda_{0,2\alpha}(\gamma)$ était nul, on aurait (prop. 2.10) $\gamma_t' = b(\gamma_t) + \sigma(\gamma_t)f_t'$ p.p. en t où $f_t' = 0$ p.p. t. Comme la solution y de $(E^o)$ issue de $\gamma_o$ vérifie aussi cette équation, on aurait $\gamma_t \equiv y_t$. Puisque $\gamma_o \in K$, ceci entraînerait $\tau_K(\gamma_o) \leqslant \alpha$, ce qui contredirait $\gamma[0,2\alpha] \subset K$. Par suite $\lambda_{0,2\alpha}(\gamma) = a > 0$ et on a $\lambda_{0,2\alpha}(g) \geqslant a$ pour tout $g \in A$.

Soit maintenant $T > 0$. Prenons k entier tel que $k\alpha \leqslant T < (k+1)\alpha$. Si $g[0,T] \subset K$, on aura

$$\lambda_{0,T}(g) \geqslant \sum_{j \leqslant k} \lambda_{(j-1)\alpha, j\alpha}(g) \geqslant ka \geqslant a(\frac{T}{2\alpha} - 1)$$

2.4 Lemme : (cf. [53]) Soit K une partie compacte de $\mathbb{R}^n$ ne contenant pas d'ensembles ω-limite. Soit $\tau_K^\varepsilon$ le temps de première sortie de K pour le processus $y_t^\varepsilon$. Il existe des constantes $a > 0$, $b > 0$ telles que :

à tout $T_o > 0$, on peut associer $\varepsilon_o > 0$ tel que les conditions $T \leqslant T_o$, $x \in K$ et $\varepsilon \leqslant \varepsilon_o$ impliquent

$$P_x(\tau_K^\varepsilon \geqslant T) \leqslant \exp\left( - \frac{aT-b}{\varepsilon^2} \right)$$

Preuve : L'ensemble $A = \{g \in C_{0,T}(\mathbb{R}^n) \mid g_o \in K,\ \tau_K(g) > T\}$ est d'après 2.3 disjoint du compact

$$C = \{g \in C_{0,T}(\mathbb{R}^n) \mid \lambda_{0,T}(g) \leqslant aT - 2b \text{ et } g_o \in K\}$$

pour $a > 0$ $b > 0$ comme en 2.5. Comme A est fermé il existe $\rho > 0$ tel que A soit disjoint de tout $[0,T]$ - tube d'axe $g \in C$ et de rayon ρ. A fortiori A est inclus dans l'ensemble B des $h \in C_{0,T}(\mathbb{R}^n)$ qui vérifient $d_{0,T}(h,g) \geqslant \rho$ pour tout $g \in C_{0,T}(\mathbb{R}^n)$ telle que $\lambda_{0,T}(g) \leqslant aT - b$, $g_o \in K$, $|g_o - x| \leqslant r$ (où $x \in K$ et $r > 0$ sont fixés quelconques). Il suffit d'appliquer la prop. 1.3 pour conclure à l'existence de $\varepsilon_o > 0$ tel que pour $T \leqslant T_o$, $x \in K$, $\varepsilon \leqslant \varepsilon_o$ on ait

$$P_x(y^\varepsilon \in A) \leqslant P_x(y^\varepsilon \in B) \leqslant \exp\left[-\frac{aT-b}{\varepsilon^2} + \frac{\eta}{\varepsilon^2}\right]$$

où $\eta > 0$ est fixé arbitrairement, ce qui prouve le lemme.

2.5 Lemme : Pour toute partie compacte K de $\mathbb{R}^n$, il existe un nombre $C > 0$ tel que pour tout $x,y \in K$ il existe un temp T et $g \in C_{0,T}(\mathbb{R}^n)$ tels que $g_o = x$, $g_T = y$ et $\lambda_{0,T}(g) \leqslant C\,|x-y|$

Preuve : Il suffit de poser $g_t = x + \frac{t}{|x-y|}(y-x)$ et $T = |x-y|$. Alors $g_t'$, $\sigma(g_t)^{-1}$, et $b(g_t)$ restent bornés en norme, de sorte que $Q^*_{g_t}(g_t') \leqslant$ cte pour $t \in [0,T]$, ce qui prouve le résultat.

2.6 Corollaire (cf. [53]) : La fonction $V_U(x,y)$ définie sur U x U par $V_U(x,y) = \inf\{\lambda_{0,T}(g) \mid g_o = x,\ g_T = y,\ g[0,T[ \subset U,\ T>0$ <u>arbitraire</u> $\}$ est finie et localement lipschitzienne sur U x U.

2.7 <u>Une hypothèse simple pour l'étude du goulot de sortie</u> :

Plaçons nous dans la situation décrite en 2.1.

Posons $V_o = \inf_{y \in \partial U} V_{\mathbb{R}^n}(0,y) = \inf_{y\in \partial U} V_U(0,y)$. Nous supposerons <u>l'existence d'une unique fonction</u> $\varphi : ]-\infty, 0] \to \mathbb{R}^n$ telle que $\varphi_o = y_o \in \partial U$, $\varphi]-\infty, 0[ \subset U$, $\lim_{t\to -\infty} \varphi_t = 0$, $V_o = V(y_o) = \lambda_{-\infty,0}(\varphi)$.

<u>Une telle fonction $\varphi$ existe toujours</u> grâce à la s.c.i. de $\lambda_{S;T}$ et à la compacité de $\{g|\lambda_{S,T}(g) \leqslant a\}$ (voir [53] ou bien le lemme 2.8 plus bas pour une preuve précise). Il se peut que $\varphi$ soit en fait constante égale à 0 pour $t \in ]-\infty,S_o]$ avec $S_o < 0$, mais ce n'est pas toujours le cas (voir [53]). L'hypothèse d'unicité par contre n'est pas toujours vérifiée. Elle facilite l'intuition, mais n'est pas essentielle (voir Azencott-Ruget [3] qui, dans une situation à priori plus retorse, évite cette hypothèse, ce qui amène à formuler les résultats de façon un peu différente).

Toutes les fonctions considérées seront à valeurs dans $\mathbb{R}^n$, et continues sur leur intervalles de définition.

Nous définirons <u>la translatée (dans le temps)</u> $\theta_t h$ <u>d'une fonction</u> $h : \mathbb{R} \to \mathbb{R}^n$ <u>par</u> $\theta_t h(u) = h(u+t)$.

Nous noterons $\tau_A(g) = \inf\{t|g_t \notin A\}$ <u>le temps de première sortie de</u> $A \subset \mathbb{R}^n$, <u>et</u> $\tau_A$ <u>la v.a.</u> $_A(y^\varepsilon)$ <u>correspondante associée au processus</u> $y^\varepsilon$.

2.8 <u>Lemme</u> : (cf. [53]) Hypothèses de 2.1, 2.7. Soit $\varphi : ]-\infty,0[ \to U$ un chemin minimisant $\lambda_{-\infty,0}$ et joignant 0 à $y_o \in \partial U$; soit

$V_o = \lambda_{-\infty,0}(\varphi)$. Alors pour tout $r > 0$ il existe $\alpha > 0$ tel que les relations (1) et (2)

(1) $f \in C_{0,T}(\mathbb{R}^n)$, $T > 0$, $\tau_U(f) \leqslant T$, $f_o = 0$

(2) $d_{-T,0}(\theta_T f, \varphi) \geqslant r$

entrainent $\lambda_{0,T}(f) \geqslant V_o + \alpha$

Preuve : Si le lemme n'était pas vrai, il existerait une suite $\alpha_k > 0$ tendant vers 0, une suite de temps $T_k > 0$ et une suite de fonctions $g^k \in C_{0,T_k}(\mathbb{R}^n)$ telles que

(3) $\tau_U(g^k) \leqslant T_k$ ; $\lambda_{0,T_k}(g^k) \leqslant V_o + \alpha_k$ ; $g_o^k = 0$;

(4) $\lim\limits_{k\to\infty} d_{-T_k,0}(\theta_{T_k} g^k, \varphi) = 0$

Posons $f^k = \theta_{T_k} g^k$, de sorte que $f^k$ est définie sur $[-T_k, 0]$. Prolongeons $f^k$ en fonction continue sur $]-\infty, 0]$, en posant $f_t^k = 0$ pour $t \leqslant -T_k$.

Pour N fini fixé, les $\lambda_{-N,0}(f^k)$ restent bornés par $(V_o + \sup\limits_k \alpha_k)$. Il existe donc une fonction continue $h^N$ sur $[-N,0]$ et une suite extraite (encore notée $f^k$) convergeant uniformément vers $h^N$ sur $[-N,0]$. Ceci entraîne

$$\lambda_{-N,0}(h^N) \leqslant \varliminf_{k\to\infty} \lambda_{-N,0}(f^k) \leqslant V_o$$

Recommencons des extractions successives de sous suites pour $N = 1, 2 \ldots$ Par le procédé diagonal classique on obtient finalement une sous suite (toujours notée $f^k$) qui converge uniformément vers $h^N$ sur $[-N,0]$ pour tout N. Alors les $h^N$ sont évidemment les restrictions à $[-N,0]$ d'une même fonction h continue sur $]-\infty,0]$. De plus on a

$$\lambda_{-\infty,0}(h) = \lim_{N\to\infty} \lambda_{-N,0}(h) = \lim_{N\to\infty} \lambda_{-N,0}(h^N) \leqslant V_o$$

Soit $B_r$ une boule ouverte de rayon $r \leqslant r_o$ petit, de centre 0. Puisque $\lambda_{-\infty,0}(f^k)$ est borné, le lemme 2.3 montre l'existence de $T=T(r)$ tel que $f^k[-a,0] \subset \overline{U}-B_r$ entraîne $a \leqslant T$. Il existe donc une suite $t_k$ dans $[-2T,0]$ telle que $f^k(t_k) \in B_r$. On peut joindre (lemme 2.5) le point 0 à tout point de $B_r$ par un chemin continu $g$ pour lequel $\lambda$ est majoré par $Cr$. En suivant un tel chemin de 0 à $f^k(t_k)$ (sur un intervalle de temps fini) puis $f^k$ de $f^k(t_k)$ à $f^k(0)$ on obtient un chemin $h : [-S,0] \to U$ tel que $\lambda_{-S,0}(h) \leqslant Cr + \lambda_{-2T,0}(f^k)$ et qui atteint $\partial U$ au temps 0. Par suite

$$V_o \leqslant \lambda_{-S,0}(h) \leqslant Cr + \lambda_{-2T,0}(f^k)$$

et donc

$$\lambda_{-\infty,-2T}(f^k) = \lambda_{-\infty,0}(f^k) - \lambda_{-2T,0}(f^k) \leqslant V_o + \alpha_k - (V_o - Cr) \leqslant \alpha_k + Cr$$

Il n'existe pas de trajectoires de $y_t' = b(y_t)$ partant de 0 et atteignant $\partial B_{r_o}$ en temps fini. Par des arguments déjà utilisés au lemme 1.3, on en déduit qu'il existe $a_o > 0$ tel que $\lambda(g) \geqslant a_o$ pour tout chemin $g$ joignant 0 à $\partial B_{r_o}$ en temps fini. Prenons $Cr \leqslant \frac{a_o}{3}$ et $k$ assez grand pour que $\alpha_k \leqslant \frac{a_o}{3}$, ce qui donne $\lambda_{-\infty,-2T}(f^k) \leqslant \frac{2a_o}{3}$. Ceci montre que $f^k[-\infty,-2T] \subset B_{r_o}$, ce qui entraîne immédiatement $h[-\infty,-2T] \subset \overline{B}_{r_o}$ où $T = T(r)$, $Cr \leqslant \frac{a_o}{3}$.

On en conclut d'abord ($r_o$ étant quelconque) que $h(t) \to 0$ quand $t \to -\infty$. Puisque $\lambda_{-\infty,0}(h) = V_o$, l'unicité de $\varphi$ minimisante entraîne $\varphi \equiv h$. Des relations $f^k[-\infty,-2T] \subset B_{r_o}$, $h[-\infty,-2T] \subset B_{r_o}$, et $f^k \to h$ uniformément sur $[-2T,0]$ on déduit que $f^k$ converge vers $h \equiv \varphi$ uniformément sur $]-\infty,0]$, ce qui contredit (4) et prouve le lemme.

2.9 Mise en place d'un tube et de deux barrières :

[La situation est celle de 2.1, 2.7]. Donnons nour $r > 0$ petit arbitraire. Determinons alors $\alpha > 0$ tel que le lemme 2.8 s'applique à $r$ et $\alpha$.

Soit $B_+$ une boule ouverte de rayon $\rho \leqslant \rho_o$ ; Pour $S > 0$ arbitraire posons

(1) $F(S) = \{g \in C_{0,S}(\mathbb{R}^n) \mid \tau_U(g) \leqslant S,\ g_o \in \overline{B}_+ ,\ d_{-S,0}(\theta_S g, \varphi) \geqslant r\}$

L'ensemble $F(S)$ est fermé dans $C_{0,S}(\mathbb{R}^n)$.

Soit $g \in F(S)$. Si $g_o = 0$ on a $\lambda_{0,S}(g) \geqslant V_o + \alpha$ d'après le lemme 2.8. Si $g_o \neq 0$, posons

$$\begin{cases} h(t) = t \dfrac{g_o}{(g_o)} & \text{pour } 0 \leqslant t \leqslant |g_o| \\ h(t) = g(t-|g_o|) & \text{pour } |g_o| \leqslant t \leqslant S + |g_o| \end{cases}$$

On a alors $\lambda_{0,S+|g_o|}(h) \leqslant C\rho + \lambda_{0,S}(g)$

(où C est une constante), et $h(0) = 0$, $\tau_U(h) \leqslant S + |g_o|$

Prenons $\rho$ assez petit pour que $C\rho \leqslant \frac{1}{16}\alpha$; si $\lambda_{0,S}(g)$ était strictement plus petit que $V_o + \frac{1}{2}\alpha$, on aurait

$$\lambda_{0,S+|g_o|}(h) < V_o + \alpha, \text{ et donc d'après le lemme 2.8,}$$

en notant $R = S + |g_o|$

$$d_{-R,0}(\theta_R h, \varphi) < r$$

d'où à fortiori $d_{-S,0}(\theta_R h, \varphi) < r$, ce qui puisque $\theta_R h$ et $\theta_S g$ coincident sur $[-S,0]$ contredit $g \in F(S)$.

Par suite on a

(2) $\qquad \lambda_{0,S}(g) \geqslant V_o + \frac{1}{2}\alpha$ pour tout $g \in F(S)$, tout $S > 0$.

Nous prendrons $\rho$ assez petit pour que tout segment de droite issu de $\overline{U}$ et de longueur moindre que $2\rho$, parcouru à vitesse 1 soit un chemin g tel que $\lambda(g) \leqslant \frac{\alpha}{32}$.

La fonction $\varphi$ est uniformément holderienne d'ordre $\frac{1}{2}$ sur $]-\infty,0]$, car $\varphi'$ est dans $L_2$. Si c est la constante associée, nous imposerons aussi $(2\rho + c\sqrt{\rho})$ suffisamment petit ce qui sera précisé plus bas.

Fixons désormais le rayon $\rho$ de la boule $B_+$. La trajectoire $\varphi$ étant minimisante ne repasse évidemment pas par 0 dans l'intervalle de temps $[\tau_{B_+}(\varphi),0]$. Il existe donc une boule fermée $\overline{B}_-$ de centre 0 disjointe du compact $\varphi([\tau_{B_+}(\varphi),0])$.

Fixons une telle boule $B_-$, telle que $\overline{B}_- \subset B_+$. Puisque le compact $\overline{U} - B_-$ ne contient pas de points $\omega$-limite (par hypothèse), le lemme 2.4 montre l'existence de $T_o > 0$, $\varepsilon_o > 0$ tels que pour $\varepsilon \leqslant \varepsilon_o$, $x \in \overline{U} - B_-$, $T \geqslant T_o$, on ait

(3) $\qquad P_x\{y^\varepsilon([0,T]) \subset \overline{U} - B_-\} \leqslant \exp\left(-\frac{V_o + 1}{\varepsilon^2}\right)$

Enfin pour $T \geqslant T_1$, on aura

(4) $\qquad \varphi_t \in B_+$ pour $t \leqslant -T$

Choisissons $T = T_1 \vee T_o$. Dans tout le paragraphe suivant les nombres $\alpha$, $r$, $\rho$, $T$, $\varepsilon_o$, les boules $B_+$, $B_-$ resteront fixés comme ci-dessus.

2.10 Etude du segment terminal de $y^\varepsilon$ :

Définissons les temps d'oscillations successives entre $B_+$ et $B_-$ par

$$\begin{aligned}
\tau_+^o &= \text{1ère sortie de } B_+ \\
\eta_-^1 &= \text{1ère entrée dans } B_- \text{ après } \tau_+^o \\
\tau_+^n &= \text{1ère sortie de } B_+ \text{ après } \eta_-^n \\
\eta_-^{n+1} &= \text{1ère entrée dans } B_- \text{ après } \tau_+^n
\end{aligned}$$

Soit $\nu$ le dernier entier $n \geqslant 0$ tel que $\tau_+^n < \tau_U$ qui existe et est fini dès que $\{\tau_+^o < \tau_U \,;\, \tau_U \text{ fini}\}$(il s'agit ici de trajectoires continues, tant pour les processus que pour les "chemins" considérés éventuellement). Nous poserons $\nu = +\infty$ si $\tau_+^o \geqslant \tau_U$ ou si $\tau_U$ est infini.

Le segment terminal d'une trajectoire $y^\varepsilon$ sera noté segter $(y^\varepsilon)$ :

c'est la restriction de $y^\varepsilon$ à $[\tau_+^\nu, \tau_+^\nu + T]$, translatée par $\tau_+^\nu$, c'est à dire segter $y^\varepsilon$ = restriction de $\theta_{\tau_+^\nu} y^\varepsilon$ à $[0,T]$ .

La variable aléatoire segter $y^\varepsilon$ est à valeurs dans $C_{0,T}(\mathbb{R}^n)$, et définie sur $\{\nu \text{ fini}\}$, fixée arbitraire sur $\{\nu \text{ infini}\}$ . Lorsque $y_o^\varepsilon = x \in \partial B_+$, l'évènement $\{T < \tau_U < \eta_-^1\}$ est inclus dans $\{y^\varepsilon[0,T] \subset \bar{U} - B_-\}$ et donc d'après 2.9 (3)

(1) $$P_x\{T < \tau_U < \eta_-^1\} \leqslant \exp\left(-\frac{V_o + 1}{\varepsilon^2}\right)$$

pour $x \in \partial B_+$, $\varepsilon \leqslant \varepsilon_o$. [<u>Toutes les probabilités</u> $P_x$ <u>concernent le processus</u> $y_t^\varepsilon$ <u>tel que</u> $y_o^\varepsilon = x$]

Notons $\Gamma$ le $[0,T]$ - tube ouvert d'axe $\theta_{-T}\varphi$ et de rayon r, c'est à dire l'ensemble des fonctions h définies <u>au moins</u> sur $[0,T]$ , à valeurs dans $\mathbb{R}^n$ et telles que $|h_t - \varphi_{t-T}| < r$ pour $t \in [0,T]$ Pour $y_o^\varepsilon = x \in \partial B_+$, avec la notation 2.9 (1), on a l'inclusion

(2) $$\{\tau_U < \eta_-^1 \ ; \ \tau_U \leqslant T \ ; \ y^\varepsilon \notin \Gamma\} \subset \{y^\varepsilon_{[0,T]} \in F(T)\}$$

Mais F(S) est fermé dans $C_{0,S}(\mathbb{R}^n)$ et vérifie, d'après 2.9 (2), $\Lambda_{0,S}[F(S)] \geqslant V_o + \frac{1}{2}\alpha$. D'après le cor. 1.5, où l'on pose $\eta = \frac{\alpha}{8}$ , $j = V_o + \frac{3\alpha}{8}$ , on peut donc trouver $\varepsilon_1 > 0$ tel que $\varepsilon \leqslant \varepsilon_1$, $x \in \partial B_+$ entraînent, en tenant compte de (2),

(3) $$P_x\{\tau_U < \eta_-^1 \ ; \ \tau_U \leqslant T \ ; \ y^\varepsilon \notin \Gamma\} \leqslant \exp\left(-\frac{V_o}{\varepsilon^2} - \frac{\alpha}{4\varepsilon^2}\right)$$

L'évènement $E = \{\tau_U < \eta_-^1\}$ est inclus dans

$$E \cap [(\tau_U > T) \cup (\tau_U \leqslant T \ ; \ y^\varepsilon \notin \Gamma) \cup (y^\varepsilon \in \Gamma)]$$

de sorte que (1) et (3) entraînent pour $\varepsilon \leqslant \varepsilon_o \wedge \varepsilon_1$, $x \in \partial B_+$ ,

(4) $$P_x(E) \leqslant \exp\left(-\frac{V_o+1}{\varepsilon^2}\right) + \exp\left(-\frac{V_o}{\varepsilon^2} - \frac{\alpha}{4\varepsilon^2}\right) + P_x\{E \cap (y^\varepsilon \in \Gamma)\}$$

Grâce à 2.9 (4), on peut trouver un chemin $\Psi$ tel que $\Psi_o = x$, $\Psi_{S_1} = \varphi_{-T}$ , $S_1 \leqslant \rho$, et $\lambda_{0,S_1}(\Psi) \leqslant \frac{\alpha}{32}$ . On pose $\Psi_{S_1+t} = \varphi_{-T+t}$ pour $0 \leqslant t \leqslant T$ ; puis on prolonge $\Psi$ sur $[T+S_1, S]$ par un petit segment de droite sortant de U et atteignant un point à distance $2\rho$ de $\partial U$, en s'assurant que $S - T \leqslant 3\rho$ et $\lambda_{T+S_1,S}(\Psi) \leqslant \frac{\alpha}{32}$. On a alors $\lambda_{0,S}(\Psi) \leqslant V_o + \frac{\alpha}{16}$. Un calcul simple montre que $d_{-T,0}(\theta_T\Psi,\varphi)$ est majoré par la somme de $2\rho$ et de l'oscillation maximale de $\varphi$ sur un intervalle de longueur $\rho$, donc (puisque $\varphi'$ est de carré intégrable sur $]-\infty,0]$) par $2\rho + c\sqrt{\rho}$ où la constante ne dépend que de $\varphi$. On peut toujours supposer qu'au moment du choix de $B_+$, on a imposé $2\rho + c\sqrt{\rho} \leqslant \frac{r}{2}$. Par suite, tout chemin g appartenant au $[0,S]$ - tube ouvert d'axe $\Psi$ et de rayon inférieur à $\frac{r}{2}$ vérifiera $g \in \Gamma$. D'autre part par le choix même de $B_-$, le compact $\varphi[\tau_{B_+}(\varphi),0]$ est disjoint de $\overline{B}_-$. Il est alors clair que si $2\rho + c\sqrt{\rho}$ est suffisamment petit, le chemin $\Psi$ et tout chemin g appartenant à un $[0,S]$ - tube d'axe $\Psi$ suffisamment étroit vérifiera $\tau_U(g) < \eta^1_-(g)$. Le rayon d'un tel tube peut être choisi indépendant de $x \in \partial B_+$, bien que $\Psi$ dépende ("peu") de x.

La prop. 1.2 montre alors l'existence de $\varepsilon_2 > 0$ tel que $x \in \partial B_+$, $\varepsilon \leqslant \varepsilon_2$ impliquent

(5) $P_x\{E \cap (y^\varepsilon \in \Gamma)\} \geqslant \exp(-\frac{\lambda_{0,S}(\Psi)}{\varepsilon^2} + \frac{\alpha}{32\varepsilon^2}) \geqslant \exp(-\frac{V_o}{\varepsilon^2} - \frac{\alpha}{8\varepsilon^2})$

De (4) et (5) on déduit pour $x \in \partial B_+$ $\varepsilon \leqslant \varepsilon_o \wedge \varepsilon_1 \wedge \varepsilon_2 = \varepsilon_3$

(6) $P_x(E) \leqslant \exp(-\frac{V_o}{\varepsilon^2} - \frac{\alpha}{8\varepsilon^2})\left[1 + 2\exp(-\frac{\alpha}{8\varepsilon^2})\right]$

d'où pour $x \in \partial B_+$, $\varepsilon \leqslant \varepsilon_3$

(7) $q\, P_x(\tau_U < \eta^1_-) \leqslant P_x(\tau_U < \eta^1_- \; ; \; y^\varepsilon \in \Gamma)$

avec $q = 1 - 2\exp(-\frac{\alpha}{8\varepsilon^2})$

Posons $y_n = y^\varepsilon_{\tau^n_+}$. La propriété de Markov forte, en espace temps, montre que la variable aléatoire

(8) $Z_n = P_{y_n}(\tau_u < \eta^1_- \; ; \; y^\varepsilon \in \Gamma)$

s'écrit aussi, sur l'ensemble $\tau^n_+ < \tau_U$,

(9) $Z_n = P\left[\text{segter } y^{\varepsilon} \in \Gamma \ ; \ \tau_+^n < \tau_U < \eta_-^{n+1} \ \middle| \mathcal{F}_{\tau_+^n} \right|$

où $\mathcal{F}_{\tau_+^n}$ est l'algèbre des évènements antérieurs à $\tau_+^n$.

De (7), (8) et (9) on déduit, sur $\{\tau_+^n < \tau_U\}$

(10) $P\left[\text{segter } y^{\varepsilon} \in \Gamma ; \ \nu = n \ \middle| \mathcal{F}_{\tau_+^n}\right] \geqslant q \ P_{y_n} \ (\tau_U < \eta_-^1 )$

Mais on a sur $\{\tau_+^n < \tau_U\}$

(11) $P_{y_n}(\tau_U < \eta_-^1) = P(\tau_+^n < \tau_U < \eta_-^{n+1} \ |\mathcal{F}_{\tau_+^n}) = P(\nu = n \mid \mathcal{F}_{\tau_+^n})$

De (10) et (11) on déduit, en prenant les espérances ,

$$P_z\left[\text{segter } y^{\varepsilon} \in \Gamma ; \ \nu = n\right] \geqslant q \ P_z(\nu = n)$$

pour tout $z \in U$, tout $n \geqslant 0$. En sommant sur n on obtient pour $z \in U$, $\varepsilon \leqslant \varepsilon_3$.

(12) $P_z(\text{segter } y^{\varepsilon} \in \Gamma) \geqslant q \ P_z(\nu \text{ fini})$

L'évènement $\{\nu = +\infty\}$ est $P_z$ - p.s. égal à $\{\tau_U < \tau_+^o\}$ . Donc si $z \in B_+$ on a $P_z(\nu \text{ fini}) = 1$.

Si $z \in U - B_+$ il est facile de voir que

$P_z(\tau_u < \tau_+^o) \leqslant \exp(- \frac{cte}{\varepsilon^2})$ où la constante est positive stricte et dépend de z à priori. Dans tous les cas (12) donne donc pour $z \in U$, $\varepsilon \leqslant \varepsilon_3$

(13) $P_z(\text{segter } y^{\varepsilon} \in \Gamma) \geqslant 1 - \exp\ (- \frac{cte}{\varepsilon^2})$

En particulier $\lim\limits_{\varepsilon \to 0} P_z \ (\text{segter } y^{\varepsilon} \in \Gamma) = 1$.

2.11. Reformulation du résultat

Comme plus haut, soit $\Gamma$ le $[0\ T]$ -tube ouvert d'axe $\Theta_{-T}\varphi$ , de rayon $r$ , et soit $\rho$ le rayon de $B_+$ . Soit $V_a$ la boule de centre 0 et de rayon $(2a+\rho)$. Montrons que

(1) pour tout $a > 0$, il existe $r_o > 0$, $T_o > 0$ tels que $r \leqslant r_o$ et $T > T_o$

$g \in \Gamma$ , $g_o \in \partial B_+$ , $\tau_U(g) \leqslant T$ entraînent

$|g_{s+\tau_U(g)} - \varphi_{s+\tau_U(\varphi)}| \leqslant a$ pour tout $s$ tel que

$\tau_{V_a}(g) \leqslant s + \tau_U(g) \leqslant \tau_U(g)$ ;

En effet, notons d'abord que par convention, nous avons $\tau_U(\varphi) = 0$. Soit $W$ un voisinage ouvert régulier de 0 tel que $\overline{W} \subset U$ et $|x-y| \geqslant a_1$ pour $x \in \partial U$, $y \in \partial W$.

Pour tout $a_2 > 0$ , on peut trouver $a_1$ et $W$ pour garantir $\tau_U(\varphi) - \tau_W(\varphi) \leqslant a_2$.

Pour $g \in \Gamma$ , on a $|g\,[\tau_U(g)] - \varphi\,[\tau_U(g) - T]| < r$ , et donc $\tau_U(g) - T \geqslant \tau_W(\varphi)$ si $r < a_1$ , ce qui entraîne $0 \geqslant \tau_U(g) - T \geqslant -a_2$.

Pour tout $a_3 > 0$, on peut trouver $a_2$ assez petit pour que l'oscillation de $\varphi$ sur les intervalles de longueur $a_2$ soit majorée par $a_3$. On a alors pour $t \leqslant 0$, $|t| \leqslant T - a_2$

$$|g_{t+\tau_U(g)} - \varphi_t| \leqslant |g_{t+\tau_U(g)} - \varphi_{t+\tau_U(g)-T}| + |\varphi_{t+\tau_U(g)-T} - \varphi_t|$$

ce qui est majoré par $r + a_3$ , grâce à $g \in \Gamma$ . Pour tout $a_4 > 0$ , fixons $a_3$ et $r$ assez petits pour que $r + a_3 < a_4$. On a alors l'inégalité (1) annoncée, pour les temps $s + \tau_U(g)$ vérifiant

$$2\,a_2 \leqslant s + \tau_U(g) \leqslant \tau_U(g)$$

Il est facile d'en déduire que, si l'on a pris $T$ assez grand pour que $\tau_{V_{a_5}}(\varphi) \geqslant -T + 2\,a_2$ , on aura pour $a$, $a_5$ convenables l'inégalité (1) sur $\tau_{V_a}(g) \leqslant s + \tau_U(g) \leqslant \tau_U(g)$.

On peut donc formuler le résultat de Ventsel et Freidlin sous la forme suivante, "interne" à U ce qui est utile pour rendre caduques les conventions adoptées sur le comportement des coefficients en dehors de U.

2.12. Théorème (Ventsel-Freidlin [53] ) : Soit U un voisinage très régulier d'un point d'équilibre stable pour le système dynamique $y'_t = b\,(y_t)$ (hypothèses 2.1, 2.7). Soit $\lambda_{S,T}$ la fonctionnelle de Cramer associée au système perturbé $(E^\varepsilon)$. On suppose l'existence d'une unique $\varphi : \,]-\infty, 0[\, \longrightarrow U$ joignant 0 à $\partial U$ et minimisant $\lambda$.

Soit K un voisinage compact arbitraire de 0 dans U, et soit $\alpha_K$ le temps de dernière sortie de K avant $\tau_U$. Considérons l'évènement

$$E = \{ |y^\varepsilon_{t+\tau_U(y^\varepsilon)} - \varphi_t| \leqslant r \text{ pour } \alpha_K(y^\varepsilon) \leqslant t + \tau_U(y^\varepsilon) \leqslant \tau_U(y^\varepsilon) \}$$

Alors, pour tout $r > 0$, pour tout $z \in U$, $P_z$ (E) tend vers 1 (vitesse en $\exp\left(-\frac{cte}{\varepsilon^2}\right)$) lorsque $\varepsilon \to 0$. La convergence est uniforme pour z dans une partie compacte arbitraire de U.

2.13. Commentaire : On a construit ainsi un tube d'axe $\varphi$ qui constitue un goulot de sortie, le seul canal par lequel la sortie de U puisse se faire. En particulier le point de sortie $y^\varepsilon_{\tau_U(y^\varepsilon)}$ tend en probabilité (à vitesse $\exp - \frac{c}{\varepsilon^2}$) vers $\varphi(0) \in \partial U$ qui est le point de $\partial U$ où le "quasi-potentiel" V (0,y) défini en 2.6, 2.7 atteint son minimum.

## 3. CALCUL DES EXTREMALES ET DU QUASI-POTENTIEL

3.1. Les résultats de ce paragraphe, sont dus à Ventsel-Freidlin [54] Nous ne donnerons pas de preuves rigoureuses renvoyant à [54] où les dites preuves sont d'ailleurs parfois omises ou très brèves ; pour des approches rigoureuses détaillées concernant le type d'équations aux dérivées partielles qui interviennent ici, nous renvoyons à P.L. Lions [32] [33] et à la bibliographie de [32] [27] , particulièrement Krylov. Nous préférons esquisser des raisonnements heuristiques qui feront peut être sentir plus rapidement pourquoi de telles équations interviennent ici.

3.2. La situation considérée est celle de 2.1. On définit comme en 2.6, 2.7 le quasi potentiel $V : \mathbb{R}^n \to [0, +\infty[$

$V(y) = \inf \{\lambda_{0,T}(g) \mid g_0 = 0,\ g_T = y\ ,\ T > 0 \text{ arbitraire}\}$ .

En général sous les hypothèses 2.1, la fonction V est seulement Lipschitzienne (localement) sur $\mathbb{R}^n$ (cf. [53] ).

Mais si une certaine équation aux dérivées partielles admet une solution, alors cette solution coïncide avec V. Construisons "à la main" cette équation.

3.3. Une équation aux dérivées partielles "vérifiée" par V

Considérons deux courbes de niveau $L_a$ et $L_{a-\Delta a}$ de la fonction V , avec $\Delta a > 0$ petit. Soit $x \in L_a$ et soit $\varphi$ une courbe minimisant $\lambda$ , joignant 0 à x sur l'intervalle de temps $]-\infty, T]$. Soit $(T - \Delta T)$ le temps de sortie de $L_{a-\Delta a}$ pour $\varphi$ , de sorte que $y = \varphi_{T-\Delta T} \in L_{a-\Delta a}$ . Soit z un point de $L_{a-\Delta a}$ et $\psi : [S - \Delta S, S] \to \mathbb{R}^n$ un chemin joignant z à x. Puisque $z \in L_{a-\Delta a}$ , il existe un chemin $g : ]-\infty, S - \Delta S]$ joignant 0 à z et vérifiant $\lambda_{-\infty, S-\Delta S}(g) = a - \Delta a$. Le chemin obtenu en suivant g puis $\psi$ joint 0 à $x \in L_a$ , ce qui force donc

(1) $$\lambda_{S-\Delta S,S}(\psi) \geqslant \Delta a$$

Comme $\varphi$ est minimisante entre 0 et $L_a$ , on en déduit que

(2) $$\lambda_{T-\Delta T,T}(\varphi) = \Delta a = \inf_{\psi,\Delta S} \lambda_{S-\Delta S,S}(\psi)$$

où $\psi$ joint un point quelconque z de $L_{a-\Delta a}$ à x donné sur $L_a$.

"Au second ordre près" on peut écrire

$$- \Delta a = V(y) - V(x) = < (y-x), V'(x) > = - \Delta T < \varphi'_T , V'(x) >$$

$$- \Delta a = V(z) - V(x) = < (z-x), V'(x) > = - \Delta S < \psi'_S , V'(x) >$$

d'où

(3) $$\Delta a = \Delta S < \psi'_S , V'(x) > = \Delta T < \varphi'_T , V'(x) >$$

où les produits scalaires et les gradients V'(x) sont calculés dans la métrique associée à la forme quadratique $Q^*_x$ de matrice $[\sigma(x)\sigma^*(x)]^{-1}$. En notant $|| \quad ||$ la norme associée à $Q^*_x$ , on a aussi, "au second ordre près"

(4) $$\begin{cases} \lambda_{T-\Delta T,T}(\varphi) = \frac{1}{2} \Delta T \, || \varphi'_T - b(x)||^2 \\ \lambda_{S-\Delta S,S}(\psi) = \frac{1}{2} \Delta S \, ||\psi'_S - b(x)||^2 \end{cases}$$

de sorte que (2) entraîne

(5) $$\Delta T \, || \varphi'_T - b(x)||^2 = \inf_{\psi,\Delta S} \Delta S \, ||\psi'_S - b(x)||^2$$

(6) $$\frac{1}{2} \Delta T \, || \varphi'_T - b(x)||^2 = \Delta a$$

grâce à (3), les égalités (5), (6) deviennent

(7) $$\frac{|| \varphi'_T - b(x)||^2}{< \varphi'_T , V'(x) >} = \inf_{\psi'_S} \frac{||\psi'_S - b(x)||^2}{< \psi'_S , V'(x)>} = \inf_{v \in \mathbb{R}^n} \frac{||v - b(x)||^2}{< v , V'(x)>}$$

(8) $$\frac{|| \varphi'_T - b(x)||^2}{< \varphi'_T , V'(x) >} = 2$$

Pour abréger l'écriture, posons $b(x) = b$ , $V'(x) = V'$. Le gradient de

$v \longrightarrow \dfrac{||v-b||^2}{< v,V'>}$ (x est fixé dans la suite) s'écrit

$2 \dfrac{v-b}{< v,V' >} - \dfrac{||v-b||^2}{< v,V' >^2} V'$ . D'après (7) ce gradient doit s'annuler

pour $v = \varphi'_T$ , ce qui en tenant compte de (8) donne en posant $v = \varphi'_T$

(9) $\qquad v = \varphi'_T = b + V'$

De (8) (9) on tire alors

$$||V'||^2 = 2 < b + V' \, , V' > \qquad \text{d'où}$$

(10) $\qquad ||V'||^2 + 2 < b, V' > = 0$

Il est facile de voir que (9) (10) équivalent à (7) (8). Notons une conséquence de (9) (10), tout à fait immédiate,

(11) $\qquad || \varphi'_T||^2 = ||b + V'||^2 = ||b||^2$

Nous pouvons donc énoncer le résultat "heuristique" suivant

3.4 "Résultat heuristique" Le quasi potentiel V (y) et les trajectoires $\varphi$ issues de 0 minimisant $\lambda$ entre 0 et leur point d'arrivée "doivent" vérifier les équations suivantes :

(12) $\qquad ||V'(x)||^2_x + 2 < b(x), V'(x) >_x = 0$

(produits scalaires et gradient V' relatifs à $Q^*_x$ de matrice $[\sigma (x) \sigma^*(x)]^{-1}$)

(13) $\qquad \varphi'_t = b(\varphi_t) + V'(\varphi_t)$

En particulier $||\varphi'_t|| = ||b(\varphi_t)||$

Remarquons que ces équations deviennent très simples lorsque le drift b(x) est un champ de gradient pour la métrique $Q^*_x$ ; en effet si

$b(x) = H'(x)$ avec $H : \mathbb{R}^n \to \mathbb{R}$ de classe 1, on voit que $V(x) = -2H(x)$ est une solution triviale de (12), et qu'on a alors $b + V' = -H' = -b$ d'où $\varphi'_t = -b(\varphi_t)$. Il est facile de voir alors directement (cf [53]) que V et $\varphi$ sont bien les solutions du problème cherché. Intuitivement les chemins de sortie sont les plus brutaux possibles : ils "remontent" le champ b avec une vitesse juste suffisante pour compenser le drift.

En ce qui concerne les questions d'existence et unicité des solutions de (12) nous renvoyons à [32] [27]. Le résultat rigoureux associé à 3.4 est énoncé dans [54] .

3.5. <u>Théorème</u> : (Ventsel-Freidlin [53] [54]). Supposons (hypothèses 2.1 sur $b(x)$, $\sigma(x)$, U ouvert de $\mathbb{R}^n$) l'existence d'une fonction V définie et continue sur un voisinage de U, strictement positive sur $\overline{U} - \{0\}$, nulle en 0, et qui sur $W - \{0\}$ , où W est un voisinage de $\overline{U}$ , est de classe 1, de gradient V' non nul vérifiant l'équation aux dérivées partielles (12). Alors V(x) coïncide avec le quasi-potentiel [associé au système perturbé $(E^\varepsilon)$ ] sur l'ensemble $\tilde{U}$ des $x \in \overline{U}$ tels que $V(x) \leqslant \inf_{y \in \partial U} V(y)$.

Pour tout autre x , on peut seulement affirmer que $\inf_{y \in \partial U} V(y)$ est un minorant du quasi-potentiel cherché. Si $x \in \tilde{U}$ , les trajectoires $\varphi$ solution de (13) sur $]-\infty, T]$ et vérifiant $\varphi_T = x$ sont effectivement des trajectoires minimisant la fonctionnelle de Cramer entre 0 et x.

<u>Preuve</u> : nous renvoyons à [53] [54] pour la preuve, ainsi que pour quelques exemples explicites. Nous verrons d'autres exemples plus géométriques au chapitre V.

## 4. STABILISATION (ET DESTABILISATION) DES EQUILIBRES

(d'après Ventsel-Freidlin)

4.1. Nous esquissons pratiquement sans preuves quelques jolis résultats de Ventsel-Freidlin [54] qui font sentir de façon frappante la portée pratique des résultats sur les goulots et temps de sortie pour les équilibres perturbés. Pour les détails nous renvoyons à [54] dont la lecture est vivement conseillée aux "utilisateurs" éventuels de la théorie dans "l'esprit" du contrôle optimal.

4.2. Comportement du temps de sortie (cf [54] ) :

La situation étant celle de 2.1, on montre que

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log E\,[\tau_U\,(y^\varepsilon)] = V\,(y_o) = \inf_{y\in\partial U} V(y)$$

On peut alors poser un premier type de problème [54]

(1) Etant donné un champ b sur $\mathbb{R}^n$ , un champ de matrices inversibles $\sigma$ sur $\mathbb{R}^n$ (décrivant la structure des perturbations possibles), trouver un voisinage ouvert U de l'équilibre stable 0 de volume donné C , et pour lequel le temps de sortie $\tau_U\,(y^\varepsilon)$ soit "asymptotiquement maximal", c'est-à-dire telle que $\lim\limits_{\varepsilon\to 0} \varepsilon^2 \log E\,[\tau_U\,(y^\varepsilon)]$ soit maximal.

Par exemple si C représente le volume "controlâble à coût donné", on cherche ici la forme de la zone de contrôle la plus sûre, à coût donné. On peut remplacer le volume par un coût du type $\int_U h(x)\,dx$ où $h > 0$ est une fonction fixée.

La solution du problème (1) est simple ; la région U optimale est de la forme $U = \{x \mid V(x) \leqslant cte\}$ où V est le quasi-potentiel du problème (cf [54]).

4.3. <u>Choix optimal du drift</u> (cf [54])

En chaque point $x \in \mathbb{R}^n$ on se donne un ensemble $\Pi(x) \subset \mathbb{R}^n$ de drifts "possibles", et une structure de perturbation $\sigma(x)$, matrice inversible $(n,n)$. On se donne un ouvert $U$ de $\mathbb{R}^n$. On pose le problème

(2) Chercher un champ $b$ tel que $b(x) \in \Pi(x)$ pour $x \in U$, tel que $b$ admette un point d'équilibre stable $x_o$, et tel que le temps de sortie $\tau_U(y^\varepsilon)$ du système perturbé associé à $(b, \sigma, \varepsilon)$ soit asymptotiquement maximal.

On peut aussi poser le problème analogue

(3) Chercher $b$ comme en (2), mais tel que $\tau_U(y^\varepsilon)$ soit asymptotiquement minimal.

Dans le problème (2) il s'agit de chercher une structure dynamique <u>résistant de façon optimale aux perturbations</u> ; dans le problème (3) il s'agit de trouver une structure dynamique stable <u>qui détecte le plus vite possible la présence d'une perturbation</u>.

Ventsel et Freidlin traitent (2) (3) en introduisant les quasi potentiels "enveloppants"

$$\overline{V}(x_o,x) = \sup\{V_b(x_o,x) \mid b \text{ champ acceptable, nul en } x_o\}$$
$$\underline{V}(x_o,x) = \inf\{V_b(x_o,x) \mid b \text{ champ acceptable, nul en } x_o\}$$

Il est possible d'écrire explicitement les équations aux dérivées partielles que "doivent" vérifier $\overline{V}(x_o, x)$ et $\underline{V}(x_o, x)$ en tant que fonctions de $x$, et de donner des conditions suffisantes assurant que les solutions de ces équations soient les fonctions $\overline{V}$, $\underline{V}$ cherchées. C'est ensuite à partir de $\overline{V}$, $\underline{V}$ que l'on peut calculer les champs $\overline{b}$, $\underline{b}$ optimaux pour les problèmes (2) (3). Nous renvoyons à [54] pour plus de détails.

## CHAPITRE V

# EXTENSION AU CAS DES VARIETES ET DIFFUSIONS EN TEMPS PETIT

### 1. DIFFUSIONS SUR UNE VARIETE

1.1. Définitions : Soit M une variété différentiable connexe, de dimension $n$, de classe 2. Soit $\Delta$ un opérateur différentiel du second ordre sur M, annulant les constantes, et semi-elliptique. Ceci équivaut à dire que pour toute carte locale $\varphi : U \to \mathbb{R}^n$, U ouvert de M, l'image $\Delta_\varphi$ de $\Delta$ par $\varphi$ peut s'écrire

$$(1)\quad \Delta_\varphi = \frac{1}{2} \sum_{1 \leq i,j \leq n} a_{ij}(x) \frac{\partial}{\partial x_i\, \partial x_j} + \sum_{1 \leq i \leq n} h_i(x) \frac{\partial}{\partial x_i}$$

où la matrice symétrique $a_\varphi(x) = [a_{ij}(x)]$ est positive au sens large ; nous noterons $h_\varphi(x)$ le vecteur de coordonnées $h_i(x)$. Les coefficients $a_\varphi$, $h_\varphi$ de $\Delta$ seront toujours supposés (au moins) boréliens et localement bornés.

Soit $\mathcal{C}(M)$ l'espace des chemins continus "explosifs" à valeur dans M (cf. III.1.2). Nous appelons $\Delta$-diffusion sur M tout processus de Markov fort, à trajectoires p.s. dans $\mathcal{C}(M)$, dont le semi-groupe de transition $R_t$ vérifie

$$(2)\quad R_t f(x) - f(x) = \int_0^t R_s \Delta f(x)\, ds \quad , \qquad x \in M$$

pour toute fonction f à support compact, de classe 2, définie sur M.

Notons $X_t = \mathcal{C}(M) \to M \cup \delta$ les applications coordonnées et $\mathcal{F}_t \subset \mathcal{B}[\mathcal{C}(M)]$ la $\sigma$-algèbre des évènements antérieurs à t. Notons $P_x$ la probabilité sur $(\mathcal{C}(M), \mathcal{B}[\mathcal{C}(M)])$ qui définit la loi des trajectoires issues de $x \in M$ pour la $\Delta$-diffusion considérée. Alors il est classique ([49] [2] [19] [40]) et immédiat que (2) équivaut à dire :

(3) Le processus $\{(X_t), (\mathcal{F}_t), (P_x)\}$ est Markov fort à trajectoires dans $\mathcal{C}(M)$ et pour toute fonction $f : M \to \mathbb{R}$, à support compact, de classe 2, le processus $t \to f(X_t) - \int_0^t \Delta f(X_s)\, ds$ est une $\mathcal{F}_t$-martingale pour la loi $P_x$.

On peut aussi dire que $P_x$ est solution du $\Delta$-problème des martingales.

Deux $\Delta$-diffusions seront dites équivalentes si pour tout $x \in M$, les lois $P_x$, $P'_x$ de leurs trajectoires coïncident sur $\mathcal{C}(M)$.

Si $X = \{(X_t)\ (\mathcal{F}_t)\ (P_x)\}$ est un processus de Markov fort à trajectoires dans $\mathcal{C}(M)$, alors le processus induit $X^U$ sur un ouvert $U$ quelconque de $M$ ("défini" par $X_t^U = X_t$ si $t < \tau_U$ et $X_t = \delta$ si $t \geqslant \tau_U$) est un processus de Markov fort à trajectoires dans $\mathcal{C}(U)$ (cf. [13] [18]).

On sait que si $X$ est une $\Delta$-diffusion, alors $X^U$ est une $\Delta$-diffusion sur $U$, et inversement, que si il existe pour chaque $U$ d'un recouvrement ouvert de $M$ une $\Delta$-diffusion sur $U$ unique à équivalence près, alors il existe une $\Delta$-diffusion sur $M$ unique à équivalence près (cf. [1] [2] [19] [40]).

Explicitions deux types d'hypothèses locales sur $\Delta$ qui réalisent cette situation.

## 1.2. Deux bonnes classes d'opérateurs $\Delta$ :

(1) Cas elliptique : on suppose que pour toute carte locale $\varphi : U \to \mathbb{R}^n$, les coefficients $a_\varphi$, $h_\varphi$ de $\Delta$ vérifient :

- la matrice $a_\varphi(x)$ est définie positive pour tout $x \in \varphi(U)$
- $a_\varphi$ est continue, $h_\varphi$ est borélien localement borné

(2) Cas hypoelliptique : on suppose que M est de classe $C^\infty$ et qu'il existe une base d'ouverts de M telle que sur chaque ouvert V de cette base on puisse écrire $\Delta = \sum_{i=1}^{r} X_i^2 + Y$ où $X_1 \dots X_r$ , Y sont des champs de vecteurs de classe $C^\infty$ sur V , et où l'algèbre de Lie engendrée par $X_1 \dots X_r$ , Y est de dimension n = dim M en chaque point de V.

Notons une forme équivalente plus explicite de (2) :

(2 bis) M est de classe $C^\infty$ et il existe un atlas de cartes locales $\varphi$ ayant la propriété suivante : $h_\varphi$ est $C^\infty$ ; $a_\varphi$ s'écrit $a_\varphi = \sigma_\varphi \, \sigma^*_\varphi$ où $\sigma_\varphi (x) = [\sigma_{i,j} (x)]$ est un champ $C^\infty$ de matrices rectangulaires ; le champ de vecteurs $y_\varphi$ défini par

$$y_\varphi (x) = \sum_i \left[h_i (x) - \sum_{j,k} \sigma_{j,k} (x) \frac{\partial \sigma_{i,k} (x)}{\partial x_j}\right] \frac{\partial}{\partial x_i}$$

et les vecteurs colonnes de $a_\varphi$ engendrent (par crochet de Lie) une algèbre de Lie qui est de dimension n en tout point.

Comme $\sigma_\varphi$ et $a_\varphi$ ont même image, on peut d'ailleurs remplacer ci-dessus les vecteurs colonnes de $a_\varphi$ par ceux de $\sigma_\varphi$ .

Pour le cas elliptique, l'existence (locale) d'une $\Delta$-diffusion unique à équivalence près se déduit de [49] (cf [1] [2] [40]). Pour le cas hypoelliptique, les spécialistes ne prennent généralement pas la peine de donner d'énoncé prècis. L'existence locale d'une $\Delta$-diffusion se déduit de l'existence de solution unique pour les équations stochastiques du type $d\,x_t = \sigma_\varphi(x_t)\, d\,\beta_t + h_\varphi (x_t)\, dt$ où $\sigma_\varphi$ est comme en (2 bis). L'unicité locale (à équivalence près) est une conséquence d'un résultat de Bony [9] sur les ouverts très réguliers. Nous renvoyons à [46] pour une preuve plus détaillée de la proposition suivante.

1.3. Proposition : Soit M une variété différentiable connexe, $\Delta$ un opérateur différentiel semi-elliptique du second ordre sur M , tel que $\Delta 1 = 0$. Dans le "cas elliptique 1.2.(1)" , et dans le "cas hypoelliptique 1.2.(2)" , il existe une $\Delta$-diffusion sur M , unique à équivalence près.

Rappelons que si on a existence et unicité des solutions des équations stochastiques locales

$$d\, x_t = \sigma_\varphi(x_t)\, d\, \beta_t + h_\varphi\,(x_t)\, dt \quad , \quad x_o = x \quad \text{où} \quad x \in \varphi(U)$$

ce qui est le cas dès que $\sigma_\varphi$ , $h_\varphi$ sont localement lipschitziens, alors $x_t$ "est" une $\Delta_\varphi$-diffusion sur $\varphi\,(U)$ , où $\Delta_\varphi$ est comme en 1.1.(1) avec $a_\varphi = \sigma_\varphi \sigma_\varphi^*$ ; en particulier d'après 1.3, le processus $\varphi^{-1}\,(x_t)$ est équivalent au processus induit sur U par la $\Delta$-diffusion de M.

## 2. SYSTEMES DYNAMIQUES PERTURBES SUR M

2.1. Le modèle : Soit M une variété différentiable connexe de classe 2, de dimension n. Soit $\Delta$ un opérateur différentiel du second ordre sur M , semi-elliptique, tel que $\Delta 1 = 0$. Soient b et $b_\varepsilon$ , $\varepsilon > 0$ des champs de vecteurs sur M tels que $\lim_{\varepsilon \to 0} b_\varepsilon = b$ où la convergence est uniforme sur tout compact de M . Nous allons considérer les $D_\varepsilon$-diffusions , où $D_\varepsilon = b_\varepsilon + \varepsilon^2 \Delta$ , comme des perturbations du système dynamique $y'_t = b\,(y_t)$ sur M . Nous utiliserons deux types d'hypothèses (notations 1.1.(1) pour l'image $\Delta_\varphi$ de $\Delta$ par une carte $\varphi$ ).

(1) Cas elliptique :

- Les champs b et $b_\varepsilon$ , $\varepsilon > 0$ sont localement lipschitziens, et $\lim_{\varepsilon \to 0} b_\varepsilon = b$ uniformément sur tout compact
- Il existe un atlas de M tel que pour toute carte $\varphi$ de cet atlas le champ $h_\varphi$ est localement lipschitzien, les matrices $a_\varphi\,(x)$ sont

définies positives, et s'écrivent $a_\varphi = \sigma_\varphi \sigma_\varphi^*$ où $\sigma_\varphi$ est un champ de matrices rectangulaires de classe 1 en x.

(2) Cas hypoelliptique :

- la variété M , et les champs $b_\varepsilon$ sont $C^\infty$
- le champ b est localement lipschitzien et $\lim\limits_{\varepsilon \to 0} b_\varepsilon = b$ uniformément sur tout compact.
- pour chaque $\varepsilon > 0$ , l'opérateur $D_\varepsilon = b_\varepsilon + \varepsilon^2 \Delta$ se met localement sous la forme $\sum\limits_{i=1}^{r} X_i^2 + Y$ où l'algèbre de Lie engendrée par les champs $C^\infty$ $X_1, \ldots, X_r$ , Y est de dimension n en tout point (de l'ouvert où $X_1 \ldots X_r$ Y sont définis).

Notons que cette dernière assertion est automatiquement vérifiée pour tout $\varepsilon > 0$ si pour toutes les cartes $\varphi$ d'un atlas de M , les coefficients de $\Delta_\varphi$ vérifient : $\{h_\varphi$ est $C^\infty$ ; $a_\varphi = \sigma_\varphi \sigma_\varphi^*$ où $\sigma_\varphi$ est un champ $C^\infty$ de matrices rectangulaires ; l'algèbre de Lie engendrée par les colonnes de $a_\varphi$ (ou bien par celles de $\sigma_\varphi$ ) est de dimension n en tout point$\}$.

Dans les cas (1) et (2) nous noterons $(y_t^\varepsilon)$ la $D_\varepsilon$ -diffusion sur M définie à équivalence près par $D_\varepsilon = b_\varepsilon + \varepsilon^2 \Delta$ . D'après 1.3, le processus induit par $y_t^\varepsilon$ sur le domaine U d'une carte $\varphi$ telle que $a_\varphi = \sigma_\varphi \sigma_\varphi^*$ avec $\sigma_\varphi$ comme ci-dessus est équivalent à l'image par $\varphi^{-1}$ du processus de Markov fort solution de $d z_t^\varepsilon = \varepsilon \sigma_\varphi(z_t^\varepsilon) + h_\varphi (z_t^\varepsilon)\, dt$ dans l'ouvert $\varphi(U) \subset \mathbb{R}^n$ .

2.2. Transformée de Cramer de $(D_\varepsilon)_{\varepsilon>0}$ :

Plaçons nous dans la situation 2.1.(1) ou bien 2.1.(2). Pour f de classe 2 au voisinage de $x \in M$ , l'expression $[\Delta (f^2)(x) - 2 f(x) \Delta f(x)]$ ne dépend que de $d f_x$ . On définit alors une forme quadratique $Q_x$ sur l'espace cotangent $T_x^*(M)$ en posant pour $w \in T_x^*(M)$

(1) $Q_x(w) = \Delta (f^2)(x) - 2 f(x) \Delta f(x)$

où f est $C^2$ au voisinage de x et $d f_x = w$.

L'opérateur $\Delta$ détermine complètement le champ de formes quadratiques $Q_x$, $x \in M$. La forme $Q_x$ lue dans une carte locale quelconque $\varphi$ n'est autre que $a_\varphi(y)$ avec $y = \varphi(x)$ ; en particulier le champ de tenseurs $Q_x$ est au moins de classe 1.
(La définition intrinsèque ci-dessus est extraite de [45] ).

Définissons le champ de formes quadratiques duales $Q_x^* : T_x(M) \to [0, +\infty]$ par

(2) $$\frac{1}{2} Q_x^*(v) = \operatorname*{Sup}_{w \in T_x^*(M)} \left[ < v, w > - \frac{1}{2} Q_x(w) \right] \quad , \quad v \in T_x(M)$$

Lues dans les cartes locales, ces définitions coïncident exactement avec celles adoptées sur les ouverts de $\mathbb{R}^n$ au chapitre III.

Du résultat local III.2.9.(2) on déduit directement que

(3) l'application $Q^* : T(M) \to [0, +\infty]$ définie par $Q^*(x,v) = Q_x^*(v)$ pour $(x, v) \in T(M)$, est s.c.i. sur le fibré tangent $T(M)$.

Notons $\mathcal{E}_{S,T}(M)$ l'espace des trajectoires "explosives" à valeur dans $M \cup \delta$, définies sur l'intervalle $[S, T]$ (cf. III.1.2, IV.1.1) et $\tau(g)$ le temps d'explosion de $g \in \mathcal{E}_{S,T}(M)$.

Soit $g \in \mathcal{E}_{S,T}(M)$. Si $g'_t$ existe (au sens de Lebesgue) pour presque tout $t \in [S, \tau(g) \wedge T]$ nous posons

(4) $$\lambda_{S,T}(g) = \frac{1}{2} \int_S^{T\wedge\tau(g)} Q^*_{g_t} [g'_t - b(g_t)] \, dt$$

Si $g_t$ n'est pas absolument continue en $t \in [S, \tau(g) \wedge T]$ , nous posons $\lambda_{S,T}(g) = +\infty$.

Nous appelons $\lambda_{S,T} : \mathcal{C}_{S,T}(M) \to [0, +\infty]$ la transformée de Cramer de $(D_\varepsilon)$ et nous définissons la fonctionnelle de Cramer associée $\Lambda_{S,T}$ par

$$\Lambda_{S,T}(A) = \inf_{g \in A} \lambda_{S,T}(g) \qquad \text{pour} \qquad A \subset \mathcal{C}_{S,T}(M).$$

## 3. QUELQUES PROPRIETES "TOPOLOGIQUES" DE LA TRANSFORMEE DE CRAMER

Dans tout ce paragraphe, on se place dans la situation 2.1.(1) ou 2.1.(2).

3.1. Définition : appelons provisoirement "bonne" carte locale toute carte locale $\varphi : U \to \mathbb{R}^n$ telle que

- $\overline{U}$ est compact et $\varphi$ est la restriction à U d'une carte locale $\psi : V \to \mathbb{R}^n$ où $\overline{U} \subset V$
- $a_\varphi = \sigma_\varphi \sigma_\varphi^*$ avec $\sigma_\varphi$ champ $C^1$ de matrices rectangulaires.

Bien entendu, les "bonnes" cartes locales forment un atlas de M.

3.2. Lemme : pour toute bonne carte locale $\varphi : U \to \mathbb{R}^n$ , et tout compact K de U il existe $\alpha > 0$ , $\theta > 0$ tels que les relations $t, S, T \in \mathbb{R}$, $S \leqslant t \leqslant T$ , $g \in \mathcal{C}_{S,T}(M)$, $\lambda_{S,T}(g) \leqslant \alpha$ , $g_t \in K$ entraînent $g_s \in U$ pour $t \leqslant s \leqslant (t+\theta) \wedge T$.

Preuve : le champ b est continu, donc borné "en norme" sur le compact $\overline{U}$. La distance entre K et $\partial U$ étant minorée, il existe $\theta > 0$ tel que pour $x \in K$ , l'unique solution maximale y de $y'_s = b(y_s)$ telle que $y_o = x$ vérifie $y[0, \theta] \subset U$ . L'ensemble $\Gamma$ des $h \in \mathcal{C}_{0,\theta}(U)$ telles que $h_o \in K$ et $\tau(h) \leqslant \theta$ est fermé.

Par lecture dans la carte $\varphi$ et application de III.2.10, $\lambda_{0,\theta}$ atteint son minimum $\alpha$ sur $\Gamma$ en un point $f \in \Gamma$. Le nombre $\alpha = \lambda_{0,\theta}(f)$ est non nul, grâce au choix de $\theta$. Soit maintenant $g \in \mathcal{C}_{0,\theta}(M)$ vérifiant $g_o \in K$ et $\lambda_{0,\theta}(g) \leqslant \frac{\alpha}{2}$. Si $\tau_U(g) \leqslant \theta$, la fonction h qui coïncide avec g sur $[0, \tau_U(g)[$ et vaut $\delta$ sur $[\tau_U(g), \theta]$ est dans $\mathcal{C}_{0,\theta}(U)$ ; elle vérifie $\lambda_{0,\theta}(h) \leqslant \lambda_{0,\theta}(g) \leqslant \frac{\alpha}{2}$, ce qui entraîne $\tau(h) > \theta$ et contredit $\tau_U(g) \leqslant \theta$ puisque $\tau_U(g) = \tau(h)$. Par suite les conditions $g \in \mathcal{C}_{0,\theta}(M)$, $g_o \in K$, $\lambda_{0,\theta}(g) \leqslant \frac{\alpha}{2}$ entraînent $g[0, \theta] \subset U$.
Pour prouver le lemme, on peut toujours supposer $t = 0$ ; soient maintenant S, T tels que $S \leqslant 0 \leqslant T$. Posons $v = \theta \wedge T$.

Soit $g \in \mathcal{C}_{S,T}(M)$ telle que $g_o \in K$, $\lambda_{S,T}(g) \leqslant \frac{\alpha}{2}$. Posons $h = g$ sur $[0, v]$ ; sur $[v, \theta]$ définissons h comme l'unique solution de $y'_s = b(y_s)$ issue de $g_v$ au temps v, si $g_v \neq \delta$, et $h \equiv \delta$ si $g_v = \delta$.
On a alors $h \in \mathcal{C}_{0,\theta}(M)$

$$\lambda_{0,\theta}(h) = \lambda_{0,v}(g) \leqslant \lambda_{S,T}(g) \leqslant \frac{\alpha}{2}$$

et $h_o \in K$, d'où $g[0, v] \subset h[0, \theta] \subset U$, ce qui prouve le lemme.

3.3. <u>Proposition</u> : (Hypothèse 2.1.(1) ou bien 2.1.(2)). L'espace $\mathcal{C}_{S,T}(M)$ étant muni de la topologie III.1.2, l'application $\lambda_{S,T} : \mathcal{C}_{S,T}(M) \to [0, +\infty]$ est s.c.i. ; de plus, pour tout compact K de M, tout a fini, l'ensemble des $g \in \mathcal{C}_{S,T}(M)$ tels que $g_S \in K$, $\lambda_{S,T}(g) \leqslant a$ est compact.

<u>Preuve</u>

Soit $g^k \in \mathcal{C}_{S,T}(M)$ une suite telle que $g^k_o$ converge vers $x \in M$, et telle que $\lambda_{S,T}(g^k)$ converge vers a fini, lorsque $k \to +\infty$. Il nous suffit de prouver que pour toute telle suite il existe $g \in \mathcal{C}_{S,T}(M)$ vérifiant $\lambda_{S,T}(g) \leqslant a$ et une suite extraite de $g^k$ convergeant vers g dans $\mathcal{C}_{S,T}(M)$.

Soit $\varphi : U \to \mathbb{R}^n$ une bonne carte locale, avec $x \in U$ , et soit $V$ un ouvert tel que $x \in V$ , $\overline{V} \subset U$ . Le résultat local III.2.10, lu dans la carte $\varphi$ fournit $g \in \mathcal{C}_{S,T}(U)$ telle que les fonctions $f^k$ obtenues en "tuant" les $g^k$ au temps $\tau_U(g^k)$ convergent (après passage à une sous-suite) vers $g$ dans $\mathcal{C}_{S,T}(U)$. Par la s.c.i. locale de $\lambda_{S,T}$ ceci donne

$$\lambda_{S,T\wedge\tau_U(g)} = \lambda^U_{S,T}(g) \leqslant \underline{\lim}_{k\to\infty} \lambda^U_{S,T}(f^k) = \underline{\lim}_{k\to\infty} \lambda_{S,T\wedge\tau_U(g^k)}(g^k) \leqslant a$$

En particulier $g^k$ converge vers $g$ uniformément sur $[S, T \wedge \tau_V(g)]$ et $\lambda_{S,T\wedge\tau_V(g)}(g) \leqslant a$.

Soit $W$ un ouvert relativement compact de $M$ contenant $x$. On peut trouver un recouvrement ouvert fini $W_1, \dots, W_r$ de $\overline{W}$ , et des ouverts $V_i$ , $U_i$ tels que $\overline{W}_i \subset V_i$ , $\overline{V}_i \subset U_i$ où chaque $U_i$ est le domaine d'une bonne carte locale $\varphi_i$ .

Considérons l'ensemble $\mathcal{N}$ des entiers $L \geqslant 1$ ayant la propriété suivante : il existe une suite croissante stricte de temps $\tau_\ell \in [S, T]$ , $0 \leqslant \ell \leqslant L$ , avec $\tau_o = S$, une suite d'indices $i_\ell \in [1, r]$ , $0 \leqslant \ell \leqslant L-1$, et une suite extraite des $g^k$ convergeant uniformément sur $[S, \tau_L]$ vers une fonction continue $g : [S, \tau_L] \to M$ , telles que pour $0 \leqslant \ell \leqslant L-1$ on ait

(1) $g_{\tau_\ell} \in W_{i_\ell}$ ; $g_{\tau_{\ell+1}} \in \partial V_{i_\ell}$ ; $g\,[\tau_\ell , \tau_{\ell+1}[\, \subset V_{i_\ell}$ ;

$$\lambda_{\tau_\ell,\tau_{\ell+1}}(g) \leqslant \underline{\lim}_{k\to+\infty} \lambda_{\tau_\ell,\tau_{\ell+1}}(g^k)$$

L'ensemble $\mathcal{N}$ contient toujours $L = 1$ grâce à la construction ci-dessus. Le lemme 3.2 fournit $\alpha > 0$ , $\theta > 0$ tels que les conditions

(2) $S \leqslant s \leqslant t \leqslant T$ , $h \in \mathcal{C}_{S,T}(M)$ , $h_s \in W_i$ et $h_t \in \partial W_i$ pour un $i \in [1, r]$ , $\lambda_{S,T}(h) < \alpha$

entraînent $t - s \geq \theta$.

Soit $L \in \mathcal{L}$. De (1) on déduit, puisque $\sum \underline{\lim} \leq \underline{\lim} \sum$,

$$(3) \quad \sum_{\ell=0}^{L-1} \lambda_{\tau_\ell,\tau_{\ell+1}}(g) \leq \underline{\lim}_{k\to+\infty} \lambda_{S,\tau_L}(g^k) \leq a$$

Soit $E = \{ \ell \mid 0 \leq \ell \leq L-1 \, , \, \lambda_{\tau_\ell,\tau_{\ell+1}}(g) \geq \alpha \}$
Alors d'après (3) on a card $E \leq \frac{a}{\alpha}$. D'autre part si $\ell \in [0, L-1]$ et $\ell \notin E$, on a d'après (2), $\tau_{\ell+1} - \tau_\ell \geq \theta$ et par suite card $\{[0, L-1] - E\} \leq \frac{T-S}{\theta}$. Finalement on voit que $L \leq \frac{a}{\alpha} + \frac{T-S}{\theta}$ pour $L \in \mathcal{L}$. L'ensemble $\mathcal{L}$ est donc fini non vide. Soit L le maximum de $\mathcal{L}$. Alors $g_{\tau_L}$ ne peut pas pas appartenir à $\overline{W}$, sinon $g_{\tau_L}$ appartiendrait à l'un des $W_i$ et on pourrait itérer la construction indiquée plus haut avec $x \in V$, $\overline{V} \subset U$, pour prolonger $g$ sur $[\tau_L,\tau_{L+1}]$, ce qui contredirait la maximalité de L. Par suite il existe $L$, $g$, $(\tau_\ell)$, $(i_\ell)$ vérifiant (1) et la condition supplémentaire

$$(4) \quad T \wedge \tau_{\overline{W}}(g) \leq \tau_L < \tau(g).$$

Une suite extraite des $g^k$ converge alors uniformément vers $g$ sur $[S, T \wedge \tau_{\overline{W}}(g)]$. De plus (3) et (4) entraînent $\lambda_{S,T\wedge\tau_{\overline{W}}(g)}(g) \leq a$.
Prenons maintenant une suite croissante d'ouverts relativement compacts $W_p$ telle que $\bigcup_p W_p = M$. Appliquons le résultat précédent à chaque $W_p$ ; par le procédé diagonal on trouve alors une fonction $g \in \mathcal{E}_{S,T}(M)$, et une suite extraite des $g^k$ qui converge uniformément vers $g$ sur $[S, T \wedge \tau_{\overline{W}_p}(g)]$ pour chaque $p$, avec $\lambda_{S,T\wedge\tau_{\overline{W}_p}(g)}(g) \leq a$. Il est clair que $\lambda_{S,T}(g) = \lim_{p\to\infty} \lambda_{S,T\wedge\tau_{\overline{W}_p}}(g) \leq a$, et que $g^k$ tend vers $g$ dans $\mathcal{E}_{S,T}(M)$ (cf. III.1.2). Ceci prouve la proposition 3.3.

## 4. RECOLLEMENT DES ESTIMATIONS DE BASE

4.1. Les informations de type "local" sur $\lim\limits_{\varepsilon\to 0} \varepsilon^2 \log P(y^\varepsilon \in A)$ obtenues au chapitre III ne peuvent pas se recoller telles quelles sur une variété. Il nous faudra passer par l'intermédiaire des estimations uniformes de IV.1. En effet si on tronçonne un tube de chemins, les points de départ de chaque tronçon et leurs axes ne sont plus fixés mais varient dans des compacts. La méthode de recollement utilisée ici est proche de celle adoptée par Azencott-Ruget [3] dans un contexte voisin.

4.2. Il existe toujours sur M des métriques riemanniennes complètes (cf. Helgason [26] ). Fixons une telle métrique et notons d la distance associée. En particulier d (x, y) tend vers + ∞ lorsque y tend vers le point à l'infini $\delta$ de M . Posons $d(x, \delta) = +\infty$ pour $x \in M$ .

Pour $g$ , $h \in \mathcal{C}_{S,T}(M)$ et $h_{[S,T]} \subset M$ , posons

$$d_{S,T}(g, h) = \operatorname{Sup} \{d(g_t , h_t) \mid S \leqslant t \leqslant T\}$$

4.3. Lemme : [Hypothèses 2.1 (1), ou bien 2.1 (2)] ; Etant donnés des compacts K et L de M , et des nombres $T > 0$ , a fini , on peut trouver $\theta > 0$ et une famille finie de bonnes cartes locales $\varphi_i : U_i \to \mathbb{R}^n$ , $i = 1 \dots r$ de sorte que les conditions $g \in \mathcal{C}_{0,T}(M)$ , $g_0 \in K$ , $\lambda_{0,T}(g) \leqslant a$ entraînent que si $g_t \in L$ pour un $t \in [0, T]$, alors $g[t, T \wedge (t + \theta)]$ est inclus dans l'un des $U_i$ .

Preuve : Soit $\Gamma$ l'ensemble des g considérées dans le lemme. Soit $\xi_L(g) = \sup\{t \mid g_t \in L\}$ . On peut toujours trouver une famille finie d'ouverts $V_i$ recouvrant L et des ouverts tels que $\overline{V}_i \subset V'_i$ , $\overline{V}'_i \subset U'_i$ , $\overline{U}'_i \subset U_i$ , où les $U_i$ sont les domaines de bonnes cartes locales. Soit $g \in \Gamma$ . L'uniforme continuité de g sur $[0 , T \wedge \xi_L(g)]$ fournit $\theta_g > 0$ tel que si $g_t \in V'_i$ alors $g_{t+s} \in U'_i$ pour $0 \leqslant s \leqslant \theta_g$ ,

$t + s \leqslant T$ . Il existe alors un voisinage ouvert $A_g$ de $g$ dans $\mathcal{C}_{0,T}(M)$, tel que si $f \in A_g$ , la condition $f_t \in V_i$ implique $f_{t+s} \in U_i$ pour $0 \leqslant s \leqslant \theta_g$ , $t + s \leqslant T$ . Comme $\Gamma$ est compact, on le recouvre par un nombre fini de $A_g$ , ce qui prouve 4.3.

4.4. Proposition (Hypothèses 2.1 (1) ou 2.1 (2)

Sur la variété M, soit $y^\varepsilon$ la $D_\varepsilon$-diffusion, où $D_\varepsilon = b_\varepsilon + \varepsilon^2 \Delta$

Soit $\lambda_{S,T}$ la transformée de Cramer de $(D_\varepsilon)_{\varepsilon>0}$ . Soit d une distance riemannienne complète arbitraire sur M. Donnons-nous des compacts K , L de M et des nombres positifs stricts T , a , $\rho$ , $\eta$ . Alors il existe $\varepsilon_o > 0$ , $r > 0$ tels que les conditions

(1) $x \in K$ , $g \in \mathcal{C}_{0,T}(M)$ , $d(g_o, x) \leqslant r$ , $g[0T] \subset L$ ,

$\lambda_{0,T}(g) \leqslant a$ , $\varepsilon \leqslant \varepsilon_o$

entraînent

(2) $-\lambda_{0,T}(g) - \eta \leqslant \varepsilon^2 \log P_x \{d_{0,T}(y^\varepsilon, g) \leqslant \rho\}$

Preuve : Soit N un entier plus grand que 1 . Posons $t_k = k \frac{T}{N}$ $k = 1 \dots N$. D'après 4.3 on détermine une famille finie de bonnes cartes locales $\varphi_i : U_i \to \mathbb{R}^n$ $i = 1 \dots p$ et N assez grand pour que

(3) quel que soit $k = 1 \dots N$ , quel que soit g vérifiant $g[0\ T] \subset L$ et $\lambda_{0,T}(g) \leqslant a$ , il existe un $i \in [1, p]$ tel que $U_i$ contienne $g[t_{k-1}, t_k]$.

On peut toujours trouver des ouverts $V_i \supset \overline{U}_i$ tels que les $\varphi_i$ soient de bonnes cartes locales définies sur $V_i$ . Donnons-nous $\eta > 0$ et notons $\theta_t$ les opérateurs de translation dans le temps (cf. IV.2.7). La proposition IV.1.2 appliquée dans les ouverts $\varphi(V_i)$ fournit des fonctions $\rho \to \varepsilon_o(\rho) > 0$ et $\rho \to r(\rho) > 0$ tendant vers 0 avec $\rho$, telles que les conditions

$$(4) \quad \begin{cases} k \in [1, N] \ , \ \lambda_{0,T}(g) \leqslant a \ , \ g[0, T] \subset L \ , \ x \in M \ , \\ d(g_{t_k}, x) \leqslant r(\rho) \ , \ y_o^\varepsilon = x \ , \ \varepsilon \leqslant \varepsilon_o(\rho) \end{cases}$$

entraînent

$$-\lambda_{t_{k-1},t_k}(g) - \frac{\eta}{N} \leqslant \varepsilon^2 \log P_x \{d_{0,1/N}(y^\varepsilon, \theta_{t_k} g) \leqslant \rho\}.$$

Grâce à la propriété de Markov aux instants $t_k$ , on voit que les hypothèses

$$(5) \quad k \in [1, N] \ , \ \lambda_{0,T}(g) \leqslant a \ , \ g[0, T] \subset L \ , \ \varepsilon \leqslant \varepsilon_o(\rho) \ ,$$

entraînent l'inégalité $P_x$ - p.s. <u>sur l'ensemble</u> $\{\omega \mid d(y^\varepsilon_{t_{k-1}}, g_{t_{k-1}}) \leqslant r(\rho)\}$

$$-\lambda_{t_{k-1},t_k}(g) - \frac{\eta}{N} \leqslant \varepsilon^2 \log P [d_{t_{k-1},t_k}(y^\varepsilon, g) \leqslant \rho \mid \mathcal{F}_{t_{k-1}}]$$

On peut toujours supposer vu la forme de (4), que $r(\rho) \leqslant \rho$ .

Définissons successivement des fonctions $r_k(\rho) > 0$ par :

$$r_N(\rho) = \rho \ , \quad r_{k-1} = r \circ r_k \quad \text{pour} \quad k = 1 \dots N$$

On a alors $r_{k-1} \leqslant r_k$ pour $k = 1 \dots N$ . Posons :

$$\Omega_k = \{ \omega \mid d_{t_{k-1},t_k}(y^\varepsilon, g) \leqslant r_k(\rho) \} \qquad k = 1 \dots N .$$

Prenons g telle que $\lambda_{0,T}(g) \leqslant a$ , $g[0, T] \subset K$ , et soit $x \in M$ tel que $d(x, g_o) \leqslant r_o(\rho)$ . Posons $\varepsilon_1 = \inf_{1 \leqslant k \leqslant N} \varepsilon_o \circ r_k$ . Puisque l'évènement $\{d_{0,T}(y^\varepsilon, g) \leqslant \rho\}$ contient $\bigcap_{1 \leqslant k \leqslant N} \Omega_k$ , (5) montre que pour $\varepsilon \leqslant \varepsilon_1(\rho)$ on aura

$$\varepsilon^2 \log P\{d_{0,T}(y^\varepsilon, g) \leqslant \rho\} \geqslant \varepsilon^2 \log P(\bigcap_{1<k<N} \Omega_k) \geqslant \sum_{1 \leqslant k \leqslant N} [-\lambda_{t_{k-1},t_k}(g) - \frac{\eta}{N}]$$

$$= \lambda_{0,T}(g) - \eta$$

ce qui prouve le résultat.

4.5. Du résultat précédent on déduira facilement plus bas la minoration $-\Lambda(\mathring{A}) \leqslant \underline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A)$. La majoration par contre est plus délicate et nous obligera à introduire une hypothèse supplémentaire (constance du rang de la matrice $a_\varphi$).

4.6. <u>Proposition</u> : (Hypothèses 2.1 (1) ou bien 2.1 (2)).

Soit $\varphi : U \to \mathbb{R}^n$ une <u>bonne</u> carte locale. Supposons le rang de la forme quadratique $Q_x$ (définie par $\Delta$ sur $T_x^*(M)$) constant égal à $r$ pour tout $x \in U$. Ceci équivaut à dire que le rang de $a_\varphi$ (cf. 2.2) reste égal à $r$. Alors il existe une <u>base</u> d'ouverts $V$ de $U$ ayant la propriété suivante :

pour tout compact $L$ de $V$, pour tout $T$, $\rho$, $\eta$ positifs stricts, il existe $r > 0$ tel que dès qu'une fonction continue $g$ vérifie $g[0\ T] \subset L$, $\lambda_{0,T}(g)$ fini, dès que $x \in V$ vérifie $d(x, g_o) \leqslant r$, alors on peut construire une fonction continue $f : [0, T] \to V$ vérifiant $f_o = x$, $d_{0,T}(f, g) \leqslant \rho$, $(1-\eta)\ \lambda_{0,T}(g) \leqslant \lambda_{0,T}(f) \leqslant (1+\eta)\ \lambda_{0,T}(g)$

<u>Preuve</u> :

Il est clair qu'on peut supposer que $U$ est un ouvert de $\mathbb{R}^n$ ; écrivons alors, grâce aux hypothèses 2.1

$$\Delta f(x) = \frac{1}{2} \sum_{1\leqslant i,j\leqslant n} a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i\, \partial x_j} + \sum_{1\leqslant i\leqslant n} h_i(x) \frac{\partial f}{\partial x_i}(x)$$

où $a(x) = [a_{ij}(x)] = \sigma(x)\,\sigma^*(x)$, avec $\sigma(x)$ champ $C^1$ de matrices $(n, k)$. Notons $\sigma(x) = [\sigma_{ij}(x)]$ $\quad i = 1 \dots n \quad j = 1 \dots k$.

Le rang $r$ de $a(x)$ étant constant, $U$ est réunion d'un nombre fini d'ouverts $U_i$ tels que dans chaque $U_i$ un mineur fixe d'ordre $r$ de $\sigma(x)$ reste non nul. Il suffit évidemment de prouver le résultat annoncé dans le cas $U = U_1$, ce que nous supposons maintenant. Sans perdre de généralité on peut supposer alors que les colonnes $\sigma_{.1}(x) \dots \sigma_{.r}(x)$ de

$\sigma$ (x) restent indépendantes pour tout $x \in U$, et que les lignes $\sigma_{1.}(x) \ldots \sigma_{r.}(x)$ ont la même propriété. Les espaces $H_x = \sigma(x) \,.\, \mathbb{R}^k$ et $H_x^* = \sigma^*(x) \,.\, \mathbb{R}^n$ admettent alors resp. les bases $\sigma_{.1}(x) \ldots \sigma_{.r}(x)$ et $\sigma^*_{.1}(x) \ldots \sigma^*_{.r}(x)$. Pour $v \in H_x$, on a $Q_x^*(v) = ||w||^2$ où $w \in \mathbb{R}^k$ est de norme minimale parmi les vecteurs w tels que $\sigma(x)\, w = v$ (voir III.2.8 (5)). Comme le noyau de $\sigma$ (x) est l'orthogonal dans $\mathbb{R}^k$ de $H_x^*$, il est clair que w est l'unique vecteur de $H_x^*$ tel que $\sigma(x)\, w = v$.

Ecrivons

(1) $v = \sum\limits_{1 \leq j \leq r} R_j \;\sigma_{.j}(x)$ pour $v \in H_x$, avec $R_j = R_j\,(x, v)$

$w = \sum\limits_{1 \leq j \leq r} S_j \;\sigma^*_{.j}(x)$ pour $w \in H_x^*$, avec $S_j = S_j\,(x, w)$

Soit $e_1 \ldots e_n$ la base usuelle de $\mathbb{R}^n$, d'où $\sigma^*_{.j}(x) = \sigma^*(x)\, e_j$ et $\sigma(x)\, w = \sum\limits_{1 \leq j \leq r} S_j \; a_{.j}(x)$, où les $a_{.j}$ sont les colonnes de a.

L'équation $\sigma(x)\, w = v$ s'écrit pour $w \in H_x^*$ $v \in H_x$

(2) $\sum\limits_{1 \leq j \leq r} S_j \; a_{.j}(x) = \sum\limits_{1 \leq j \leq r} R_j \;\sigma_{.j}(x)$

et admet donc une unique solution $S_1 \ldots S_r$ pour chaque $R_1 \ldots R_r$.

Par suite, un des mineurs d'ordre r de $[a_{.1}(x) \ldots a_{.r}(x)]$ est non nul en x, et par continuité dans un voisinage $V_x$ de x. Soit m (x) la matrice d'un tel mineur ; puisque a(x) est $C^1$, $m(x)^{-1}$ est alors $C^1$ au voisinage de x.

Soit $\pi_j$ $j = 1 \ldots r$ la base usuelle de $\mathbb{R}^r$.

Posons $S\,(x, w) = \sum\limits_{1 \leq j \leq r} S_j\,(x, w)\; \pi_j$ , $R\,(x, v) = \sum\limits_{1 \leq j \leq r} R_j\,(x,v)\; \pi_j$

Alors (2) donne

(3) $S\,(x, w) = m\,(x)^{-1}\; R\,(x, v)$

tandis que, en notant $J_n = \mathbb{R}^r \to \mathbb{R}^n$ $J_k : \mathbb{R}^r \to \mathbb{R}^k$ les injections naturelles, on a

(4) $w = \sigma^* (x) \circ J_n \,.\, S (x, w)$ , $v = \sigma (x) \circ J_k \,.\, R (x, v)$

d'où finalement,

$$w = \sigma^* (x) \circ J_n \circ m(x)^{-1} \,.\, R (x, v) \quad \text{et}$$

(5) $Q_x^* (x) = ||w||^2 = [m(x)^{-1} \,.\, R (x,v)]^* \quad A(x) \; [m(x)^{-1} \; R (x,v)]$

où

(6) $A(x) = J_n^* \; a(x) \; J_n$ définit une forme quadratique non dégénérée sur $\mathbb{R}^r$, de classe $C^1$ en x.

Pour $x, y \in U$ et $v \in \mathbb{R}^n$ , définissons $\tilde{v} = F (x, y, v) \in \mathbb{R}^n$ par la formule

(7) $\tilde{v} = F (x, y, v) = \sigma (y) \circ J_k \,.\, R (x, v)$

Alors, $F (x, y, v) \in H_y$ et il est clair que

(8) $R (y, \tilde{v}) = R (x, v)$

Les formules (5) (8) donnent pour $y \in V_x$

$$Q_y^* (\tilde{v}) = [m(y)^{-1} \; R (x,v)]^* \quad A(y) \quad [m(y)^{-1} \; R (x,v)]$$

Puisque A(x) est à coefficients $C^1$ , et définie positive sur $\mathbb{R}^r$ on sait que $\sqrt{A(x)}$ est continue et inversible. Posons $B(y) = \sqrt{A(y)} \; m(y)^{-1}$ . Alors on a pour $y \in V_x$ .

(9) $Q_y^* (\tilde{v}) = ||B(y) \; R (x,v)||^2$ , $Q_x^* (v) = ||B(x) \,.\, R (x, v)||^2$

ce qui entraîne pour $y \in V_x$ ,

$$\frac{1}{||B(x) \; B(y)^{-1}||^2} \; Q_x^*(v) \leqslant Q_y^*(\tilde{v}) \leqslant \quad ||B(y) \; B(x)^{-1}||^2 \; Q_x^*(v)$$

Soit K un compact de $V_x$ ; l'inégalité précédente fournit une fonction monotone $d \to \eta_K (d) > 0$ tendant vers 0 avec d , telle que pour $y, z \in K$

on ait

(10) $\left[1 - \eta_K (|z-y|)\right] Q_z^* (v) \leqslant Q_y^* \left[F (z,y,v)\right] \leqslant \left[1 + \eta_K (|z-y|)\right] Q_z^* (v)$

Pour $v \in H_x$ , on a d'après (1)

$$v_i = \sum_{1\leqslant j\leqslant r} \sigma_{ij} (x) R_j (x,v) \qquad i = 1 \dots n$$

ce qui entraîne

(11) $R (x, v) = \left[s (x)^{-1} \circ J_n^*\right] . v$

où $s(x) = \left[\sigma_{ij} (x)\right]$ $\quad 1\leqslant i, j\leqslant r$ est inversible, à coefficients $C^1$.
L'application $u \to R (x, \sigma(x) u)$ de $\mathbb{R}^k$ dans $\mathbb{R}^r$ s'écrit donc
$R (x, \sigma(x) u) = \left[s (x)^{-1} \circ J_n^* \circ \sigma(x)\right] . u$
Elle est linéaire, à coefficients $C^1$. D'après (7) on a donc, pour
$u \in \mathbb{R}^k$ , $y, z \in V_x$

(12) $F (z, y, \sigma(z) u) = \left[\sigma(y) \circ J_k \circ s(z)^{-1} \circ J_n^* \circ \sigma(z)\right] . u = G (z,y)$ .

où $G (z, y) : \mathbb{R}^k \to \mathbb{R}^n$ est une application linéaire, à coefficients $C^1$
en z, y. De plus (4) implique

(13) $G (z, z) u = \left[\sigma (z) \circ J_k\right] . R (z, \sigma(z) u) = \sigma(z) u$

Soit g une fonction telle que $g \left[0, T\right] \subset K$ et $\lambda_{0,T} (g) \leqslant a$.
On peut écrire (cf III.2.10)

(14) $g_t' = \sigma (g_t) \varphi_t' + b (g_t)$ $\qquad$ p.p. $t \in \left[0, T\right]$

où $\varphi : \left[0, T\right] \to \mathbb{R}^k$ vérifie

(15) $\frac{1}{2} \int_0^T |\varphi_t'|^2 dt = \lambda_{0,T} (g) \leqslant a.$

Considérons l'équation différentielle

$$(16)\quad f'_t = G\,(g_t,\, f_t)\ \varphi'_t + b\,(f_t) \qquad \text{p.p. } t \in [0,\, T]$$
$$f_o = z \quad , \quad \text{où} \quad z \in V_x \ .$$

Comme $\varphi' \in L_2\,[0,\, T]$ et comme $G\,(.\ ,\ .)$ est $C^1$ sur $V_x \times V_x$, il existe une unique solution maximale de (16) appartenant à $\mathcal{C}_{0,T}\,(V_x)$. Notons que si $z = g_o$, la solution maximale de (16) n'est autre que $g_t$ grâce à (13) et (14). Donc on peut s'attendre à ce que f et g soient proches si z est proche de $g_o$. En effet, soit W un voisinage compact de K dans $V_x$. En intégrant (14) et (16), et en utilisant (13), on a pour $t \leqslant \tau_W\,(f)$,

$$f_t - g_t = \int_0^t \left[G\,(g_s, f_s) - G\,(g_s, g_s)\right]\ \varphi'_s\, ds + \int_o^t \left[b\,(f_s) - b\,(g_s)\right] ds$$
$$+ f_o - g_o$$

les fonctions G et b étant $C^1$ au voisinage du compact W , on peut écrire pour $z\ ,\ y \in W$

$$|G\,(z,\, y) - G\,(z,\, z)| + |b(z) - b(y)| \leqslant \text{cte}\ |z - y|$$

d'où pour $t \leqslant \tau_W\,(f)$

$$|f_t - g_t| \leqslant |f_o - g_o| + \text{cte} \int_o^t |f_s - g_s|\, ds$$
$$+ \text{cte}\left[\int_o^t |\varphi'_s|^2\, ds\right]^{1/2} \left[\int_o^t |f_s - g_s|^2\, ds\right]^{1/2}$$

La fonction $\rho\,(t) = \sup\limits_{0 \leqslant s \leqslant t} |f_s - g_s|$ vérifie donc pour $t \leqslant \tau_W\,(f)$

$$(17)\quad \rho\,(t) \leqslant \text{cte} \int_o^t \rho\,(s)\, ds + \text{cte}\,\sqrt{a}\left[\int_o^t \rho\,(s)^2\, ds\right]^{1/2} + \rho\,(o)$$

On majore $\int_o^t \rho\,(s)^2\, ds$ par $t\,\rho\,(t)^2$ pour obtenir, lorsque $\text{cte}\,\sqrt{a}\ \sqrt{t} \leqslant \frac{1}{2}$, la majoration $\rho\,(t) \leqslant c \int_o^t \rho\,(s)\, ds + 2\,\rho\,(o)$ où c est une constante. Un argument classique permet d'en déduire l'inégalité $\rho\,(t) \leqslant 2\,\rho\,(o)\, e^{ct}$ pour $t \leqslant \tau_W\,(f)$. Il est alors clair que pour T, K et $V_x$ fixés, on peut pour tout $\rho > 0$ fixé trouver $r > 0$ tels que les conditions $\{|z - g_o| \leqslant r \, , \, g\,[0, T] \subset K \, , \; \lambda_{0,T}(g) \text{ fini}\}$ forcent la solution f de (16) à vérifier

(18) $d_{0,T}(f, g) \leqslant \rho$ et $f\,[0, T] \subset V_x$

De plus d'après (16) (12) on a

$$Q^*_{f_t}\,[f'_t - b\,(f_t)] = Q^*_{f_t}\,[G\,(g_t\,,\,f_t)\;\varphi'_t] = Q^*_{f_t}\,[F\,(g_t, f_t\,,\,\sigma(g_t)\;\varphi'_t]$$

puis en tenant compte de (10)

$$Q^*_{f_t}\,[f'_t - b\,(f_t)] \leqslant [1 + \eta_K\,(|f_t - g_t|)]\quad Q^*_{g_t}\,[\sigma\,(g_t)\;\varphi'_t]$$

Grâce à (18) et (14) ceci devient

$$Q^*_{f_t}\,[f'_t - b\,(f_t)] \leqslant [1 + \eta_K\,(\rho)]\quad Q^*_{g_t}\,[g'_t - b\,(g_t)]$$

La moitié gauche de (10) fournit la minoration analogue, d'où finalement par intégration

(19) $[1 - \eta_K\,(\rho)]\;\lambda_{0,T}(g) \leqslant \lambda_{0,T}(f) \leqslant [1 + \eta_K\,(\rho)]\;\lambda_{0,T}(g)$

Ceci prouve la proposition 2.7, puisque les $V_x$ recouvrent U.

4.7. Remarque : Etant donné l'importance de 4.6 dans la suite de l'argument de recollement, il serait pertinent d'examiner la validité de 4.6 lorsque le rang de $a_\varphi$ n'est pas constant sur U , ce que nous n'avons pas fait. D'autre part, le résultat 4.6 n'utilise en fait que l'existence d'une décomposition locale $a_\varphi = \sigma_\varphi\,\sigma^*_\varphi$ avec $\sigma_\varphi$ rectangulaire et $x \to \sigma_\varphi(x)$ lipschitzienne de rang constant.

## 4.8. Régularisation d'un chemin

Donnons nous $\rho$ et T positifs, et un compact $L \subset M$. Recouvrons le compact L de M par un nombre fini de boules V de centre $x_V \in L$ , de rayon $\rho$. Notons

$V^k$ la boule de centre $x_V$ de rayon $k\rho$. On peut toujours supposer $\rho$ assez petit pour que les $V^{10}$ soient les domaines de bonnes cartes locales. Soit V l'une quelconque de ces boules.

Donnons-nous $A \geqslant 0$. Choisissons $a > 0$ assez petit pour que les trajectoires du système $y'_t = b\,(y_t)$ issues de $\overline{V}^2$ n'atteignent pas $\partial V^3$ en un temps inférieur à a. Posons $F_a = \{g \in \mathcal{C}_{0,a}\,(M) \mid g_o \in \overline{V}\,,\ \lambda_{0,a}(g) \leqslant A\}$. Comme on l'a vu plus haut, il existe alors $\theta \in ]0,\, a]$ tel que $g \in F_a$ implique $\tau_{V^2}\,(g) \geqslant \theta$ . Soit alors $g \in F_\theta$. Si $\tau_{V^2}\,(g) < \theta$ prolongeons g sur $[\tau_{V^2}\,(g)\ ,\ a]$ par la solution de $y'_t = b\,(y_t)$ issue du point $g_{\tau_{V^2}(g)}$ au temps $\tau_{V^2}\,(g)$ , ce qui est possible grâce au choix de a ; on obtient ainsi une fonction f telle que

$$\lambda_{0,a}\,(f) = \lambda_{0,\tau_{V^2}(g)}\,(g) \leqslant \lambda_{0,\theta}\,(g) \leqslant A$$

ce qui prouve que $f \in F_a$. D'après la construction de $\theta$ on devrait avoir $\tau_{V^2}\,(f) \geqslant \theta$ , alors que par construction de f on a $\tau_{V^2}\,(f) = \tau_{V^2}(g) < \theta$.

Cette contradiction montre que $g \in F_\theta$ entraîne $\tau_{V^2}\,(g) \geqslant \theta$ , et donc $\tau_{V^3}\,(g) > \theta$ .

On peut donc appliquer l'estimation IV.1.4, lue dans une bonne carte locale de domaine $V^{10}$ pour obtenir une fonction $\rho \to \varepsilon_o\,(\rho) > 0$ telle que

$$P_x\,\{d_{0,\theta}\,(z^\varepsilon,g) \geqslant \rho \quad \text{pour tout } g \in \mathcal{C}_{0,\theta}(V^{10})$$
$$\text{tel que } g_o = x,\ \lambda_{0,\theta}\,(g) \leqslant A\} \leqslant \exp\,(-\,\frac{A-1}{\varepsilon^2}\,)$$

pour $x \in \overline{V}$ , $\varepsilon \leqslant \varepsilon_o\,(\rho)$ , $z^\varepsilon$ = processus induit par $y^\varepsilon$ sur $V^{10}$. Mais si il existe g telle que $d_{0,\theta}\,(z^\varepsilon,\, g) < \rho$ et $g \in \mathcal{C}_{0,\theta}\,(V^{10})$, $g_o = x$ , $\lambda_{0,\theta}\,(g) \leqslant A$ , alors on a $\tau_{V^2}\,(g) \geqslant \theta$ , et donc $\tau_{V^3}\,(y^\varepsilon) = \tau_{V^3}\,(z^\varepsilon) \geqslant \theta$

On obtient donc

(1) $P_x \left\{ \tau_{V^3} (y^\varepsilon) < \theta \right\} \leqslant \exp \left(- \frac{A-1}{\varepsilon^2}\right)$

pour $x \in \overline{V}$ , $\varepsilon \leqslant \varepsilon_o (\rho)$. Puisque les ouverts V considérés sont en nombre fini, les mêmes $\theta$ et $\varepsilon_o$ pourront être utilisées pour tous les V du recouvrement fini $\mathcal{R}$ choisi plus haut.

Nous allons construire par morceaux un chemin aléatoire proche de $y^\varepsilon$ sur $[0, \tau_L]$ pour lequel la valeur (aléatoire) de $\lambda_{0, \tau_L}$ ait une loi dont on puisse estimer la queue quand $\varepsilon \to 0$.

Etant donné $T > 0$ , on peut quitte à diminuer $\theta$ prendre $\theta = \frac{T}{N}$ avec N entier. Définissons des v.a. $\Lambda_{r,k}$ pour $r > 0$ , k entier, par

$$\Lambda (r, k) = \inf \left\{\lambda_{k\theta,(k+1)\theta} (g) \mid d_{k\theta,(k+1)\theta} (y^\varepsilon, g) \leqslant r\right\}$$

La mesurabilité de $\Lambda (r, k)$ en $\omega \in \Omega$ est prouvée plus bas en 4.10. Définissons une v.a. K par

$$K = \operatorname{Sup} \left\{k \mid k \text{ entier} ,\ k\,\theta \leqslant \tau_L (y^\varepsilon) ,\ k \leqslant N\right\}$$

Soit U une application mesurable de L dans l'ensemble fini $\mathcal{R}$ , telle que $x \in U(x)$ pour tout $x \in L$ . Pour $j \leqslant K$ , nous poserons $V_j = U (y^\varepsilon_{j\theta})$ , de sorte que $V_j$ est un ouvert (aléatoire) du recouvrement $\mathcal{R}$ , tel que $y^\varepsilon_{j\theta} \in V_j$ . Donnons-nous des $r_j > 0$ à préciser plus bas. Grâce au lemme 4.10 ci-dessous, on peut trouver des variables aléatoires $g^j$ , $j = 0, \ldots, N$ , à valeurs dans $\mathcal{E}_{j\theta,(j+1)\theta} (M)$, avec $g^j$ mesurable par rapport à $\mathcal{F}_{(j+1)\theta}$ , telle que

(2) $d_{j\theta,(j+1)\theta} (g^j, y^\varepsilon) \leqslant r_j$ et $\Lambda (r_j, j) = \lambda_{j\theta,(j+1)\theta} (g^j)$, $j = 1 \ldots N$

Plaçons-nous désormais <u>sous l'hypothèse de rang constant</u> introduite dans la proposition 4.6 ; <u>nous supposons donc que la forme quadratique $Q_x$ définie par $\Delta$ sur $T^*_x (M)$ reste de rang constant quand x varie dans M.</u> On peut alors grâce à la proposition 4.6, supposer que les $V^{10}$ sont les domaines de "très bonnes" cartes locales, c'est-à-dire que pour tout $\eta > 0$ , une fonction $u \to \varphi(u) > 0$ définie sur $]0, \rho[$ telle que les conditions

(3) $f \in \mathscr{C}_{0,\theta}(V^{10})$, $f\,[0, \theta] \subset V^9$ , $z \in M$ , $d\,(f_o, z) \leqslant \varphi(u)$

entraînent

(4) il existe $\tilde{f} \in \mathscr{C}_{0,\theta}\,(V^{10})$ telle que $d_{0,\theta}\,(\tilde{f}, f) \leqslant u$

et $\lambda_{0,\theta}\,(\tilde{f}) \leqslant (1+\eta)\ \lambda_{0,\theta}\,(f)$

La proposition 4.6 montre que en fait on peut choisir $\tilde{f} = H\,(z, f)$ où H est mesurable sur $M \times \mathscr{C}_{0,\theta}\,(V^{10})$.

Soient $B_j$ des nombres positifs à préciser plus bas. Supposons construite une variable aléatoire $G \in \mathscr{C}_{0,j\theta}\,(M)$, mesurable par rapport à $\mathscr{F}_{j\theta}$ , telle que sur la partie $\Omega_j$ de $\Omega$ définie par $\Omega_j = \{j \leqslant K$ ; et $\tau_{V^3_m}\,(y^\varepsilon) \geqslant \theta$ pour $m = 0, 1 \ldots (j-1)\}$ on ait

(5) $d_{0,j\theta}\ (G, y^\varepsilon) \leqslant B_j$ (sur $\Omega_j$)

Ceci entraîne

(6) $d\,(G_{j\theta}\ , g^j_{j\theta}) \leqslant B_j + r_j$

On peut toujours prendre $\varphi(u)$ inversible dans (3), et $\varphi(u) < u$. Nous noterons $\psi = \varphi^{-1}$ . Plaçons-nous sur la partie $\Omega_{j+1}$ de $\Omega$ , définie comme en (5). Alors si $r_j < \rho$ , (2) montre que $\tau_{V^4_j}\ (g^{j+1}) \geqslant \theta$ . On note f la translatée (temporelle) de $g^{j+1}$ par $(-\,j\theta)$ et on pose (cf. (3) (4)), $\tilde{f} = H\,(G_{j\theta}, f)$, ce qui est possible d'après (3) (6) si $B_j + r_j < \varphi(\rho)$.

Le chemin $F : [0, (j+1)\,\theta] \to M$ qui coïncide avec G sur $[0, j\theta]$ et avec la translatée dans le temps de $\tilde{f}$ sur $[j\theta, (j+1)\,\theta]$ , définit une v.a. à valeurs dans $\mathscr{C}_{0,(j+1)\theta}(M)$, sur l'évènement $\Omega_{j+1}$ . Sur $\Omega - \Omega_{j+1}$ on prend F = chemin fixe arbitraire. On a, grâce à (3) (4) (5) (2), sur $\Omega_{j+1}$

$$(7) \begin{cases} d_{0,j\theta}(F, y^{\varepsilon}) \leqslant \psi(B_j + r_j) = B_{j+1} \\ \lambda_{j\theta,(j+1)\theta}(F) = \lambda_{0,\theta}(\tilde{f}) \leqslant (1+\eta)\,\lambda_{j\theta,(j+1)\theta}(g^{j+1}) = \Lambda(r_j, j) \end{cases}$$

Définissons maintenant les $r_j$ et les $B_j$ par

$$(8) \quad r_N = B_N = \rho \ , \ r_{j-1} = B_{j-1} = \frac{1}{2}\,\varphi(r_j) \qquad j = N, \ldots, 1.$$

La méthode indiquée ci-dessus fournit clairement, par récurrence, une variable aléatoire G , à valeurs dans $\mathscr{C}_{0,T}(M)$, telle que $G|_{[0,\, j\theta]}$ soit $\mathscr{F}_{j\theta}$-mesurable pour j = 0 ... N , et telle que pour j = 0, 1, ..., N-1 on ait

$$(9) \quad \lambda_{j\theta,(j+1)\theta}(G) \leqslant (1+\eta)\,\Lambda(r_j, j) \leqslant (1+\eta)\,\Lambda(r_o, j) \qquad \text{sur } \Omega_{j+1}$$

$$(10) \quad d_{0,T\wedge\tau_L(y^{\varepsilon})}(G, y^{\varepsilon}) \leqslant \rho \qquad \text{sur } \Omega_{j+1} \cap (K \geqslant j)$$

Nous pouvons maintenant aborder la preuve des majorations uniformes analogues à IV.1.4.

4.9. <u>Proposition</u> : Les hypothèses sont celles de 2.1.(1), ou bien celles de 2.1.(2) ; nous supposons de plus que le rang de la forme quadratique $Q_x$ définie par $\Delta$ sur $T_x^*(M)$ reste constant lorsque x varie dans M (cf. 2.2.(1)). Alors pour tout compact L de M , pour tous T, A, $\rho$, $\eta$ positifs stricts, il existe $\varepsilon_o > 0$ tel que les relations $\varepsilon \leqslant \varepsilon_o$ , $a \leqslant A$ , $x \in L$ entraînent

$$P_x \{ d_{0,T\wedge\tau_L(y^{\varepsilon})}(g, y^{\varepsilon}) > \rho \text{ pour tout } g \in \mathscr{C}_{0,T}(M) \text{ tel que } g_o = x \text{ et }$$
$$\lambda_{0,T\wedge\tau_L(g)}(g) \leqslant a \} \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{\eta}{\varepsilon^2}\right)$$

Preuve : Soit G : $\Omega \to \mathcal{E}_{0,T}$ (M) le chemin aléatoire construit en 4.8. Reprenons les notations de 4.8, et posons $Y_j = \Lambda\,(r_o, j)$. Sur l'ensemble $(K = k) \cap \Omega_k$ , on a d'après 4.8 (9)

(1) $\lambda_{0,T\wedge\tau_L(y^\varepsilon)}\,(G) \leqslant (1+\eta)\,(Y_o + \ldots + Y_k)$

D'autre part, la majoration triviale

(2) $P_x\left[(K \geqslant k) \cap \Omega^c_{k+1}\right] \leqslant \sum\limits_{0\leqslant j\leqslant k} P_x\left[y^\varepsilon_{j\theta} \in L\ ,\ \tau_{V^3_j}\,(y^\varepsilon) < \theta\right]$

donne, grâce à 4.8.(1) (en conditionnant par rapport à $\mathcal{F}_{j\theta}$)

(3) $P_x\left[(K \geqslant k) \cap \Omega^c_{k+1}\right] \leqslant (1+k)\exp\left(-\frac{A-1}{\varepsilon^2}\right)$ pour $\varepsilon \leqslant \varepsilon_o$ , $x \in L$

De (1) et (3) on déduit pour $x \in L$ , $\varepsilon \leqslant \varepsilon_o$ , $a > 0$

(4) $$P_x\left[\lambda_{0,T\wedge\tau_L(y^\varepsilon)}\,(G) > a\right] \leqslant \sum_{k=0}^{N-1} P_x\left[\Omega_{k+1} \cap (K\geqslant k) \cap (Y_o+\ldots+Y_k > \frac{a}{1+\eta})\right]$$

$$+\ (N+1)^2 \exp\left(-\frac{A-1}{\varepsilon^2}\right)$$

L'évènement $E = \{y^\varepsilon_o \in L\ ,\ \tau_{V^3_o}\,(y^\varepsilon) \geqslant \theta\ ,\ Y_o > a\}$

implique $d_{0,\theta}\,(y^\varepsilon, g) > r_o$ pour tout $g \in \mathcal{E}_{0,\theta}$ (M) tel que $g_o \in \overline{V}_o$ et $\lambda_{0,\theta}\,(g) \leqslant a$ . D'après le début de 4.8, de telles fonctions g vérifient $\tau_{V^3_o}\,(g) > \theta$ . A fortiori, E est donc inclus dans l'évènement $D = \{y^\varepsilon_o \in L\ ;$ $\tau_{V^3_o}\,(y^\varepsilon) \geqslant \theta\ ;\ d_{0,\theta}\,(y^\varepsilon, g) > r_o$ pour tout $g \in \mathcal{E}_{0,\theta}\,(V^4_o)$ tel que $g_o = y^\varepsilon_o\ ,\ \lambda_{0,\theta}\,(g) \leqslant a\}$ . On peut "lire" $P_x$ (D) dans une carte locale de domaine $V^{10}$ où V = U (x) (cf. 4.8), et appliquer la majoration locale IV.1.4 à $P_x$ (D). Pour tout $\gamma > 0$ , on obtient ainsi $\varepsilon_1 > 0$ tel que pour $\varepsilon \leqslant \varepsilon_1$ , $a \leqslant A$ , $x \in L$ , on ait $P_x\,(E) \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{\gamma}{\varepsilon^2}\right)$.

Conditionnons par rapport à $\mathcal{F}_{k\theta}$ pour déduire de ce résultat l'inégalité presque sûre

$$(5) \quad P\left[Y_k > a \mid \mathcal{F}_{k\,\theta}\right] \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{\gamma}{\varepsilon^2}\right) \quad \text{sur l'ensemble}$$

$$\{k \leqslant K \ , \ \tau_{V_k^3}(y^\varepsilon) \geqslant \theta\}$$

Posons $Z_k = Y_o + \ldots + Y_k$ . Nous allons prouver, par récurrence sur $k$ , que pour $a \leqslant A$ , $x \in L$ , $\varepsilon$ "petit" on a

$$(6) \quad P_x\left[\Omega_{k+1} \cap (k \leqslant K) \cap (Z_k > a)\right] \leqslant \exp\left(-\frac{a}{\varepsilon^2} + \frac{2\,k\,\gamma}{\varepsilon^2}\right)$$

Pour $k = 0$ , ceci est une conséquence de (5). Supposons (6) vraie pour un $k \in [0, N-1]$. Reprenons un argument utilisé par Azencott-Ruget [3] . On écrit $Z_{k+1} = Z_k + Y_{k+1}$ , d'où

$$(7) \quad P_x\left[\Omega_{k+2} \cap (k+1 \leqslant K) \cap (Z_{k+1} > a)\right] \leqslant P_x\left[(Z_k > a-\gamma) \cap \Omega_{k+1} \cap (k \leqslant K)\right]$$
$$+ E\left[I_{(Z_k \leqslant a-\gamma)} \; I_{\Omega_{k+2}} \; I_{(k+1 \leqslant K)} \; P\left(Y_{k+1} > a - Z_k \mid \mathcal{F}_{k\theta}\right)\right]$$

Notons $\pi$ la mesure sur $\mathbb{R}^+$ définie par

$$(8) \quad \pi\ [0, a] = P_x\left[(Z_k < a) \cap \Omega_{k+2} \cap (k+1 \leqslant K)\right]$$

Le second membre de (7) est grâce à (5) (6) (8) majoré par

$$(9) \quad \exp\left(-\frac{a - 2\,k\,\gamma - \gamma}{\varepsilon^2}\right) + \int_o^{a-\gamma} d\pi\,(u) \;\; \exp\left(-\frac{a - u - \gamma}{\varepsilon^2}\right)$$

pour $\varepsilon \leqslant \varepsilon_1$ , $a \leqslant A$.

Ecrivons $e^{u/\varepsilon^2} = \frac{1}{\varepsilon^2} \int_{-\infty}^{u} e^{v/\varepsilon^2}\, dv$ et utilisons le théorème de Fubini pour obtenir

$$\int_o^{a-\gamma} e^{u/\varepsilon^2}\, d\pi(u) = \frac{1}{\varepsilon^2} \int_{-\infty}^{a-\gamma} e^{v/\varepsilon^2} \;\; \pi\left[v \vee 0,\ a - \gamma\right]\, dv$$

Majorons $\pi\left[v \vee 0\ ,\ a - \gamma\right]$ par $\exp\left[-\frac{1}{\varepsilon^2}(v - \gamma)\right]$ , grâce à (6), lorsque $2\,k\,\gamma < v < a - \gamma$ , et par 1 ailleurs ; ceci donne pour $a \geqslant 2\,k\,\gamma + \gamma$

$$\int_o^{a-\gamma} e^{u/\varepsilon^2} \, d\pi\,(u) \leqslant e^{2k\gamma/\varepsilon^2} \quad (1 + \frac{a - (2k+1)\,\gamma}{\varepsilon^2})$$

Reportons dans (9) pour conclure que le second membre de (7) est majoré par $(2 + \frac{A}{\varepsilon^2})\exp - \frac{1}{\varepsilon^2}\,[a - (2k+1)\,\gamma]$ , et donc par $\exp\,[- \frac{a - 2\,(k+1)\,\gamma}{\varepsilon^2}]$ si $(2 + \frac{A}{\varepsilon^2})\exp\,(- \frac{\gamma}{\varepsilon^2}) \leqslant 1$, ce qui est vrai dès que $\varepsilon \leqslant \varepsilon_2$ . Finalement (6) est vraie pour $k = 0 \ldots N$ , $\varepsilon \leqslant \varepsilon_3$ , $x \in L$ , $a \leqslant A$.

On utilise (4) et (6) pour conclure que

$$P_x\left[\lambda_{0,T\wedge\tau_L(y^\varepsilon)}\,(G) > a\right] \leqslant (N+1)\exp\,(- \frac{a}{(1+\eta)\varepsilon^2} + \frac{2\,N\gamma}{\varepsilon^2}) + (N+1)^2 \exp - \frac{A-1}{\varepsilon^2}$$

et finalement pour $\varepsilon \leqslant \varepsilon_4$ , $a \leqslant A - 1$ , $x \in L$ , on obtient

$$(10)\quad P_x\left[\lambda_{0,T\wedge\tau_L(y^\varepsilon)}\,(G) > a\right] \leqslant \exp\,(- \frac{a}{\varepsilon^2} + \frac{\gamma_1}{\varepsilon^2})$$

où $\gamma_1 = (2\,N + 1)\,\gamma + 2\,\eta$ , à condition d'avoir pris $\eta$ assez petit pour que $\frac{1}{1+\eta} \leqslant 1 - 2\,\eta$ , et $\varepsilon_4$ assez petit pour que $(N+1)^2 \exp\,(- \frac{\gamma}{\varepsilon^2}) \leqslant \frac{1}{2}$ lorsque $\varepsilon \leqslant \varepsilon_4$ .

D'autre part 4.8.(10) donne

$$P_x\left[d_{0,T\wedge\tau_L(y^\varepsilon)}\,(G,\, y^\varepsilon) > \rho\right] \leqslant \sum_{j=0}^{N-1} P_x\left[(K \geqslant j) \cap \Omega_{j+1}^c\right]$$

D'après (3) on voit que pour $\varepsilon \leqslant \varepsilon_4$ , $a \leqslant A-1$, $x \in L$ on a

$$(11)\quad P_x\left[d_{0,T\wedge\tau_L(y^\varepsilon)}\,(G,\, y^\varepsilon) > \rho\right] \leqslant (N + 1)^2 \exp\,(- \frac{A-1}{\varepsilon^2})$$

L'évènement $\{d_{0,T\wedge\tau_L(y^\varepsilon)}\,(g,\, y^\varepsilon) > \rho$ pour tout $g \in \mathcal{E}_{0,T}\,(M)$ tel que $g_o = x$ , $\lambda_{0,T\wedge\tau_L(g)} \leqslant a\}$ est inclus dans l'union de $\{d_{0,T\wedge\tau_L(y^\varepsilon)}\,(G,\, y^\varepsilon) > \rho\}$ et de $\{\lambda_{0,T\wedge\tau_L(y^\varepsilon)}\,(G) > a\}$ . Sa probabilité est donc majorée par $\left[\exp\left(-\frac{a}{\varepsilon^2}+\frac{\gamma_1}{\varepsilon^2}\right)+(N+1)^2\exp-\frac{A-1}{\varepsilon^2}\right]$,

d'après (10) et (11), donc par $\exp\left(-\frac{a}{\varepsilon^2}+\frac{\gamma_2}{\varepsilon^2}\right)$ pour $x \in L$ , $a \leqslant A-1$,

$\varepsilon \leqslant \varepsilon_5$ , avec $\gamma_2 = (2N+2)\gamma + 2\eta$ . A un changement de notations près, ceci prouve la proposition 4.9.

Il nous reste à prouver un lemme technique de mesurabilité utilisé dans la construction 4.8.

4.10. Lemme technique (Hypothèses 2.1.(1), ou bien 2.1.(2))

Pour S, T, r > 0 donnés la fonction $F : \Omega \to \mathbb{R}$ définie par $F(\omega) = \inf \{ \lambda_{S,T}(g) \mid d_{S,T}[y^\varepsilon(\omega), g] \leqslant r \}$ est mesurable.

Il existe une variable aléatoire Z à valeurs dans $\mathcal{E}_{S,T}(M)$ mesurable par rapport à la σ-algèbre engendrée par les $y_t^\varepsilon$ , $S \leqslant t \leqslant T$ , telle que $F = \lambda_{S,T}(Z)$.

Preuve. Pour $h \in \mathcal{E}_{S,T}(M)$, soit $G(h)$ l'inf. de $\lambda_{S,T}$ sur le fermé $A = \{g \in \mathcal{E}_{S,T}(M) \mid d_{S,T}(h, g) \leqslant r\}$ .

Soit $h_n \in \mathcal{E}_{S,T}(M)$ une suite telle que $\lim\limits_{n\to+\infty} G(h_n) = a$ existe et soit finie, et telle que $\lim\limits_{n\to\infty} h_n = h$ . D'après la proposition 3.3 , il existe $g_n$ telle que $d_{S,T}(h_n, g_n) \leqslant r$ et $G(h_n) = \lambda_{S,T}(g_n)$.

Puisque $\lambda_{S,T}(g_n)$ reste bornée, il existe $g \in \mathcal{E}_{S,T}(M)$ et une suite extraite de $g_n$ , (encore notée $g_n$) telles que $g = \lim\limits_{n\to\infty} g_n$ ; on a alors $\lambda_{S,T}(g) \leqslant \lim\limits_{n\to\infty} \lambda_{S,T}(g_n) \leqslant a$ , et $d_{S,T}(h, g) \leqslant r$ . Par conséquent $G(h) \leqslant \lambda_{S,T}(g) \leqslant a$ . Ceci montre que G est s.c.i. sur $\mathcal{E}_{S,T}(M)$.

Par conséquent $F(\omega) = G \circ y^\varepsilon_{S,T}(\omega)$ est mesurable.

Pour chaque h telle que G (h) soit fini, l'ensemble

$C\ (h) = \{g \in \mathcal{E}_{S,T}\ (M) \mid \lambda_{S,T}\ (g) = G\ (h)\ ,\ d_{S,T}\ (g, h) \leq r\}$

est un compact non vide. Soit B un fermé , et soit N entier.

L'ensemble $E = \{h \mid G\ (h) \leq N\ ,\ C\ (h) \cap B = \emptyset\}$ est mesurable. En effet C (h) $\cap$ B est vide si et seulement si pour tout $f \in B$ on a $G\ (h) < \lambda_{S,T}\ (f)$ .

Mais B étant fermé on sait que $\Lambda_{S,T}\ (B) = \lambda_{S,T}\ (f_o)$ pour un $f_o \in S$ ; donc $E = \{h \mid G\ (h) \leq N\ ,\ G\ (h) < \Lambda_{S,T}\ (B)\}$ est mesurable.

Posons $\Gamma\ (h) = C\ (h)$ si G (h) est fini, et $\Gamma\ (h) = \{h\}$ si $G\ (h) = +\infty$. Alors pour tout h , $\Gamma(h)$ est un compact non vide, et pour tout fermé B, l'ensemble des h telles que $\Gamma(h) \cap B = \emptyset$ est mesurable. Il existe donc (cf. Rockafellar [41]) une application mesurable L de $\mathcal{E}_{S,T}\ (M)$ dans $\mathcal{E}_{S,T}\ (M)$ telle que $L\ (h) \in \Gamma\ (h)$ pour tout h . Par construction on a $\lambda_{S,T} \circ L = G$. La v.a. $Z = L \circ y^{\varepsilon}_{S,T}$ vérifie alors $F = \lambda_{S,T}\ (Z)$.

5. <u>Estimation de $\lim\limits_{\varepsilon\to 0} \varepsilon^2 \log P(y^\varepsilon \in A)$</u>

5.1 <u>Théorème</u> :

Soient M une variété différentiable, $\Delta$ un opérateur semi-elliptique du second ordre sur M, $(b_\varepsilon)_{\varepsilon > 0}$ une famille de champs sur M telle que $\lim\limits_{\varepsilon\to 0} b_\varepsilon = b$. Les hypothèses précises sur les opérateurs $D_\varepsilon = \varepsilon^2\Delta + b_\varepsilon$ sont <u>ou bien</u> celles du "cas hypoelliptique" 2.1 (2), <u>ou bien</u> celles du "cas elliptique" 2.1 (1). De plus nous supposons que la forme quadratique $Q_x$ naturellement définie par $\Delta$ sur $T_x^*(M)$ (cf. 2.2 (1)) reste de rang constant quand x décrit M. Soit $y^\varepsilon$ la $D_\varepsilon$-diffusion sur M. Soit $\lambda_{S,T} : \mathcal{E}_{S,T}(M) \to [0, +\infty]$ la transformée de Cramer de $(D_\varepsilon)_{\varepsilon > 0}$, définie en 2.2, et soit $\Lambda_{S,T}$ la fonctionnelle de Cramer associée. Alors pour tout $x \in M$ et toute partie A de $\mathcal{E}_x(M) = \{g \in \mathcal{E}_{S,T}(M) | g_o = x\}$ on a

$$-\Lambda_{S,T}(\overset{\circ}{A}) \leqslant \underline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \leqslant \overline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \leqslant -\Lambda_{S,T}(\overline{A})$$

où $\overset{\circ}{A}$, $\overline{A}$ sont respectivement l'intérieur et l'adhérence de A dans $\mathcal{E}_x(M)$.

<u>Preuve</u> : Prenons S = 0. Soit $g \in \overset{\circ}{A}$ avec $\lambda_{0,T}(g)$ fini. Puisque A est un voisinage de g, il existe $t \in [0,T]$, $t < \tau(g)$, et $\rho > 0$ tels que si u vérifie $t \leqslant u < T \wedge \tau(g)$ alors

$$B = \{ h \in \mathcal{E}_x(M) \mid d_{o,u}(h,g) \leqslant \rho\} \subset A.$$

Les définitions 2.2 montrent que $\lambda_{0,T}(g)$ est la limite de $\lambda_{0,s}(g)$ quand s croit vers $T \wedge \tau(g)$. Fixons u pour le moment. Puisque $g[0,u]$ est compact, la proposition 2.6 entraîne

$$\underline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \geqslant \underline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P_x(y^\varepsilon \in B) \geqslant -\lambda_{0,u}(g)$$

Faisons maintenant croitre u vers $T \wedge \tau(g)$ pour obtenir

$$\underline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \geqslant -\lambda_{0,T}(g)$$

On remplace maintenant le second membre par son sup. en $g \in \overset{\circ}{A}$, avec $\lambda_{0,T}(g)$ fini, ce qui donne

(1) $$\varliminf_{\varepsilon \to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \geq - \Lambda_{0,T}(\overset{\circ}{A})$$

La majoration de $\varlimsup \varepsilon^2 \log P_x(y^\varepsilon \in A)$ est plus délicate.

Supposons provisoirement que b et $\Delta$ soient tels que

(2) $\{\lambda_{0,T}(g)$ fini$\}$ implique $T < \tau(g)$.

Soit $A \subset \mathcal{E}_x(M)$. Fixons $a < \Lambda_{0,T}(\bar{A})$ ; $\bar{A}$ est alors disjoint de $\Gamma = \{g \mid g_o = x \, , \, \lambda_{0,T}(g) \leq a\}$. Puisque $\tau(g) > T$ pour tout $g \in \Gamma$ un argument déjà utilisé plus haut montre l'existence d'un compact F de M tel que $g \in \Gamma$ implique $g[0,T] \subset F$ (x est fixé dans ce passage). D'autre part $\bar{A}$ est fermé, et est disjoint du compact $\Gamma$. Il existe donc $\rho > 0$ tel que $h \in \bar{A}$ implique $\{d_{0,T}(h,g) > \rho$ pour tout $g \in \Gamma\}$.

Soit L un compact de M tel que $F \subset L$ et $d(F, \partial L) \geq 2\rho$. L'évènement $\Omega_L = \{d_{0,T \wedge \tau_L(y^\varepsilon)}(y^\varepsilon, g) \leq \rho$ pour au moins un $g \in \Gamma\}$ implique $\tau_L(y^\varepsilon) > T$, puisque $g[0,T] \subset F$ quand $g \in \Gamma$. A fortiori $\Omega_L$ entraîne $\{d_{0,T}(y^\varepsilon, g) \leq \rho$ pour au moins un $g \in \Gamma\}$, ce qui implique $y^\varepsilon \notin \bar{A}$. On en conclut que

$$P_x(y^\varepsilon \in A) \leq P_x\{d_{0,T \wedge \tau_L(y^\varepsilon)}(y^\varepsilon, g) > \rho \text{ pour tout } g \in \Gamma\}$$

et la proposition 4.9 montre alors que

$$\varlimsup_{\varepsilon \to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \leq -a$$

On peut alors faire croitre a vers $\Lambda_{0,T}(\bar{A})$ pour obtenir, <u>sous l'hypothèse provisoire (2)</u>

(3) $$\varlimsup_{\varepsilon \to 0} \varepsilon^2 \log P_x(y^\varepsilon \in A) \leq - \Lambda_{0,T}(\bar{A})$$

Revenons au cas général. Soit U un ouvert relativement compact de M, et soit $U_n$ une suite d'ouverts relativement compacts telle que $\bar{U}_n \subset U_{n+1}$, $U_o = U$, $M = \bigcup_n U_n$. Considérons le champ $\tilde{b} = fb$ où $f : M \to \mathbb{R}$ est une fonction lisse tendant vers 0 à l'infini, égale à 1 sur U.

Il existe clairement des nombres $a_n > 0$ tels que les conditions $\{f \leqslant a_n$ sur $(U_{n+1} - \overline{U}_n)\}$ $n = 0, 1, 2 \ldots$ garantissent qu'aucune trajectoire du champ $\tilde{b}$ ne passe de $U_n$ à $\partial U_{n+1}$ en un temps inférieur à 2.

Posons $\tilde{\Delta} = f_1 \Delta$ où $f_1$ est une fonction analogue à f, et positive stricte. Soit $\mu$ la transformée de Cramer associée aux opérateurs $\tilde{b} + \varepsilon^2 \Delta$, et soit $\lambda$ celle des opérateurs $\tilde{b} + \varepsilon^2 \tilde{\Delta}$.

Par un argument déjà utilisé plusieurs fois, il existe des constants $c_n > 0$ telles que pour tout chemin $g \in \mathcal{E}_{0,1}(M)$ vérifiant $g_o \in \partial U_n$, et $\tau_{U_{n+1}}(g) \leqslant 1$ on ait $\mu_{0,1}(g) \geqslant c_n$. D'autre part la définition de $Q_x^*$ (cf. 2.2) montre que si $R_x^*$ , $S_x^*$ soit les formes quadratiques duales associées à $\Delta$ et $\tilde{\Delta}$ respectivement, on a $S_x^* = \dfrac{1}{f_1(x)} R_x^*$. Soit alors $g \in \mathcal{E}_{0,1}(M)$ vérifiant $g_o \in \partial U_n$, $\tau_{U_{n+1}}(g) \leqslant 1$. Soit s le dernier passage de g sur $\partial U_n$ avant le temps $\tau = \tau_{U_{n+1}}(g)$ ; soit h le chemin obtenu en prolongeant g sur $[\tau, 1 + s]$ par une trajectoire du champ $\tilde{b}$. Alors $\mu_{s,\tau}(g) = \mu_{s,s+1}(h) \geqslant c_n$. Soit $d_n$ le maximum de f sur $\overline{U}_{n+1} - U_n$. Par définition de $\lambda$, $\mu$ on a

$$\lambda_{s,\tau}(g) = \frac{1}{2} \int_s^\tau S^*_{g(t)}\left[g_t' - \tilde{b}(g_t)\right] dt = \frac{1}{2} \int_s^\tau \frac{1}{f_1(g_t)} R^*_{g_t}(g_t' - \tilde{b}(g_t))\, dt$$

d'où

$$\lambda_{s,\tau}(g) \geqslant \frac{1}{d_n} \mu_{s,1}(g) \geqslant \frac{c_n}{d_n}$$

Par conséquent les conditions $\{f_1 \leqslant c_n$ sur $\overline{U}_{n+1} - U_n\}$ garantissent $\lambda_{s,\tau}(g) \geqslant 1$ pour tout g vérifiant $g_o \in \partial U_n$, $\tau_{U_{n+1}}(g) \leqslant 1$. Pour tout chemin g posons $t_n = \tau_{U_n}(g)$ et $s_n$ = dernier passage sur $\partial U_{n-1}$ avant le temps $t_n$. On vient de voir que l'on a pour chaque n ou bien $t_n - s_n > 1$, ou bien $\lambda_{s_n,t_n}(g) \geqslant 1$. Comme $s_n \leqslant t_n \leqslant s_{n+1}$ pour tout n, on en déduit que si $\lambda_{0,T}(g)$ est fini, l'ensemble des n tels que $\lambda_{s_n,t_n}(g) \geqslant 1$ et

$t_n \leqslant T$ est fini. D'autre part l'ensemble des n tel que $t_n - s_n > 1$ et $t_n \leqslant T$ est évidemment fini aussi. Finalement, l'ensemble de n tels que $t_n \leqslant T$ est fini, et il existe donc N tel que $\tau_{U_N}(g) > T$ ce qui force $\tau(g) > T$.

Ainsi les opérateurs $\tilde{D}_\varepsilon = \tilde{b} + \varepsilon^2\tilde{\Delta}$ on les propriétés suivants (pour le choix adéquat de f, $f_1$ indiqué plus haut) : ils coincident avec $D_\varepsilon$ sur U, et leur transformée de Cramer $\tilde{\lambda}$ est telle que si $\tilde{\lambda}_{0,T}(g)$ est fini, alors $\tau(g) > T$.

Notons $z^\varepsilon$ la $\tilde{D}_\varepsilon$-diffusion sur M, de sorte que les processus $y^\varepsilon$ et $z^\varepsilon$ induisent la même diffusion sur U. A tout fermé A de $\mathcal{E}_{0,T}(M)$ associons l'ensemble A(U) défini par

$$A(U) = \{g \in \mathcal{E}_{0,T}(M) \mid \text{il existe } f \in A \text{ telle que } f = g \text{ sur } [0,\tau_U(g)[\}$$

Soit $\phi$ : $\mathcal{E}_{0,T}(M) \to \mathcal{E}_{0,T}(U)$ l'application continue définie par $\phi(g)_t = g_t$ pour $t \in [0,\tau_U(g)[$. Il est clair que $A(U) = \phi^{-1}[\phi(A)]$. Comme $\mathcal{E}_{0,T}(M)$ et $\mathcal{E}_{0,T}(U)$ sont des espaces polonais, $\phi(A)$ est analytique (cf. Meyer [35]) et donc A(U) est analytique. Les v.a. $y^\varepsilon$, $z^\varepsilon$ sont définies sur l'espace de probabilité $\Omega = \mathcal{E}_{0,T}(M)$ , que l'on peut bien entendu supposer complet pour $P_x$. Par conséquent $\{y^\varepsilon \in A(U)\}$ et $\{z^\varepsilon \in A(U)\}$ sont deux parties mesurables de $\Omega$ ; ces deux évènements ont clairement même $P_x$-probabilité puisque les processus induits par $y^\varepsilon$ et $z^\varepsilon$ sur U sont équivalents. Par conséquent, en appliquant (3) au processus $z^\varepsilon$, qui vérifie l'hypothèse provisoire (2) par construction, on obtient

$$(4)\quad \overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P_x[y^\varepsilon \in A(U)] = \overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P_x[z^\varepsilon \in A(U)] \leqslant -\Lambda_{0,T}[\overline{A(U)}]$$

où $y^\varepsilon$ et $z^\varepsilon$ représentent les $[0,T]$-trajectoires des processus $y_t^\varepsilon$, $z_t^\varepsilon$ comme $A(U) \supset A$, (4) implique

(5) $$\overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P_x(y^\varepsilon \in A) \leq -\Lambda_{0,T}\left[\overline{A(U)}\right]$$

pour tout $x \in M$, $A$ borélien de $\mathcal{E}_{0,T}(M)$ et $U$ ouvert relativement compact de $M$.

Appliquons (5) à la suite croissante $U_n$ d'ouverts relativement compacts considérée plus haut. Pour chaque $n$ il existe $f^n \in \overline{A(U_n)}$ telle que

(6) $$\lambda_{0,T}(f^n) = \Lambda_{0,T}\left[\overline{A(U_n)}\right]$$

Puisque $A(U_n)$ décroit avec $n$, la suite $\lambda_{0,T}(f^n)$ est croissante d'après (6), et (5) entraîne

(7) $$\overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P_x(y^\varepsilon \in A) \leq -\lim_{n\to+\infty} \lambda_{0,T}(f^n)$$

Si la suite $\lambda_{0,T}(f^n)$ n'est pas bornée, l'inégalité (7) implique trivialement

$$\overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P_x(y^\varepsilon \in A) = -\infty \leq -\Lambda(\overline{A})$$

Supposons donc $\lambda_{0,T}(f^n)$ bornée. Il existe alors une suite extraite de $f^n$ (encore notée $f^n$) convergeant vers $f \in \mathcal{E}_{0,T}(M)$, ce qui implique

(8) $$\lambda_{0,T}(f) \leq \lim_{n\to\infty} \lambda_{0,T}(f^n)$$

Posons $t_r = \tau_{U_r}(f)$. Alors il existe $n(r) > r$ tel que $d_{0,t_r}(f,f^{n(r)}) \leq \frac{1}{r}$ et $d(U_r, \partial U_{n(r)}) > 2$. Puisque $f^n \in \overline{A(U_n)}$, il existe $g^r \in A\left[U_{n(r)}\right]$ telle que $d_{0,t_r}(f^{n(r)}, g^r) \leq \frac{1}{r}$. Par définition de $A(U)$ il existe $h^r \in A$ telle que $g^r$ et $h^r$ sortent de $U_{n(r)}$ au même instant $s_r$ et coincident sur $[0,s_r[$. Puisque $d_{0,t_r}(f,g^r) \leq \frac{2}{r}$) on a $t_r < s_r$ et par suite $d_{0,t_r}(f,h^r) \leq \frac{2}{r}$. Il est évident que $h^r$ tend vers $f$ dans $\mathcal{E}_{0,T}(M)$ (cf. III 1.2) ce qui entraîne $f \in \overline{A}$. Mais alors (8) implique

(9) $$\Lambda_{0,T}(\overline{A}) \leq \lim_{n\to+\infty} \lambda_{0,T}(f^n)$$

tandis que l'inclusion $A(U_n) \supset A$ donne l'inégalité

(10) $\lambda_{0,T}(f^n) = \Lambda_{0,T}(\overline{A(U_n)}) \leqslant \Lambda_{0,T}(\bar{A})$

Confrontant (9), (10) et (7) on conclut que

(11) $\overline{\lim_{\varepsilon\to 0}}\ \varepsilon^2 \log P_x(y^\varepsilon \in A) \leqslant - \Lambda_{0,T}(\bar{A})$

pour tout $x \in M$, tout borélien A de $\mathcal{E}_{0,T}(M)$. Les inégalités (1) et (11) prouvent le théorème.

5.2 Remarques sur les hypothèses du théorème 5.1 :

(1) Nous avons déjà vu au Ch.III que si M est un ouvert de $\mathbb{R}^n$ sur lequel la matrice a(x) des coefficients d'ordre 2 de $\Delta$ s'écrit $a(x) = \sigma(x)\sigma(x)^*$ avec $\sigma(x)$ rectangulaire et de classe $C^1$ en x, l'hypothèse de "rang constant" pour la matrice a(x) est superflue. Il est "plausible" qu'il en soit de même pour M quelconque, sous les hypothèses 2.1(2) ou 2.1(1) mais nous n'avons pas clarifié ce point.

(2) Lorsque les opérateurs $D_\varepsilon = \varepsilon^2\Delta + b_\varepsilon$ vérifient les hypothèses 2.1(1) [ou bien 2.1(2)] seulement sur un ouvert relativement compact $U_\varepsilon$ de M tel que $U_\varepsilon$ croisse vers M pour $\varepsilon$ tendant vers 0, la fin de la preuve du théorème 5.1 montre que les résultats 5.1 restent valides pourvu que l'on soit assuré de l'existence d'une $D_\varepsilon$-diffusion $y^\varepsilon$ sur M, ou plus généralement de l'existence d'un processus de Markov $y^\varepsilon$ induisant sur $U_\varepsilon$ une $D_\varepsilon$-diffusion (nécessairement unique). Cette dernière condition est toujours réalisée évidemment ; plus précisément si $V_\varepsilon$ est une suite d'ouverts relativement compacts croissant vers M lorsque $\varepsilon \to 0$, et telle que $\overline{V}_\varepsilon \subset U_\varepsilon$, il existe à équivalence près un unique processus de Markov fort $y^\varepsilon$ tel que $y_t^\varepsilon \equiv y^\varepsilon_{\tau_{V_\varepsilon}}$ pour $t \geqslant \tau_{V_\varepsilon}$, et coincidant sur $\overline{V}_\varepsilon$ avec le processus obtenu en arrêtant au temps $\tau_{V_\varepsilon}$

(cf. Dynkin [18] pour la terminologie) l'unique $D_\varepsilon$-diffusion définie par $D_\varepsilon$ sur $V_\varepsilon$. Alors les résultats 5.1 s'appliquent au processus $y^\varepsilon$.

Remarquons que <u>la situation (2) décrite ci-dessus est automatiquement vérifiée si</u>

(3) - M et les champs b, $b_\varepsilon$ sont $C^\infty$

- l'opérateur $b + \Delta$ se met <u>localement</u> sous la forme $\sum_{i=1}^{r} X_i^2 + Y$ où l'algèbre de Lie engendrée par les champs $C^\infty$ $X_1 \dots X_r$ Y est de dimension n = dim M en tout point de l'ouvert où $X_1 \dots X_r$ Y sont definis.

- dans toute carte locale, les coordonnées de $b_\varepsilon$ ainsi que leurs dérivées partielles d'ordre arbitraire convergent, uniformément sur tout compact de la carte, respectivement vers les coordonnées de b et les dérivées partielles correspondantes de ces coordonnées.

L'ordre "arbitraire" des dérivées à considérer peut évidemment être remplacé par "ordre $\leqslant$ p", où p se calcul localement à partir du nombre de crochets nécessaire pour engendrer $T_x(M)$ à partir de $X_1 \dots X_r$ Y.

## 6. <u>Diffusions en temps petit</u> :

6.1 Le comportement de la densité p(t,x,y) d'une diffusion lorsque le temps t tend vers 0 peut s'étudier par des calculs directs délicats (cf. Varadhan [50] Molchanov [37]) qui font intervenir une structure riemannienne liée au générateur différentiel de la diffusion, dans le cas elliptique. Dans le cas hypoelliptique, la situation est nettement plus compliquée, et nécessite des calculs encore plus pénibles, qui sont loin d'éclaircir la situation hypoelliptique générique (cf. Gaveau [24] [25]) . Ces problèmes sont liés aux résultats à la Ventsel-Freidlin sur les petites perturbations, comme l'ont déjà noté Varadhan et Gaveau.

Nous allons esquisser le rapport entre les deux types de problèmes.

6.2 Les Hypothèses : Soit M une variété différentiable connexe, de dimension n, de classe $\geqslant 2$. Soit $\Delta$ un opérateur différentiel d'ordre 2, semi-elliptique, défini sur M, tel que $\Delta 1 = 0$. Soit b un champ de vecteurs sur M. Nous allons considérer deux types d'hypothèses :

(1) "Cas elliptique"

- b est localement lipschitzien, et $\Delta$ est elliptique ;

- il existe un atlas de M tel que pour toute carte $\varphi$ de cet atlas l'image $\Delta_\varphi$ de $\Delta$ par $\varphi$ s'écrive

$$\Delta_\varphi = \frac{1}{2} \sum_{1\leqslant i,j\leqslant n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{1\leqslant i\leqslant n} h_i(x) \frac{\partial}{\partial x_i}$$

où les $h_i$ sont localement lipschitziens, et où $a(x) = [a_{ij}(x)]$ soit de la forme $a(x) = \sigma(x)\sigma(x)^*$, avec $\sigma(x)$ champ de matrices (rectangulaires) , de classe 1 en x.

(2) "Cas hypoelliptique"

- M et b sont de classe $C^\infty$

- L'opérateur $D = \Delta + b$ se met localement sous la forme $\sum_{i=1}^{r} X_i^2 + Y$ où l'algèbre de Lie engendrée par les champs $X_1 \dots X_r$ Y (qui sont de classe $C^\infty$) est de dimension n en tout point de l'ouvert où $X_1 \dots X_r$ Y sont définis. (voir 1.2 (2 bis) pour une expression locale explicite de la condition de Hormander ci-dessus).

- le rang de la "matrice des coefficients du second ordre" de $\Delta$, lu dans une carte locale quelconque, reste constant égal à r.

Dans les situations 6.2 (1) ou bien 6.2 (2), il existe (cf. §1) une unique D-diffusion sur M, à équivalence près, où $D = \Delta + b$, que nous noterons $y_t$.

6.3 La transformée de Cramer de $\Delta$ :

Les hypothèses sont celles de 6.2 (1) ou bien celles de 6.2 (2). Soit $Q_x$ la forme quadratique définie par $\Delta$ sur l'espace cotangent $T_x^*(M)$

(cf. 2.2 (1)). Soit $Q_x^*$ la forme quadratique duale de $Q_x$ définie sur $T_x(M)$ (et à valeurs dans $[0, +\infty]$) par la formule 2.2 (2). Soit $\mathcal{E}_{0,T}(M)$ l'espace des trajectoires "explosives" à valeurs dans $M \cup \delta$, définies sur $[0,T]$ (cf. III 1.2, IV 1.1), et soit $\tau(g)$ le temps d'explosion de $g \in \mathcal{E}_{0,T}(M)$.

Soit $g \in \mathcal{E}_{0,T}(M)$. Si $g_t'$ existe (au sens de Lebesque) pour presque tout $t \in [0,T \wedge \tau(g)]$, nous posons

(1) $$\lambda_{0,T}(g) = \frac{1}{2} \int_0^{T\wedge\tau(g)} Q_{g_t}^*(g_t')\, dt$$

Ceci a un sens car $Q_x^*(v)$ est s.c.i. en $(x,v)$ sur $T(M)$ (cf. 2.2 (3)). Si $g_t$ n'est pas absolument continue en $t \in [0,T \wedge \tau(g)]$ nous posons $\lambda_{0,T}(g) = +\infty$.

Appelons $\lambda_{0,T}$ <u>la transformée de Cramer de $\Delta$</u> et définissons <u>la fonctionnelle de Cramer</u> associée à $\Delta$ par

(2) $$\Lambda_{0,T}(A) = \inf_{g \in A} \lambda_{0,T}(g), \quad \text{où } A \subset \mathcal{E}_{0,T}(M).$$

Il est clair (cf. 2.2) que $\lambda_{0,T}$ n'est autre que la transformée de Cramer de n'importe quel "système perturbé" de la forme $\varepsilon^2\Delta + b_\varepsilon$ où $\varepsilon \to 0$ et $b_\varepsilon$ <u>tend vers zero</u> quand $\varepsilon \to 0$ ; <u>en particulier, et c'est ce qui servira ci-dessous $\lambda_{0,T}$ est aussi la transformée de Cramer au sens 2.2 du système perturbé $\varepsilon^2\Delta + \varepsilon^2 b$</u>.

<u>Examinons plus précisément le cas elliptique 6.2 (1)</u> : dans ce cas, en coordonnées locales, lues dans la carte $\varphi$ (notations 6.2.1), on a $Q_x = a(x)$ et $Q_x^* = a(x)^{-1}$. Par suite $Q_x^*$ définit une métrique riemannienne sur M. Rappelons (cf. Milnor[36]) la définition de <u>l'énergie</u> E(g) du chemin absolument continu $g : [0,T] \to M$

(3) $E(g) = T \int_0^T ||g_t'||_{g_t}^2\, dt$ , où $||\ ||_x$ est la métrique riemannienne en x. Il est immédiat que pour un tel chemin g on a

(4) $$\frac{E(g)}{2T} = \lambda_{0,T}(g)$$

En particulier, si la métrique riemannienne est complète, on sait que pour $x$, $y \in M$, il existe des chemins $g$ d'énergie minimale allant de $x$ à $y$ entre les temps $0$ et $T$, que l'énergie minimale $E(g)$ vaut $d^2(x,y)$, où $d(x,y)$ est la distance riemannienne de $x$ à $y$, et que de tels chemins sont des géodésignes brisées, (cf. Milnor [36]). Donc si la métrique riemannienne $Q_x^*$ est complète, on a

(5) $$\Lambda_{0,T}[g \text{ continu} \mid g_0 = x \text{ et } g_T = y] = \frac{d^2(x,y)}{2T}$$

Le triple résultat mentionné sur l'énergie minimale reste vrai pourvu que

(6) $$d(x,y) < d(x,\delta) = \varliminf_{z\to\delta} d(x,z)$$

où $\delta$ est le point à l'infini de $M$. Ainsi (6) implique (5).

Notons en passant que l'opérateur de Laplace-Beltrami $L$ (cf. Helgason [26]) associé à la métrique riemannienne $Q^*$ s'écrit $L = \Delta + Z$ où $Z$ est un champ de vecteurs sur $M$, et que d'après nôtre définition $L$ et $\Delta$ ont même transformée de Cramer.

<u>Dans le cas hypo-elliptique 6.2 (2)</u> la forme quadratique $Q_x^*$ peut prendre des valeurs infinies ; en chaque point $x$ le sous espace vectoriel $H_x$ des $v \in T_x(M)$ tels que $Q_x^*(v)$ soit fini sera appelé <u>sous - espace horizontal en x</u> ; en coordonnées locales cet espace est simplement l'image de $\mathbb{R}^n$ par $a(x)$, [où de $\mathbb{R}^k$ par $\sigma(x)$ lorsque $a(x) = \sigma(x)\sigma(x)^*$ et $\sigma(x)$ est une matrice $n$, $k$]. On peut définir (cf. Gaveau [25] pour un cas particulier de cette situation) "<u>l'action S(g) du chemin absolument continu</u> $g : [0,T] \to M$, par

(7) $$S(g) = T \int_0^T Q_{g_t}^*(g_t')\, dt$$

de sorte que pour un tel $g$ on a

(8) $$\frac{S(g)}{2T} = \lambda_{0,T}(g)$$

Les seules courbes d'action finie sont les courbes "horizontales", c'est à dire telles que $g'_t \in H_{g_t}$ pour presque tout t. Les courbes minimisant l'action S(g) pour $g_o = x$ et $g_T = y$ avec T, x, y donnés existent d'après les propriétés de $\lambda_{0,T}$ (cf. prop. 3.3). Lorsqu'elles sont lisses par morceaux, ce qui n'est pas toujours le cas (cf. Gaveau [24] [25]) elles coincident par morceaux avec les bicaractéristiques de $\Delta$ (cf. Gaveau [24] pour un cas particulier)

Nous noterons

(9) $$S(x,y) = \inf \{ S(g) \mid g_o = x,\ g_1 = y \}$$

l'action minimale qui permet de passer de x à y entre les temps 0 et 1 de sorte que

(10) $$\Lambda_{0,T} \{g \mid g_o = x,\ g_T = y \} = \frac{S(x,y)}{2T}$$

6.4 Théorème : Soit M une variété ; soit $\Delta$ un opérateur différentiel semi-elliptique du second ordre sur M, annulant les constantes ; soit b un champ de vecteurs sur M. Plaçons-nous sous les hypothèses précises du cas elliptique 6.2 (1), ou bien sous celles du cas hypo-elliptique 6.2 (2). Soit $(y_t)$ l'unique $(\Delta + b)$-diffusion sur M.

Notons $\lambda = \lambda_{0,1} = \mathcal{E}_{0,1}(M) \to [0, +\infty]$ la transformée de Cramer de $\Delta$, et $\Lambda = \Lambda_{0,1}$ la fonctionnelle associée. Posons

$$\mathcal{E}_x(M) = \{ g \in \mathcal{E}_{0,1}(M) \mid g_o = x \}.$$

Pour chaque $\theta > 0$ définissons le processus $z^\theta$ par $z^\theta_u = y_{\theta u}$, $0 \leq u \leq 1$, à valeurs dans $\mathcal{E}_{0,1}(M)$. Alors pour toute partie borélienne A de $\mathcal{E}_{0,1}(M)$ on a

$$-\Lambda(\mathring{A}) \leq \varliminf_{\theta\to0} \theta \log P_x(z^\theta \in A) \leq \varlimsup_{\theta\to0} \theta \log P_x(z^\theta \in A) \leq -\Lambda(\bar{A})$$

où $\mathring{A}$ , $\bar{A}$ sont respectivement l'intérieurs et l'adhérence de A dans $\mathcal{E}_x(M)$.

<u>Preuve</u> : Si $R_t$ est la fonction de transition de $(y_t)$, et $R_t^\theta$ celle de $(z_t^\theta)$ , il est évident que $R_t^\theta = R_{t\theta}$ . L'équation

$$R_t f(x) - f(x) = \int_o^t R_s(\Delta + b)\, f(x)\, ds$$

vraie pour toute $f : M \to R$ de classe 2 à support compact entraîne trivialement

$$R_t^\theta f(x) - f(x) = \int_o^{\theta t} R_s(\Delta + b)\, f(x)\, ds = \int_o^t R_u^\theta(\theta\Delta + \theta b)\, f(x)\, du$$

Par conséquent $z^\theta$ est la diffusion de générateur différentiel $D_\theta = \theta\Delta + \theta b$. Posons $\varepsilon = \sqrt{\theta}$ et $b_\varepsilon = \theta b$ pour écrire $D_\theta = \varepsilon^2\Delta + b_\varepsilon$ avec $\lim\limits_{\varepsilon\to 0} b_\varepsilon = 0$, uniformément sur tout compact. Le théorème est alors une conséquence directe du théorème 5.1, car $\lambda$ n'est autre que la transformée de Cramer du système perturbé $(\varepsilon^2\Delta + b_\varepsilon)_{\varepsilon>0}$ comme on l'a remarqué plus haut.

<u>Remarque</u> : <u>On notera que $\lambda$ ne dépend pas du drift b, mais seulement de $\Delta$.</u>

6.5 <u>Corollaire</u> : (Hypothèses "elliptiques" 6.2 (1) <u>ou bien</u> "hypo-elliptiques" 6.2 (2)). Soit M une variété, $\Delta$ un opérateur différentiel du second ordre semi-elliptique sur M, b un champ de vecteurs sur M. Soit $y_t$ la $(\Delta + b)$-diffusion sur M. Soit $S(x,y)$, $x,y \in M$ l'action minimale permettant d'aller de x à y entre les temps 0 et 1 (cf. 6.3(9)) ; rappelons que $S(x,y)$ <u>est complètement déterminée par la seule donnée de $\Delta$</u>. Alors pour tout borélien F de M on a

$$(1) \quad -\frac{1}{2} \inf_{z\in \overset{\circ}{F}} S(x,z) \leqslant \varliminf_{t\to 0} t\log P_x(y_t \in F) \leqslant \varlimsup_{t\to 0} t\log P_x(y_t \in F) \leqslant -\frac{1}{2} \inf_{z\in \bar{F}} S(x,z)$$

Dans le cas elliptique 6.2 (1), notons d la distance riemannienne associée à $\Delta$ (cf. 6.3) et posons $d(x,F) = \inf_{z \in F} d(x,y)$. Supposons que $F$ et $\overset{\circ}{F}$ aient même adhérence dans M, et que $d(x,\overline{F}) < d(x,\delta)$ où $\delta$ est la point à l'infini de M. Alors on a

(2) $$\lim_{t \to 0} t \log P_x(y_t \in F) = -\frac{1}{2} d^2(x,F)$$

Preuve : Posons pour $E \subset M$, E borélien, $x \in M$,

$$A_E = \{g \in \mathcal{E}_{0,1}(M) \mid g_o = x,\ g_1 \in E\}$$

Il est évident que $A_E$ est un borélien de $\mathcal{E}_x(M)$, que $\overline{A_E} \subset A_{\overline{E}}$ et que $A_{\overset{\circ}{E}} \subset (A_E)^\circ$. Puisque l'évenement $\{y_t \in F\}$ coincide avec $\{z^t \in A_F\}$ où $z^t$ est défini comme en 6.4, avec $\theta = t$, le théorème 6.4 montre que $\underline{\lim}_{t \to 0}$ et $\overline{\lim}_{t \to 0}$ de $t \log P_x(y_t \in E)$ sont dans l'intervalle $[-\Lambda(A_{\overset{\circ}{E}}) - \Lambda(A_{\overline{E}})]$.

De plus 6.3 (9), (10) donne par définition de S(x,y) et de $\Lambda$,

$$\Lambda(A_E) = \frac{1}{2} \inf_{z \in E} S(x,z)$$

ce qui prouve le premier résultat annoncé.

Dans le cas elliptique on a $S(x,y) = d^2(x,y)$ pourvu que $d(x,y) < d(x,\delta)$ pour tout $y \in \overline{E}$, donc si on suppose $d(x,\overline{E}) < d(x,\delta)$. Si $E$ et $\overset{\circ}{E}$ ont même adhérence, on a

$$d(x,\overset{\circ}{E}) = d(x,E) = d(x,\overline{E})$$

d'où $$\Lambda(A_{\overset{\circ}{E}}) = \Lambda(A_E) = \Lambda(A_{\overline{E}}) = \frac{1}{2} d^2(x,E)$$

ce qui prouve le second résultat annoncé.

6.6 Remarque : Le résultat 6.5 fait sentir que la densité p(t,x,y) de la fonction de transition de $y_t$, qui existe sous les hypothèses 6.2 (1) ou 6.2(2), et qui est définie par :

$$\int_F p(t,x,y)\, d\nu(y) = P_x(y_t \in F)$$

quelque soit F borélien de M, [$\nu$ est un volume riemannien arbitraire fixé] "devrait vérifier"

(1) " $\lim\limits_{t \to 0} t\log p(t,x,y) = \frac{1}{2} S(x,y)$ "

Ce résultat est correct dans le cas elliptique (Molchanov [37]) et a été prouvé pour quelques cas particuliers d'opérateurs hypo-elliptiques (Gaveau [24] [25]). Il existe cependant des exemples de diffusions gaussiennes à générateur hypo-elliptique pour lesquels ce résultat semble faux (Manankiandrianana [ 34]). Nous reviendrons (Ch.VI) sur ce point intéressant, ainsi que sur les méthodes plausibles pour passer du résultat "intégré" 6.5 aux resultats concernant le comportement de $t\log p(t,x,y)$ quand $t \to 0$. A ce propos l'argument de Gaveau [25] (qui obtient une inégalité dans un sens au lieu de l'égalité (1), pour une classe d'opérateurs hypo-elliptiques) nous semble plus que bref, pour ne pas dire elliptique.

6.7 Remarque : Considérons le cas où l'opérateur $\Delta$ s'écrit localement sous la forme $\sum\limits_{i=1}^{r} X_i^2 + Y$ avec $X_1 \dots X_r$ $Y$ champs $C^\infty$ tels que l'algèbre de Lie engendrée par $\{X_1 \dots X_r\}$ engendre $T_x(M)$ en chaque point alors $\Delta + b$ vérifie les hypothèses hypo-elliptiques 6.2 (2) quel que soit le drift $b$ de classe $C^\infty$. Ceci revient à dire que dans suffisamment de cartes locales, les colonnes de la matrice $a(x)$ des coefficients du 2^nd^ ordre de $\Delta$ engendrent (par crochet de Lie) l'espace tangent en chaque point ; ou encore qu'il existe localement des champs $X_1 \dots X_r$ qui en chaque point $x$ forment une base de l'espace horizontal $H_x$, et engendrent (par crochets de Lie) l'espace tangent en chaque point.

Dans ce cas on peut prouver que étant donnés $x,y$ quelconques de $M$, il existe dans tout voisinage de $y$ des points qui peuvent être joints à $x$ par une courbe horizontale $g$ d'action $S(g)$ finie ; [par exemple pour approcher un chemin tangent à $[X_1, X_2]$ on suit un "parallèlogramme" tangent successivement à $X_1$, $X_2$, $-X_1$, $-X_2$, dans une situation où $X_1$, $X_2$ sont horizontaux et où $[X_1, X_2]$ ne l'est pas]. Par suite pour tout ouvert F de M, les nombres $\inf_{y \in F} S(x,y)$ et $\inf_{y \in \bar{F}} S(x,y)$ sont positifs <u>finis</u> (et d'ailleurs positifs stricts si $x \notin \bar{F}$). La formule 6.5 (1) montre alors l'existence de deux nombres positifs <u>finis</u> c et d tels que pour t petit,

$$e^{-\frac{c}{2t}} \leqslant P_x(y_t \in F) \leqslant e^{-\frac{d}{2t}}$$

La formule 6.5 (1) fournit donc dans ce cas une estimation du premier terme du developpement de $\log P_x(y_t \in F)$ qui est moralement en $\frac{1}{t} \inf_{y \in F} S(x,y)$.

Cette remarque s'applique à fortiori au cas elliptique 6.2 (1) comme le montre la formule 6.5 (2). Par contre si les "coefficients des termes du 2nd ordre" de $\Delta$ ne suffisent pas à assurer l'hypoellipticité 6.2 (2) de $\Delta$, et si b est un drift $C^\infty$ tel que $(\Delta + b)$ vérifie 6.2 (2), la propriété de S(x,y) énoncée plus haut est trivialement fausse en général, c'est à dire qu'il peut exister des ouverts F tels que $\inf_{y \in F} S(x,y) = +\infty$. C'est par exemple le cas pour de nombreuses diffusions gaussiennes à générateur hypo-elliptique, (voir [34] pour des exemples de telles diffusions). Dans ce cas 6.5 (1) montre que $\lim_{t \to 0} t \log P_x(y_t \in F) = -\infty$ et fait pressentir la possibilité d'un comportement en $(-\frac{cte}{t^k})$, $k > 1$, pour $\log P_x(y_t \in F)$, pour certains ouverts F. L'examen des exemples fournis par [34] confirme cette possibilité, et on peut très probablement étendre ces calculs aux diffusions gaussiennes hypo-elliptiques. Cependant même dans ces cas il existe toujours des ouverts F pour lesquels $\inf_{y \in F} S(x,y)$ est fini (les voisinages d'extrémités de courbes horizontales issues de x par exemple) et donc pour lesquels le comportement de $\log P_x(y_t \in F)$ est en $(-\frac{cte}{t})$.

D'un point de une heuristique, il semble que la probabilité de suivre un tube "tangent" au drift b soit plus "petite" que celle de suivre des chemins horizontaux lorsque b n'est pas horizontal, et que la probabilité de suivre un chemin "tangent" à $[Z_1 [Z_2 \ldots Z_{s-1}, Z_s] \ldots]$

"soit" en $\exp(-\frac{cte}{t^k})$ où l'exposant k "est déterminé" par le nombre q de vecteurs $Z_i$ égaux à b, les autres $Z_j$ étant supposés horizontaux; l'argument de "parallèlogramme" mentionné plus haut laisse penser que k augmente avec q.

6.8 Corollaire de 6.4 : (Hypothèses "elliptiques" 6.2 (1))

Soit M une variété, $\Delta$ un opérateur différentiel du 2nd ordre semi-elliptique sur M, b un champ de vecteurs sur M. Soit $y_t$ la $(\Delta + b)$-diffusion sur M. Soit $d(x,y)$ la distance riemannienne associée à $\Delta$ (voir 6.3). Soit F une partie borélienne relativement compacte de M, telle que $\overset{\circ}{F}$, $\overline{F}$ et $M-\overset{\circ}{F}$ aient même frontière $\partial F$ . Soit $x \in F$ fixé. Supposons que $d(x, \partial F) < d(x,\delta)$ où $\delta$ est le point à l'infini de M. Soit $\tau_F$ le temps de première sortie de F pour la diffusion $(y_t)$. On a alors

(1) $$\lim_{t\to 0} t\log P_x(\tau_F \leqslant t) = -\frac{1}{2} d^2(x,\partial F)$$

Preuve : Considérons le processus $u \to z_u^t = y_{tu}$ où $0 \leqslant u \leqslant 1$ et $t > 0$ est fixé. Soit $\tau_F^t$ le temps de sortie de F pour $z^t$, qui vérifie clairement $t\,\tau_F^t = \tau_F$. Tout revient donc à calculer

(2) $$\lim_{t\to 0} t\log P_x(\tau_F^t \leqslant 1) = \lim_{t\to 0} t\log P_x(z^t \in H_F)$$

où l'on a posé

(3) $$H_F = \{g \in \mathcal{E}_{0,1}(M) \mid g_o = x,\ \tau_F(g) \leqslant 1\}$$

Posons aussi

(4) $$J_F = \{g \in \mathcal{E}_{0,1}(M) \mid g_o = x,\ \tau_F(g) < 1\}$$

Il est clair que si $\mathcal{E}_x(M) = \{g \in \mathcal{E}_{0,1}(M) \mid g_o = x\}$, l'ensemble $J_{\overline{F}}$ est ouvert dans $\mathcal{E}_x(M)$, tandis que l'ensemble $H_{\overset{\circ}{F}}$ est fermé dans $\mathcal{E}_x(M)$, pour la topologie III.1.2. Comme d'autre part on a trivialement

(5) $$J_{\overline{F}} \subset H_F \subset H_{\overset{\circ}{F}}$$

le théorème 6.4 prouve que $\underline{\lim}_{t\to 0}$ et $\overline{\lim}_{t\to 0}$ de tlog $P_x(z^t \in H_F)$ sont dans l'intervalle $[-\Lambda_{0,1}(J_{\overline{F}}), -\Lambda_{0,1}(H_{\mathring{F}})]$ où $\Lambda_{0,1}$ est la fonctionnelle de Cramer associée à $\Delta$.

Soient $g \in H_{\mathring{F}}$ et $u = \tau_{\mathring{F}}(g) \leqslant 1$. Alors on a $z = g_u \in (M-\mathring{F}) \cap \overline{\mathring{F}}$ par définition de $\tau_{\mathring{F}}$, et donc $z \in \partial F$. Par suite en notant $||v||_x$ la norme riemannienne sur $T_x(M)$ définie par $\Delta$ (cf. 6.3), on a

$$(6) \quad \lambda_{0,1}(g) \geqslant \frac{1}{2} \int_0^u ||g_t'||^2_{g_t} \geqslant \frac{1}{2u} d^2(x,z) \geqslant \frac{1}{2} d^2(x,\partial F)$$

où l'inégalité intermédiaire se déduit de 6.3 (5) et 6.3 (6). Par définition de $\Lambda_{0,1}$ nous avons donc

$$(7) \qquad \Lambda_{0,1}(H_{\mathring{F}}) \geqslant \frac{1}{2} d^2(x,\partial F)$$

Soit $z \in \partial F$ tel que $d(x,z) = d(x,\partial F)$. Soit g une géodésigne minimisante (brisée) joignant x à z, contenue dans $\overline{F}$. Il existe alors une géodésigne brisée h prolongeant strictement g du côté de z et sortant de $\overline{F}$ au point z. Soit f(t) la position au temps t d'un mobile partant de x à t = 0 et parcourant h à la vitesse constante $v = a\, d(x,y)$ où $a > 1$ est très proche de 1. On a alors

$$\tau_{\overline{F}}(f) = \frac{1}{a} < 1 \text{ et}$$

$$\lambda_{0,1}(f) = \frac{1}{2} \int_0^1 ||f_t'||^2_{f_t}\, dt = \frac{1}{2} a^2 d^2 (x,z) = \frac{1}{2} a^2 d^2 (x,\partial F)$$

ce qui, par définition de $\Lambda_{0,1}$ donne

$$\Lambda_{0,1}(J_{\overline{F}}) \leqslant \frac{1}{2} a^2 d^2 (x,\partial F)$$

Faisons tendre a vers 1 pour obtenir

$$(8) \qquad \Lambda_{0,1}(J_{\overline{F}}) \leqslant \frac{1}{2} d^2(x,\partial F)$$

De (5), (7), (8) on déduit $\Lambda_{0,1}(J_{\overline{F}}) = \Lambda_{0,1}(H_{\mathring{F}}) = \frac{1}{2} d^2(x,\partial F)$ ce qui prouve finalement $\lim_{t\to 0}$ tlog $P_x(\tau_F \leqslant t) = \frac{1}{2} d^2(x,\partial F)$.

6.9 <u>Remarque</u> : Sous les hypothèses de 6.8, on prouve facilement que si $\tau$ est le temps d'explosion de $y_t$ on a

$$\overline{\lim_{t\to 0}}\ t\log P_x(\tau \leqslant t) \leqslant -\frac{1}{2} d^2(x,\delta)$$

La minoration analogue pour $\underline{\lim}_{t\to 0}$ ne peut pas s'obtenir par la même méthode qu'en 6.8, car M n'a pas "d'extérieur" à priori. Pourtant si M est un ouvert F relativement compact d'une variété $\tilde{M}$, et si $\Delta + b$ vérifie 6.2 (1) sur $\tilde{M}$, on a $d(x,\delta) = d\,(x,\partial F)$ de sorte que grâce à 6.8, si F vérifie les conditions indiquées en 6.8 on peut garantir $\lim_{t\to 0} t\log P_x(\tau \leqslant t) = -\frac{1}{2} d^2(x,\delta)$. Il est naturel de se demander quel type de comportement à l'infini des coefficients de $\Delta + b$ sur M permet de prouver cette égalité.

Notons en passant que si la métrique riemannienne définie par $\Delta$ est complète, on a $d(x,\delta) = +\infty$ et donc certainement $\lim_{t\to 0} t\log P_x(\tau \leqslant t) = -\infty$. Ceci signifie intuitivement que du point de vue de l'estimation des probabilités "à la Cramer-Chernoff" le processus $y_t$ "n'explose pas" sur $[0,t]$ pour $t \to 0$.

# CHAPITRE VI

## QUESTIONS SANS REPONSES

Au fil du texte précédent, nous avons mentionné en passant certaines questions ouvertes. Décrivons maintenant quelques autres problèmes.

### 1. HYPOTHESES SUR LES COEFFICIENTS DE $\Delta$ :

1.1. Pour un système perturbé $d\, y_t^\varepsilon = \varepsilon\, \sigma(y_t^\varepsilon)\, d\, \beta_t + b\, (y_t^\varepsilon)\, dt$ sur un ouvert U de $\mathbb{R}^n$, nous avons étudié $\lim\limits_{\varepsilon \to 0} \varepsilon^2 \log P\, (y^\varepsilon \in A)$ en supposant en particulier que $\sigma\,(x)$ était de classe 1 en x. Que se passe-t-il quand $\sigma\,(x)$ est seulement lipschitzienne ? En effet, à priori, les résultats $C^1$ gardent encore tout leur sens dans ce cas.

1.2. Si on suppose $\{\sigma$ et b höldériens$\}$, ou bien même $\{\sigma$ de classe 1, b höldérien$\}$, on se confronte au problème de non unicité des solutions (déterministes !) de $y'_t = b\,(y_t)$, ou de $y'_t = b\,(y_t) + \sigma\,(y_t)\, f'_t$. On rejoint ainsi une question posée par De Giorgi et P. Baldi sur la limite de $y^\varepsilon_{[0,1]}$ dans ce cas.

1.3. Pour un système dynamique perturbé $y^\varepsilon$ sur une variété M, de générateur infinitésimal $(b + \varepsilon^2\ \Delta)$, nous avons introduit pour recoller les résultats locaux une hypothèse peut-être utile, celle de la constance du rang de la "matrice des coefficients d'ordre 2" de $\Delta$. Il devrait être possible de s'en passer. Le point à étudier est le suivant : étant donnés un tube $\Gamma$ de rayon r, d'axe $f : [0, 1] \to M$, un nombre $\eta > 0$, et x proche de $f_0$, trouver un chemin $g : [0, 1] \to M$ tel que $g_0 = x$, $g \in \Gamma$ et $\lambda\,(g) \leqslant (1+\eta)\ \lambda\,(f)$. Notre hypothèse de rang constant permet de garantir $\left| \dfrac{\lambda(g)}{\lambda(f)} - 1 \right| \leqslant \eta$, ce qui est plus qu'il n'en faut pour poursuivre notre argument de recollement.

2. <u>LES EQUIVALENTS PRECIS DE P $(y^{\varepsilon} \in A)$</u> :

2.1. Lorsque $\beta$ est un brownien k-dimensionnel, Schilder [44] a obtenu un développement du type

(1) $E \exp \Theta (\varepsilon\beta) = \exp (\frac{K}{\varepsilon^2}) \left[c + c_1 \varepsilon + c_2 \varepsilon^2 + \ldots\right]$

où $\Theta$ est une fonctionnelle "lisse" définie sur l'espace de trajectoires $C_{0,1} (\mathbb{R}^n)$, et où $K$, $c$, $c_1$ ... se calculent à partir de $\Theta$.

Lorsque $d\, y_t^{\varepsilon} = \varepsilon\, \sigma (y_t^{\varepsilon})\, d\, \beta_t + b (y_t^{\varepsilon})\, dt$

on a $y^{\varepsilon} = F (\varepsilon\beta)$ où $F$ est, en général, seulement mesurable. On a envie d'écrire

(2) $E \exp \Theta (y^{\varepsilon}) = E \exp \Theta \circ F (\varepsilon\beta)$

et d'en déduire un développement de $E \exp \Theta (y^{\varepsilon})$. Mais $F$, et donc $\Theta \circ F$ , n'est pas lisse en général. Il est tentant de penser que l'approximation déterministe $G$ de $F$ que nous utilisons au chapitre III, définie par

(3) $g = Gf \Longleftrightarrow g'_t = \sigma (g_t)\, f'_t + b (g_t)$ p.p. $t$

pourrait remplacer $F$. Autrement dit, est-ce que $E \exp \Theta (y^{\varepsilon}) = E \exp \Theta \circ F (\varepsilon\beta)$ est équivalent à $E (\exp \Theta \circ G (\varepsilon\beta))$. Telle quelle la question n'a pas de sens rigoureux, car $G (\varepsilon\beta)$ n'est pas défini en général. Il faudrait peut-être "améliorer" l'approximation $G$ pour obtenir la même conclusion, ou alors essayer de sauter cette étape de l'argument pour passer plus directement à 2.2. Dans la situation (exceptionnelle en dimension $n$ pour $n > 1$) où les champs définis par les colonnes de $\sigma (x)$ forment une algèbre de Lie <u>abélienne</u> , on peut choisir $F$ lisse (cf. Doss [17], et Yamato) et cette méthode fournit très directement (Doss [17]) un développement de $E \exp \Theta (y^{\varepsilon})$.

2.2. Une fois "franchie", ou contournée l'étape 2.1, il serait intéressant d'encadrer une fonction indicatrice $1_A$ où $A \subset C_{0,1} (\mathbb{R}^n)$ par

$\Theta_1 \leqslant 1_A \leqslant \Theta_2$ avec $\Theta_1$, $\Theta_2$ lisses, pour obtenir un développement de la forme

$$(4)\quad P\,(y^\varepsilon \in A) = \left[\exp\left(-\frac{\Lambda\,(A)}{\varepsilon^2}\right)\right]\ \left[a + a_1\,\varepsilon + a_2\,\varepsilon^2 + \dots\right]$$

où a, $a_1$, ... se calculent à partir de A.

2.3. Problème discret analogue au précédent : pour une suite de v.a. indépendantes $X_1 \dots X_n$ de même loi, à valeurs dans un Banach E trouver un développement du type

$$P\,(\overline{X}_n \in A) = \left[\exp\,(-\,n\,\Lambda\,(A))\right]\ \left[c + \frac{c_1}{n} + \frac{c_2}{n^2} + \dots\right]$$

où c, $c_1$ ... dépendent de A. Rappelons (cf. Chapitre Zéro, § 4) que ce problème a été traité en dimension 1.

## 3. PASSAGE DE $\lim\limits_{t\to 0} t \log P_x\,(y_t \in F)$ à $\lim\limits_{t\to 0} t \log p\,(t,x,y)$

3.1. Considérons la (b+$\Delta$)-diffusion ($y_t$) sur la variété M. On a pu étudier (chapitre V) $\lim\limits_{t\to 0} t \log P_x\,(y_t \in F)$ pour F ouvert de M. Soit p (t,x,y) la densité de la fonction de transition de ($y_t$). Pour relier $\lim\limits_{t\to 0} t \log p\,(t,x,y)$ à $\lim\limits_{t\to 0} t \log P_x\,(y_t \in F)$, on peut essayer de faire tendre vol (F) vers 0 avec t, en prenant $F_t$ = boule de centre y et de rayon t dans M. Le paquet $\Gamma_t$ de trajectoires continues sur [0, 1] qui partent de x et finissent dans $\overline{F_t}$ au temps 1 contient toujours la géodésique qui va de x à y. De sorte que $\Lambda_{0,1}\,(\Gamma_t)$ reste majoré par $\frac{1}{2}\,d^2\,(x,y)$, où d est la distance riemannienne associée à $\Delta$. Les résultats d'uniformité du chapitre IV, § 1, et le changement de temps utilisé au Chapitre V montrent alors

$$(1)\quad \overline{\lim_{t\to 0}}\ t \log P_x\ (y_t \in F_t) \leqslant \frac{1}{2}\,d^2\,(x,\,y)$$

bien que $F_t$ dépende de t. Ceci s'écrit

(2) $\overline{\lim\limits_{t\to 0}}\left[t \log \operatorname{vol}(F_t) + t \log \frac{1}{\operatorname{vol}(F_t)} \int_{F_t} p(t,x,z)\, dz\right] \leqslant \frac{1}{2} d^2(x,y)$

Comme vol $F_t$ = cte . $t^n$ , le premier terme tend vers 0 avec t. Il suffirait alors de savoir que

(3) $\{t \mid \operatorname{grad}_z \log p(t,x,z)|$ reste borné pour x fixé et z dans un voisinage fixe de y$\}$

pour conclure que $\overline{\lim\limits_{t\to 0}}\ t \log p(t,x,y) \leqslant \frac{1}{2}\ d^2(x,y)$. Sur ce point nous renvoyons au résultat énoncé par Gaveau dans [25] , cas hypoelliptique, où les détails de la preuve sont omis.

La minoration de t log p (t,x,y) , par contre, ne peut pas utiliser le même résultat d'uniformité, car le même argument que ci-dessus amène à considérer la probabilité $P(y^\varepsilon \in \Gamma_\varepsilon)$ où $y^\varepsilon$ est une $(\varepsilon^2\, \Delta)$-diffusion et $\Gamma_\varepsilon$ un tube d'axe fixe dans $C_{[0,1]}$ mais dont le rayon tend vers 0 avec $\varepsilon$.

Evidemment le cas elliptique ayant été complètement traité (Molchanov [37]) la question considérée ici devrait être étudiée dans le contexte des opérateurs $\Delta$ hypoelliptiques au sens fort suivant : $\Delta = X_1^2 + \ldots + X_r^2 + Y$ où l'algèbre de Lie engendrée par $\{X_1 \ldots X_r\}$ est de dimension n = dim M en tout point.

## 4. AMELIORATION DES RESULTATS LORSQUE $\lim\limits_{\varepsilon\to 0} \varepsilon^2 \log P(y^\varepsilon \in A) = -\infty$

Nous avons signalé (chapitre V) des situations où $y^\varepsilon$ étant un système dynamique perturbé, et A étant inclus dans $C_{0,1}(\mathbb{R}^n)$, on a $\lim\limits_{\varepsilon\to 0} \varepsilon^2 \log P(y^\varepsilon \in A) = -\infty$ [cf. situations "en temps petit" où $\lim\limits_{t\to 0} t \log P(y_t \in F) = -\infty$] . Le problème se pose alors de trouver la bonne puissance de $\varepsilon^2$ qui garantira $\{\overline{\lim\limits_{\varepsilon\to 0}}\ \varepsilon^{2\alpha} \log P(y^\varepsilon \in A)$ fini $\}$

L'exemple des diffusions gaussiennes (hypoelliptiques) en dimension n nous semblent un terrain utile d'exploration préliminaire pour ce type de résultats (cf [34] pour un cas particulier).

# BIBLIOGRAPHIE


[1] R. AZENCOTT - Methods of localisation and diffusions on manifolds (à paraître)

[2] R. AZENCOTT - Behaviour of diffusion semi-groups at infinity. Bull. Soc. Math. France 102, 1974, p. 193-240

[3] R. AZENCOTT - G. RUGET - Mélanges d'équations différentielles et grands écarts à la loi des grands nombres. Z. Wahrscheinlichkeit. verw. Gebiete 38, 1977, p. 1-54

[4] A. BADRIKIAN - Fonctions aléatoires linéaires et mesures cylindriques. Lecture Notes Math. 139 (1970)

[5] R.R. BAHADUR - Rates of convergence of estimates and test statistics Ann. Math. Stat. 38 (1967) p. 303-324

[6] R.R. BAHADUR - Some limit theorems in statistics. SIAM (1971) Philadelphia

[7] R.R. BAHADUR - S.L. ZABELL - Large deviations of the sample mean in general vector spaces (à paraître)

[8] P. BILLINGSLEY - Ergodic theory and information. Wiley - New-York (1965)

[9] J.M. BONY - Cours au C.I.M.E. (1969)

[10] H. CHERNOFF - Asymptotic efficiency for tests based on the sum of observations. Ann. Math. Stat. 23 (1952), p. 493-507

[11] J. CHRISTENSEN - Topology and borel structures. North Holland Amsterdam (1974)

[12] COURANT - HILBERT - Methods of Mathematical Physics. Interscience (1962)

[13] P. COURREGE - P. PRIOURET - Recollement de processus de Markov. Pub. Inst. Stat. Univ. Paris 14 (1965) p. 275-325

[14] H. CRAMER - Sur un nouveau théorème limite de la théorie des probabilités. Colloquium on theory of probability. Paris-Hermann (1937)

[15] M. DONSKER - S. VARADHAN - Asymptotic evaluation of certain Markov processes expectations for large time. I, II, III
Comm. Pure App. Math. 28 (1975) p. 1-47
29 (1976) p. 279-301
29 (1976) p. 389-461

[16] M. DONSKER - S. VARADHAN - Large deviations for Markov processes and asymptotic evaluation of certain expectations for large time in "Probabilistic methods in differential equations".
Lecture Notes in Math., Springer, 451 (1975) p. 82-87

[17] H. DOSS - Quelques formules asymptotiques pour les petites perturbations de systèmes dynamiques (à paraître)

[18] E. DYNKIN - Markov processes I, II. Springer

[19] N. EL KAROUI - Thèse Université PARIS VI

[20] W. FELLER - Limit theorems for probabilities of large deviations.
Z. Wahrscheinlich. verw. Geb. 14 (1969) p. 1-20

[21] W. FELLER - Generalisation of a theorem of Cramer.
Trans. Am. Math. Soc. 54 (1943) p. 361-372

[22] X. FERNIQUE - Régularité des trajectoires des fonctions aléatoires gaussiennes. Ecole d'Eté Saint-Flour IV (1975). Lecture Notes Math. 480

[23] M.I. FREIDLIN - Action functional for a class of stochastic processes
Th. Prob. App. 17 (1972) p. 511-515

[24] B. GAVEAU - Principe de moindre action. Propagation de la chaleur. Estimées sous elliptiques sur certains groupes nilpotents. Acta Math. 139 (1977) p. 96-153

[25] B. GAVEAU - Systèmes hamiltoniens associés à certains opérateurs hypoelliptiques. Bull. Sci. Math. (1978)

[26] S. HELGASON - Differential geometry and symmetric spaces. Academic Press. New-York (1962)

[27] N. KRYLOV - Control of a solution of a stochastic integral equations. Th. Prob. App. 17 (1972) p. 114-131

[28] KULLBACK - Information theory

[29] H. KUO - Gaussian measures in Banach spaces. Lecture Notes in Math. 526 (1975)

[30] O. LANFORD - Entropy and equilibrium states in classical statistical mechanics. Lecture Notes Physics 20 (1971) p. 1-113

[31] Y. LINNIK - Large deviations for the sum of independent variables. Proc. $4^{\underline{th}}$ Berkeley Symp. 2 (1961) p. 289-306

[32] P.L. LIONS - Résolution des problèmes généraux de Bellman-Dirichlet C.R. Acad. Sci. Paris 287 (1978) p. 747-750

[33] P.L. LIONS - Contrôle de diffusions dans $R^n$. C.R. Acad. Sci. Paris (à paraître)

[34] D. MANANKIANDRIANANA - Noyau de la chaleur d'un opérateur hypoelliptique dégénéré (comportement pour des temps petits). C.R. Acad. Sci. Paris (à paraître)

[35] P.A. MEYER - Probabilités et Potentiel - Hermann - Paris (1966)

[36] J. MILNOR - Morse theory - Ann. Math. Studies 51. Princeton (1963)

[37] S. MOLCHANOV - Diffusions et géométrie riemannienne. Uspetchi Mat. Nayk. 30 (1975) p. 3-59

[38] A. NAGAEV - Integral limit theorems taking large deviations into account when Cramer's conditions does not hold I, II Th. Prob. Appl. 14 (1969) p. 51-64, p. 193-208

[39] V. PETROV - Sums of independent random variables. Springer-Verlag New-York (1975)

[40] P. PRIOURET - Diffusions et équations différentielles stochastiques. in "Ecole d'Eté Saint Flour III - 1973". Lecture Notes in Math. Springer-Verlag 390 (1974)

[41] R. ROCKAFELLAR - Measurable dependence of convex sets on parameters. J Math. Ana. App. 28 (1969) p. 4-25

[42] R. ROCKAFELLAR - Convex analysis-Princeton University Press. Princeton (1970)

[43] I. SANOV - On the probability of large deviations of random variable Selected Trans. Math. Stat. Prob. 1 (1961) p. 213-244

[44] M. SCHILDER - Asymptotic formulas for Wiener integrals. Trans. Am. Math. Soc. 125 (1966) p. 63-85

[45] L. SCHWARTZ - Semi martingales sur des variétés. Martingales conformes sur des variétés analytiques complexes. Cours Ecole Polytechnique (1978) (à paraître)

[46] Séminaire de Probabilités, Paris VII (1977-79) (à paraître)

[47] Séminaire de Statistiques, Orsay 1977-78 (à paraître)

[48] A.V. SKOROKHOD - Note on Gaussian measures in a Banach space Th. Prob. Appl. 15 (1970) p. 508

[49] D. STROOCK - S.R. VARADHAN - Diffusion processes with continuous coefficients. Comm. Pure App. Math. 22 (1969) p. 345-400 et p. 479-530

[50] S. VARADHAN - Asymptotic probabilities and differential equations Comm. Pure App. Math 1 (1966), p. 261-286

[51] A.D. VENTSEL - Rough limit theorems on large deviations for Markov processes I, II. Th. Prob. Appl. 21 (1976) p. 227-242 et p. 499-512

[52] A.D. VENTSEL - Action functional for gaussian random functions. Th. Prob. Appl. 17 (1972) p. 515-517

[53] A.D. VENTSEL - M.I. FREIDLIN - On small random perturbations of dynamical systems. Russian Math. Surveys 25 (1970) p. 1-55

[54] A.D. VENTSEL - M.I. FREIDLIN - Some problems concerning stability under small random perturbations. Th. Prob. App. 17 (1972) p. 269-283