Probab. Th. Rel. Fields. 80, 169–215 (1988)

Probability Theory and Related Fields

Large Deviations and Stochastic Flows of Diffeomorphisms *

P.H. Baxendale [1], and D.W. Stroock [2]

[1] Department of Mathematics, DRB 306 University of Southern California, Los Angeles, CA 90089-1113, USA

[2] Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Summary. Previous results in the theory of large deviations for additive functionals of a diffusion process on a compact manifold M are extended and then applied to the analysis of the Lyapunov exponents of a stochastic flow of diffeomorphisms of M. An approximation argument relates these results to the behavior near the diagonal Δ in M^2 of the associated two point motion. Finally it is shown, under appropriate non-degeneracy conditions, that the two-point motion is ergodic on $M^2-\Delta$ if the top Lyapunov exponent is positive.

Introduction

Let M be a connected, compact C^∞ manifold of dimension N and consider a diffusion $\{x_t\colon t \geqq 0\}$ on M which is governed by the Statonovich stochastic differential equation

$$(0.1)\qquad dx_t = \sum_{k=1}^{d} X_k(x_t) \circ d\theta_k(t) + X_0(x_t)\, dt,$$

where $X_0, \ldots, X_d$ are C^∞-vector fields on M and $\{(\theta_1(t), \ldots, \theta_d(t))\}$ is a standard $\mathbb{R}^d$-valued Brownian motion. In particular, because the differentials are taken in the sense of Stratonovich, note that the associated generator is

$$(0.2)\qquad L = \tfrac{1}{2} \sum_{k=1}^{d} X_k^2 + X_0.$$

Denote by $\{A_t\colon t>0\}$ a (possibly Banach space-valued) additive functional of $\{x_t : t \geqq 0\}$. Section 1 contains various general results about the large deviations

* At the period when this research was initiated, both authors where guests of the I.M.A. in Minneapolis. The first author was at Aberdeen University, Scotland when this article was prepared. Throughout the period of this research, the second author has been partially supported by N.S.F. grant DMS-8611487 and ARO grant DAAL03-86-K-171

of $\frac{A_t}{t}$ as $t \to \infty$. As soon as these generalities are established they are applied to two important cases. The first of these is the case when $\frac{A_t}{t}$ is the $\mathbf{M}_1(M)$-valued ($\mathbf{M}_1(M)$ denotes the space of probability measures on M) process given by

$$(0.3) \qquad A_t = \int_0^t \delta_{x_s}\, ds,$$

where δ_x denotes the unit (Dirac) mass at $x \in M$. In particular, when L is sufficiently non-degenerate that $\{x_t\colon t>0\}$ is ergodic with a stationary measure $m_L \in \mathbf{M}_1(M)$, the ergodic theorem predicts that $\frac{A_t}{t} \Rightarrow m_L$ and the goal is to describe the rate at which large deviations of $\frac{A_t}{t}$ from m_L occur. This is the case studied originally by Donsker and Varadhan in [11]. The second case is the one most immediately related to the other topics in this article. In this case $A_t = y_t$ where

$$(0.4) \qquad y_t = \sum_{k=1}^{d} \int_0^t Y_k(x_s) \circ d\theta_k(s) + \int_0^t Y_0(x_s)\, ds,$$

for a given set $\{Y_0, \ldots, Y_d\} \subseteq C^\infty(M; \mathbb{R}^\nu)$. Such a situation was discussed earlier in [21] in connection with the rate at which the solution $\{z_t\colon t>0\}$ to a linear stochastic differential equation in $\mathbb{R}^{N+1}$ implodes or explodes. In that setting, $M = \mathbf{S}^N$, $x_t = \frac{z_t}{|z_t|}$, and $y_t = \log \frac{|z_t|}{|z_0|}$.

The main general large deviation result in Sect. 1 is Theorem 1.7. This theorem is stated in such a way that it covers both the above cases. Moreover, it contains one significant technical improvement over the results in [11], [20], and [21]. Namely, those earlier results demanded that the transition function for $\{x_t\colon t>0\}$ be "nicely" related to m_L whereas Theorem 1.7 only requires that an appropriate time-average of the transition function enjoy that property. In particular, it covers the situations when the set $\{X_0, \ldots, X_d\}$ satisfies Hörmander's condition (i.e. it generates $\mathbf{T}_x(M)$ at every $x \in M$) and the controllability hypothesis explained in Corollary 1.6. The major technical difficulties involved are dealt with in Theorem 1.5 and its Corollary 1.6, both of which seem to be new. Once these difficulties have been overcome, the proofs of Theorem 1.7 and Corollary 1.10 follows the same pattern as the arguments given in [20] and [21].

The first special case (the one when A_t is given in (0.3)) is treated in Corollary 1.11. The formula found for the rate function (denoted here by J instead of the more usual I) in terms of L is the same as the expression given in [11]. The second case (when $A_t = y_t$ and y_t is the quantity in (0.4)) is the focus of Corollary 1.12. Initially the corresponding rate function (this one is denoted by I) is identified as the Legendre transform (or convex conjugate) of the logarithmic generating function $\Lambda: \mathbb{R}^\nu \to \mathbb{R}$ defined by

$$\Lambda(\eta) = \lim_{t\to\infty} \frac{1}{t} \log(E[\exp[(\eta, y_t)_{\mathbb{R}^\nu}]]), \qquad \eta \in \mathbb{R}^\nu. \tag{0.5}$$

(The existence of this limit is part of the content of Corollary 1.12.) Subsequently, the Cameron-Martin-Girsanov transformation is used to find an alternative expression for I (cf. (1.17)). Although from the standpoint of large deviation theory it is the rate function I which is of paramount interest, it is the logarithmic generating function Λ which is most important for the analysis of Lyapunov exponents. To see this, recall the example discussed following (0.4) and note the obvious connection in this example between Λ and the stability properties of the original process $\{z_t: t \geqq 0\}$. Indeed, one sees that Λ is here what is called *the moment Lyapunov function*; and it is considerations of this sort which connect Sect. 1 with the other sections of this paper.

In addition to what has already been said, there is one more technical advance contained in Sect. 1. Namely, in the case when the vector fields $X_1, \ldots, X_d$ themselves (i.e. without X_0) generate the tangent space at each point and $X_0 = \sum_{k=1}^{d} \alpha_k X_k$ for some $\{\alpha_k\}_1^d \subseteq C^\infty(M; \mathbb{R})$, then considerable improvement can be made in the expressions for the rate function I and the logarithmic moment generating function Λ. These improved expressions appear in Theorem 1.25 and enable one to obtain what appears to be a new formula for the quadratic growth rate of $\Lambda(\eta)$ as $|\eta| \to \infty$. In an attempt not to encumber the presentation with somewhat unrelated details, most of the preparations for the proof of Theorem 1.25 have been put in the Appendix (where Theorem A.8 may be of some independent interest).

Turning to the contents in Sect. 2, denote by $\{\xi_t: t \geqq 0\}$ the stochastic flow of diffeomorphisms determined by the vector fields $X_0, \ldots, X_d$ through the stochastic differential equation (0.1). Thus, for each $x \in M$, $\{\xi_t(x): t \geqq 0\}$ coincides almost surely with the solution $\{x_t: t \geqq 0\}$ to (0.1) with initial condition $x_0 = x$. Next, denote by $D\xi_t$ the derivative of this flow of (random) diffeomorphisms:

$$D\xi_t(x): \mathbf{T}_x(M) \to \mathbf{T}_{\xi_t(x)}(M).$$

After giving M a Riemannian structure, it is shown that $t \in [0, \infty) \mapsto \det(D\xi_t(x)) \in \mathbb{R}$ fits into the framework of Corollary 1.12; and the associated Λ is interpreted as a measure of the extent to which the diffeomorphisms ξ_t fail to be measure-preserving. (See, in particular, Corollary 2.14, especially part ii).) In order to make a similar analysis of $t \in \mathbb{R} \mapsto |D\xi_t(x)(v)| \in (0, \infty)$ for a fixed $v \in \mathbf{T}_x(M) \setminus \{0\}$, it is necessary (just as in the example following (0.4)) to move to the associated sphere bundle $\mathbf{S}M$. To be precise, let $t \in [0, \infty) \mapsto \tilde{\xi}_t(v) \in \mathbf{S}M$ be defined so that $\tilde{\xi}_t(v)$ is the element of $\mathbf{S}M_{\xi_t(x)}$ obtained by normalizing $D\xi_t(x)(v)$. Theorem 2.15 says that (under suitable non-degeneracy assumptions about the vector fields on $\mathbf{S}M$ determining the flow $\tilde{\xi}_t$) once again Corollary 1.12 applies and yields information about the long time behavior of

$\log|D\xi_t(x)(v)|$. In particular, by analogy with the example following (0.4), the corresponding logarithmic generating function $\tilde{\Lambda}$ is again called the moment Lyapunov function; and, as is explained in iv) of Corollary 2.14, it measures the failure of the diffeomorphisms ξ_t to be isometric. In addition, both Λ and $\tilde{\Lambda}$ are intimately connected with the Lyapunov exponents $\lambda_1 \geqq \ldots \geqq \lambda_N$ of ξ_t. Two of these connections are made explicit in (2.11) and (2.12). The remainder of Sect. 2 contains some comparisons of Λ to $\tilde{\Lambda}$ and geometric interpretations of certain degenerate behavior of these functions (cf. Corollary 2.14 and Theorem 2.15).

Because the analysis of Lyapunov exponents is based entirely on linearization of the equations governing the flow under consideration, it is not immediately clear to what extent predictions based on such an analysis can be relied on to give reliable information about the true behavior of the flow. Sections 3 and 4 are devoted to an examination of this problem. Thus, the *two-point motion* $(\xi_t(x), \xi_t(y))$ $(x \neq y)$ is introduced. The state space for this process is the open $\hat{M} \equiv M^2 \backslash \Delta$, where Δ is the diagonal in M^2. The ability of this process to stay on $\hat{M}$ can be seen as a consequence of the degeneration of its generator $L^{(2)}$ (cf. (3.3)) at Δ; and it is the nature of this degeneracy which determines whether the two-point motion (as a process on $\hat{M}$) is transient. The principle behind the analysis given in Sect. 3 is based on the idea that, as one moves toward Δ, properties of the two-point motion are increasingly accurately reflected by properties of $D\xi_t(x)(v)$ for small $|v|$. In particular, near Δ, the generator $\mathbf{T}L$ of $D\xi_t(x)(v)$ ought to be comparable to $L^{(2)}$ and should act as a source of comparison functions. The comparison functions produced in Theorem 3.18 should be viewed as examples of this line of reasoning. (The proximity of $\mathbf{T}L$ to $L^{(2)}$ is clearest when one parametrizes a neighborhood of Δ in a polar coordinate system $(r, \theta) \in (0, \delta) \times \mathbf{S}M$; and it is for this reason that polar coordinates are introduced.)

Once the test functions in Theorem 3.18 have been constructed, their application to questions about the behavior of the two-point process is quite standard: they are used to find sub- and supermartingales with which to estimate the probability that, having gotten into a neighborhood of Δ, the two-point process will ever leave that neighborhood. Theorem 3.19 is devoted to the case in which the top Lyapunov exponent λ_1 is strictly positive; and what is shown is that the two-point motion in this case will, with probability one, exit a small neighborhood of Δ without ever touching Δ. In fact, rather precise estimates on the time spent in such a neighborhood are obtained. As a consequence of these considerations, one knows, of course, that the two-point motion is non-transient when $\lambda_1 > 0$. This same conclusion can be drawn from the results in a recent paper by Ledrappier and Young [18]; although their technique does not yield the quantitative information contained in Theorem 3.19. The case when $\lambda_1 < 0$ is not treated in this article since it has been already handled (by quite different techniques) in [7].

Finally, in Sect. 4 the results obtained in Sect. 3 about the two-point process near Δ are used to show that, under suitable non-degeneracy conditions, the two-point process is positively recurrent on $\hat{M}$ if $\lambda_1 > 0$. In fact, quite precise estimates are found for the mass assigned by the (unique) stationary measure to neighborhoods of Δ (cf. Theorem 4.6). Like the results in Sect. 3, a pleasing aspect of the results in Sect. 4 is that they not only confirm qualitative predictions based on the Lyapunov exponents but show how one can extract quantitative information from a knowledge of the functions Λ and $\tilde{\Lambda}$.

1. Some Large Deviation Results

Let $V_0, \ldots, V_d$ be elements of $C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ and define V_α for $\alpha = (\alpha_1, \ldots, \alpha_l) \in \{0, \ldots, d\}^l$ so that $V_\alpha = V_k$ if $\alpha = (k)$ and $V_\alpha = [V_k, V_\beta]$ if $l \geqq 2$, $\beta = (\alpha_1, \ldots, \alpha_{l-1})$, and $\alpha_l = k$. Assume that there exist an $l \in Z^+$ and an $\varepsilon > 0$ such that

(1.1)
$$\sum_{|\alpha| \leqq l} (V_\alpha, \eta)^2_{\mathbb{R}^N} \geqq \varepsilon |\eta|^2 \qquad \eta \in S^{N-1}.$$

Set $\Omega = \Omega_{\mathbb{R}^N} = C([0, \infty); \mathbb{R}^N)$ and for each $x \in \mathbb{R}^N$ denote by Q_x the probability measure on $(\Omega, \mathscr{B}_\Omega)$ which solves the martingale problem for $1/2 \sum_1^d V_k^2 + V_0$ starting from x. Then $\{Q_x : x \in \mathbb{R}^N\}$ is a Feller continuous strong Markov family. Next, for $c \in C_b^\infty(\mathbb{R}^N)$, define

$$Q^c(t, x, \Gamma) = E^{Q_x}\left[\exp\left(\int_0^t c(x(s))\, ds\right), x(t) \in \Gamma\right].$$

Given a function $\psi \in C_0^\infty([0, \infty))$, define

$$Q_\psi^c(x, \Gamma) = \int_0^\infty \psi(t)\, Q^c(t, x, \Gamma)\, dt$$

for $x \in \mathbb{R}^N$ and $\Gamma \in \mathscr{B}_{\mathbb{R}^N}$.

(1.2) **Theorem.** *Let $\psi \in C_b^\infty([0, \infty))$ and $c \in C_b^\infty(\mathbb{R}^N)$ be given, and define $Q_\psi^c(x, \cdot)$ accordingly. Then, for each $x \in \mathbb{R}^N$, $Q_\psi^c(x, \cdot)$ is absolutely continuous and there is a $q_\psi^c(x, \cdot) \in C^\infty(\mathbb{R}^N \setminus \{x\})$ such that $Q_\psi^c(x, dy) = q_\psi^c(x, y)\, dy$. Moreover, for each $n \geqq 0$ there is an $M_n \in (0, \infty)$, which depends only on the V_k's and c, such that*

(1.3)
$$\|q_\psi^c(x, \cdot)\|_{C^n(\mathbb{R}^N \setminus \overline{B(x,r)})} \leqq M_n \sum_{m=0}^n \left(\int_{(0,\infty)} \frac{\exp(-r^2/M_n t + M_n t)}{(t \wedge 1)^{M_n}} |\psi^{(m)}(t)|^4\, dt\right)^{1/4},$$
$$0 < r \leqq 1.$$

Proof. Set $\rho(y) = (2 + \sin(y))$ for $y \in \mathbb{R}^1$, and, for $z = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^N \times \mathbb{R}^1$, set

$$W_0(z) = \begin{bmatrix} \rho(y)\, V_0(x) \\ 0 \end{bmatrix}$$
$$W_k(z) = \begin{bmatrix} \rho(y)^{1/2}\, V_k(x) \\ 0 \end{bmatrix}, \qquad 1 \leqq k \leqq d$$
$$W_{d+1}(z) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Let $\mathscr{W}$ denote $(d+1)$-dimensional Wiener measure on

$$\Theta=\left\{\theta\in C([0,\infty);\mathbb{R}^{d+1}):\ \theta(0)=0 \text{ and } \lim_{t\to\infty}\frac{|\theta(t)|}{1+t}=0\right\};$$

and let Z: $[0,\infty)\times\mathbb{R}^{N+1}\times\Theta\to\mathbb{R}^{N+1}$ be a measurable map such that: i) for each $z\in\mathbb{R}^{N+1}$, $(t,\theta)\to Z(t,z,\theta)=\begin{bmatrix}X(t,z,\theta)\\Y(t,z,\theta)\end{bmatrix}$ is a right-continuous progressively measurable solution to

$$Z(T,x)=z+\sum_{k=1}^{d+1}\int_0^T W_k(Z(t,x))\circ d\theta_k(t)+\int_0^T W_0(Z(t,x))\,dt \qquad T\geqq 0;$$

and ii) $(t,z)\to Z(t,z,\theta)$ is an element of $C^{0,\infty}([0,\infty)\times\mathbb{R}^{N+1};\mathbb{R}^{N+1})$ for $\mathscr{W}$-almost every θ. Set

$$R(t,z,\theta)=\int_0^t \rho(Y(s,z,\theta))\,ds$$

and observe that if $z=\begin{bmatrix}x\\y\end{bmatrix}$ then

$$Q^c(t,x,\Gamma)=E^{\mathscr{W}}\left[\exp\left(\int_0^t c(X(R^{-1}(s,z),z)\,ds\right), X(R^{-1}(t,z),z)\in\Gamma\right].$$

Hence,

$$Q^c_\psi(x,\Gamma)=E^{\mathscr{W}}\left[\int_0^\infty \psi(R(t,z))\,\rho(Y(t,z))\exp\left(\int_0^t \tilde{c}(Z(s,z))\,ds\right), X(t,z)\in\Gamma\right]$$

where $\tilde{c}(\zeta)=\rho(\eta)\,c(\xi)$ for $\zeta=\begin{bmatrix}\xi\\\eta\end{bmatrix}\in\mathbb{R}^N\times\mathbb{R}^1$.

Notice that, because of (1.1),

$$\sum_{k=1}^{d+1}(W_k(z),\eta)^2_{\mathbb{R}^{N+1}}+\sum_{2\leqq|\alpha|\leqq l\vee 4}(W_\alpha(z),\eta)^2_{\mathbb{R}^{N+1}}\geqq\varepsilon|\eta|^2, \qquad z\in\mathbb{R}^{N+1} \text{ and } \eta\in S^{N+1}.$$

Hence (cf. [17], (3.4)), for each $p\in[1,\infty)$, the Malliavin covariance matrix

$$A(t,z)=\langle\!\langle Z(t,z),Z(t,z)\rangle\!\rangle$$

of $Z(t,z)$ satisfies

$$\|1/\det A(t,z)\|_p\leqq A_p/t^\nu, \qquad (t,z)\in(0,1]\times\mathbb{R}^{N+1},$$

where $A_p<\infty$ and $\nu\in(0,\infty)$ depend only on the V_k's. Moreover, using (2.6) in [17], one sees that if $\Delta_p(z)=\|1/\det(A(1,z))\|_{L^p(\mathscr{W})}$ then

$$\|1/\det(A(t+1,z))\|_{L^p(\mathscr{W})}\leqq E^{\mathscr{W}}[\Delta_p(Z(t,z))^p]^{1/p}\leqq A_p.$$

Thus,

$$\|1/\det(A(t,z))\|_{L^p(\mathscr{W})} \leqq A_p/(t \wedge 1)^\nu, \qquad (t,z)\in(0,\infty)\times\mathbb{R}^{N+1}.$$

Applying the integration by parts formula in Theorem (1.20) of [16], we conclude that for every $\beta\in\mathscr{N}^N$ there is a measurable Ψ_β: $(0,\infty)\times\mathbb{R}^{N+1}\times\Theta\to\mathbb{R}^1$ such that for all $\phi\in C_0^\infty(\mathbb{R}^N)$ and $(t,z)\in(0,\infty)\times\mathbb{R}^{N+1}$

$$E^{\mathscr{W}}\left[\psi(R(t,z))\,\rho(Y(t,z))\exp\left(\int_0^t \tilde{c}(Z(s,z))\,ds\right)(\partial^\beta\phi)(X(t,z))\right]$$
$$=E^{\mathscr{W}}[\Psi_\beta(t,z))\,\phi(X(t,z))]$$

and

$$\|\Psi_\beta(t,z)\|_{L^p(\mathscr{W})}\leqq\frac{M_\beta e^{M_\beta t}}{(t\wedge 1)^{M_\beta}}\sum_{m=0}^{|\beta|}E^{\mathscr{W}}[|\psi^{(m)}(R(t,z))|^4]^{1/4}$$

for some $M_\beta<\infty$ depending only on the V_k's and c. Combining the preceding with standard estimates for $\mathscr{W}(X(t,z)\notin B(x,r))$, one quickly arrives at the required estimates in (1.3). □

(1.4) **Corollary.** *Given an open set G in $\mathbb{R}^N$, set $\tau=\inf\{t\geqq 0\colon x(t)\notin G\}$ and define*

$$\hat{Q}_\psi^c(t,x,\Gamma)=E^{Q_x}\left[\int_0^\tau \psi(t)\exp)\int_0^t c(x(s))\,ds\right)\chi_\Gamma(x(t))\,dt\right]$$

for $(t,x)\in[0,\infty)\times\mathbb{R}^N$ and $\Gamma\in\mathscr{B}[G]$. Then $\hat{Q}_\psi^c(x,dy)=\hat{q}_\psi^c(x,y)\,dy$ where $\hat{q}_\psi^c\in C^\infty(G\setminus\{x\})$. In fact, for each $n\geqq 0$ there exists an $M_n<\infty$, which depends only on the V_k's and c, such that

$$\|q_\psi^c(x,\cdot)-\hat{q}_\psi^c(x,\cdot)\|_{C^n(G(r))}\leqq\frac{M_n e^{M_n T}}{r^{M_n}}\|\psi\|_{C^n([0,\infty))},$$

for $x\in G$ and $0<r<\operatorname{diam}(G)/2$, where $T=\sup\{t\geqq 0\colon \psi(t)\neq 0\}$ and $G(r)=\{y\in G\colon \operatorname{dist}(y,G^c)>r\}$.

Proof. Note that

$$Q_\psi^c(x,\Gamma)-\hat{Q}_\psi^c(x,\Gamma)=E^{Q_x}\left[\exp\left(\int_0^\tau c(x(s))ds\right)Q_{\psi_\tau}^c(x(\tau),\Gamma),\tau\leqq T\right],$$

where $\psi_\tau(\cdot)=\psi(\tau+\cdot)$. Thus

$$q_\psi^c(x,y)-\hat{q}_\psi^c(x,y)=E^{Q_x}\left[\exp\left(\int_0^\tau c(x(s))\,ds\right)q_{\psi_\tau}^c(x(\tau),y),\tau\leqq T\right].$$

The required estimate follows from this and Theorem 1.2 above. □

(1.5) **Theorem.** *Let M be a compact N-dimensional manifold and let m be a smooth probability measure on M (i.e. at every point there is a coordinate chart in which $m(dz)=\mu(x)\,dx_1\wedge\ldots\wedge dx_N$ where μ is a smooth positive function). Sup-*

pose that $\{X_0, \ldots, X_d\} \subseteq \Gamma(\mathbf{T}M)$ (*the smooth vector fields on* M) *has the property that* $\mathrm{Lie}(X_0, \ldots, X_d)(x) = \mathbf{T}_x M$ *for all* $x \in M$. *Given an* $x \in M$, *denote by* P_x *the solution on* $\Omega = \Omega_M$ *to the martingale problem for* $L = 1/2 \sum_1^d X_k^2 + X_0$ *starting at* x *and define* $P(t, x, \Gamma) = P_x(x(t) \in \Gamma)$. *Given* $\psi \in C_0^\infty((0, \infty))$, *define*

$$P_\psi(x, \Gamma) = \int_0^\infty \psi(t)\, P(t, x, \Gamma)\, dt, \qquad x \in M \quad \text{and} \quad \Gamma \in \mathscr{B}_M.$$

Then $P_\psi(x, dy) = p_\psi(x, y)\, m(dy)$ *where* $p_\psi \in C^\infty(M \times M)$.

Proof. Let L^* denote the formal adjoint of L with respect to m. Then $L^* = 1/2 \sum_1^d X_k^2 + Y + c$ where $Y = -X_0 + \sum_1^d b_k X_k$ for some choice of b_k's and c from $C^\infty(M)$. In particular, $\mathrm{Lie}(Y, X_1, \ldots, X_d) = \mathbf{T}M$. Moreover, if P_y^* on Ω is the solution to the martingale problem for $1/2 \sum_1^d X_k^2 + Y$ starting from $y \in M$ and

$$P_\psi^*(y, \Gamma) = E^{P_y^*}\left[\int_0^\infty \psi(t) \exp\left(\int_0^t c(x(s))\, ds\right) \chi_\Gamma(x(t))\, dt\right],$$

then for all ϕ_1 and $\phi_2 \in C^\infty(M)$

$$\int \phi_1(x)\left(\int \phi_2(y)\, P_\psi(x, dy)\right) m(dx) = \int \phi_1(y)\left(\int \phi_1(x)\, P_\psi^*(y, dx)\right) m(dy).$$

With the preceding remarks in mind, one sees that it is enough to check that for any $c \in C^\infty(M)$ the measures

$$P_\psi^c(x, \Gamma) = E^{P_x}\left[\int_0^\infty \psi(t) \exp\left(\int_0^t c(x(s))\, ds\right) \chi_\Gamma(x(t))\, dt\right]$$

satisfy $P_\psi^c(x, dy) = p_\psi^c(x, y)\, dy$ with $p_\psi^c(x, \cdot) \in C^\infty(M)$ and $\sup_{x \in M} \|p_\psi^c(x, \cdot)\|_{C^n(M)} < \infty$ for each $n \geqq 0$.

To this end, let (W, h) be a coordinate patch for M and choose open U_1 and U_2 so that $\bar{U}_1 \Subset U_2$ and $\bar{U}_2 \Subset W$. Next, define $\sigma_0 = \inf\{t \geqq 0 \colon x(t) \in \bar{U}_1\}$ and use induction to define $\tau_n = \inf\{t \geqq \sigma_n \colon x(t) \notin U_2\}$ for $n \geqq 0$ and $\sigma_n = \inf\{t \geqq \tau_{n-1} \colon x(t) \in \bar{U}_1\}$ for $n \geqq 1$. Then, for $x \in M$ and $\Gamma \in \mathscr{B}_{U_1}$,

$$P_\psi^c(x, \Gamma) = \sum_{n=0}^\infty E^{P_x}\left[\exp\left(\int_0^{\sigma_n} c(x(s))\, ds\right) \hat{P}_{\psi_{\sigma_n}}^c(x(\sigma_n), \Gamma), \sigma_n \leqq T\right]$$

where $T = \sup\{t \geqq 0 \colon \psi(t) \neq 0\}$ and

$$\hat{P}_{\psi_s}^c(y, \Gamma) = E^{P_y}\left[\int_0^{\tau_0} \psi(t+s) \exp\left(\int_0^t c(x(s))\, ds\right) \chi_\Gamma(x(t))\, dt\right]$$

for $s \geqq 0$ and $y \in W$. Now choose $\{V_0, \ldots, V_d\} \subseteq C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ so that (1.1) holds and $h_* X_k = V_k$ on a neighborhood of $\bar{U}_2$. Then, for $y \in W$ and $\Gamma \in \mathscr{B}_{U_1}$

$$\hat{P}^c_{\psi_s}(y, \Gamma) = \hat{Q}^{c \circ h^{-1}}_{\psi_s}(h(y), h(\Gamma))$$

where $\hat{Q}^{c \circ h^{-1}}(y, \cdot)$ is defined relative to the V_k's. Hence, because there is an $M < \infty$ such that

$$P_x(\sigma_n \leqq T) \leqq M e^{-nM \operatorname{dist}(\bar{U}_1, (U_2)^c)}, \qquad n \geqq 0,$$

Corollary 1.4 allows us to reach the required conclusions. □

(1.6) **Corollary.** *Let M, m and $\{X_0, \ldots, X_d\}$ be as they were in Theorem 1.5. Given $u \in C([0, \infty); \mathbb{R}^d)$ and $x \in M$, denote by $\Phi(\cdot, x; u)$ the curve which satisfies $\Phi(0, x; u) = x$ and*

$$\Phi(t, x; u) = \sum_1^d u_k(t) X_k(\Phi(t, x; u)) + X_0(\Phi(t, x; u)), \qquad t \geqq 0.$$

Assume that, for each $x \in M$, $\{\Phi(t, x; u): t \geqq 0$ and $u \in C([0, \infty); \mathbb{R}^d)\}$ is dense in M. Then there is a $\psi \in C_0^\infty((0, \infty))^+$ with $\int_{(0,\infty)} \psi(t)\, dt = 1$ for which the corresponding p_ψ is a uniformly positive element of $C^\infty(M \times M)$. In particular, there is a unique probability measure $\mu = m_L$ on M such that μ is $P(t, x, \cdot)$-invariant (i.e. $\mu = \int_M P(t, x, \cdot)\, \mu(dx)$ for all $t > 0$). Moreover, $m_L(dx) = \mu_L(x)\, m(dx)$ where μ_L is a uniformly positive element of $C^\infty(M)$. Finally, if $f \in C^\infty(M)$ and $\int_M f\, dm_L = 0$, then there is a unique $u = u_f \in C^\infty(M)$ with the properties that $Lu = -f$ and $\int_M u\, dm_L = 0$. In fact, there is a $K < \infty$ such that $\|u_f\| \leqq K \|f\|$.

Proof. Let $\phi_1 \in C_0^\infty((0, \infty))^+$ satisfying $\int_{(0,\infty)} \psi_1(t) dt = 1$ be given. Then $p_{\psi_1} \in C^\infty(M \times M)^+$. Moreover,

$$\int p_{\psi_1}(x, y)\, m(dx) = E^{P_y^*}\left[\int_0^\infty \psi_1(t) \exp\left(\int_0^t c(x(s))\, ds\right) dt\right] \geqq e^{-T_1 \|c\|}$$

where $\{P_y^*: y \in M\}$ and c are as in the proof of Theorem (1.5) and $\mathbf{T}_1 = \sup\{t \geqq 0: \psi(t) \neq 0\}$. Hence, there is an $\varepsilon > 0$ and an $r > 0$ such that, for each $x \in M$, there is a $y \in M$ for which $p_{\psi_1}(x, \cdot) \geqq \varepsilon$ on $B(y, r)$. (Here we have used $B(y, R)$ to denote the ball of radius r with center y relative to some compatible metric on M.) Next note that, under the stated hypotheses, for every pair x and y from M, there is a $T > 0$ such that $\int_0^T P(t, x, B(y, r/2))\, dt > 0$. Hence, if $n \geqq 1$ and $y_1, \ldots, y_n \in M$ are chosen so that $M = \bigcup_1^n B(y_i, r/2)$, then for each $x \in M$ there is a $T(x)>0$ for which $\min_{1\leqq i\leqq n} \int_0^{T(x)} P(t, x, B(y_i, r/2))\,dt>0$. Since $x\mapsto \min_{1\leqq i\leqq n} \int_0^T P(t, x, B(y_i, r/2))\,dt$ is lower semi-continuous for each $T>0$, it follows that there is a $T>0$ and an $\eta>0$ such that $\min_{1\leqq i\leqq n} \int_0^T P(t, x, B(y_i, r/2))\,dt\geqq\eta$ for all $x\in M$. But this means that $\int_0^T P(t, x, B(y, r))\,dt\geqq\eta$ for all $x, y\in M$, from which it follows that $\int_1^{T+1} P(t, x, B(y, r))\,dt\geqq\eta$. Now choose $\psi_2\in C_0^\infty((0, \infty))^+$ with $\int_{(0,\infty)} \psi_2(t)\,dt=1$ so that $\psi_2(t)\geqq 1/2T$ for $t\in[1, T+1]$. Then $P_{\psi_2}(x, B(y, r))\geqq\eta/2T$ for all $x, y\in M$. Finally, set $\psi=\psi_1*\psi_2$, and note that $p_\psi(x, y)=\int p_{\psi_1}(\xi, y)\,P_{\psi_2}(x, d\xi)\geqq\varepsilon\eta/2T$ for all $x, y\in M$.

Turning to the existence and uniqueness of a $P(t, x, \cdot)$-invariant μ, first observe that (because M is compact) existence is automatic. To prove the uniqueness, note that if μ is $P(t, x, \cdot)$-invariant, then $\mu=\int P_\psi(x, \cdot)\,\mu(dx)$ and so

$$\mu(dy)=\left(\int p_\psi(x, y)\,\mu(dx)\right) m(dy).$$

From this the existence of the uniformly positive density for μ is immediate; and therefore it is also clear that only one such μ can exist.

To prove the existence and asserted properties of u_f for $f\in C^\infty(M)$ with $\int f\,dm_L=0$, set $\tilde{M}=S^1\times M$ and consider the operator $\tilde{L}=1/2\dfrac{\partial^2}{\partial\theta^2}+\rho(\theta)\,L$, where $\rho(\theta)=2+\sin(\theta)$, on $C^\infty(\tilde{M})$. It is then easy to check that $\tilde{L}=\sum_{k=1}^{d+1} \tilde{X}_k^2+\tilde{X}_0$ where the $\tilde{X}_k$'s are vector fields on $\tilde{M}$ for which the hypotheses of the present theorem hold. Let $\tilde{P}(t, (\theta, x), \cdot)$ be the transition probability function for the diffusion on $\tilde{M}$ determined by $\tilde{L}$ and define $\tilde{m}_{\tilde{L}}(d\theta\times dx)=\lambda(d\theta)\times m_L(dx)$ where λ denotes the normalized rotationally invariant measure on S^1. A simple computation shows that $\int_{\tilde{M}} \tilde{L}\tilde{f}\,d\tilde{m}_{\tilde{L}}=0$ for all $\tilde{f}\in C^\infty(\tilde{M})$; and from this it is a relatively easy matter to conclude that $\tilde{m}_{\tilde{L}}$ is the unique $\tilde{P}(t, (\theta, x), \cdot)$-invariant probability measure on $\tilde{M}$. In addition, since $\mathrm{Lie}\left(\dfrac{\partial}{\partial t}+\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_{d+1}\right)(t, \theta, x)$ $=\mathbf{T}_{(t,\theta,x)}(\mathbb{R}^1\times\tilde{M})$ for each $(t, \theta, x)\in\mathbb{R}^1\times\tilde{M}$, Hörmander's theorem says that $\tilde{P}(t, (\theta, x), d\eta\times dy)=\tilde{p}(t, (\theta, x), (\eta, y))\,d\eta\times dy$ where $\tilde{p}\in C^\infty((0, \infty)\times\tilde{M}\times\tilde{M})$. Hence, by Doeblin's Theorem, $\|\tilde{P}(t, (\theta, x), \cdot)-\tilde{m}_{\tilde{L}}\|_{\mathrm{var}}\leqq Ae^{-\alpha t}$ for all $(t, \theta, x)\in(0, \infty)\times\tilde{M}$ and some $A<\infty$ and $\alpha\in(0, \infty)$. In particular, if

$$\tilde{G}\tilde{f}(\theta, x)=\int_0^\infty\left(\int_{\tilde{M}} \tilde{f}(\eta, y)(\tilde{P}(t, (\theta, x), d\eta\times dy)-\tilde{m}(d\eta, dy))\right)dt,$$

then $\|\tilde{G}\tilde{f}\|\leqq K\|\tilde{f}\|$ where $K=A/\alpha$. Also, it is easy to check that $\tilde{L}\tilde{G}\tilde{f}=-(\tilde{f}-\int\tilde{f}\,d\tilde{m}_{\tilde{L}})$ in the sense of distributions; and therefore, because $\tilde{L}$ is hypoelliptic,

$\tilde{G}\tilde{f}\in C^\infty(\tilde{M})$ and is a strong solution to this equation. At the same time, it is clear that if $\tilde{u}\in C^\infty(\tilde{M})$ and $\tilde{L}\tilde{u}=0$, then $\tilde{u}=\int\tilde{u}\,d\tilde{m}_{\tilde{L}}$.

Now suppose that $f\in C^\infty(M)$ with $\int f\,dm_L=0$ is given, and set $\tilde{f}(\theta, x)=\rho(\theta)f(x)$. Then $\tilde{u}=\tilde{G}\tilde{f}$ is the unique element of $C^\infty(\tilde{M})$ which satisfies $\tilde{L}\tilde{u}=-\tilde{f}$ and $\int\tilde{u}\,d\tilde{m}_{\tilde{L}}=0$. Hence the function

$$u=\frac{\int\limits_{S^1}\rho(\theta)\,\tilde{u}(\theta,\cdot)\,\lambda(d\theta)-\int\limits_{\tilde{M}}\rho(\theta)\,\tilde{u}(\theta,x)\,\tilde{m}_{\tilde{L}}(d\theta\times dx)}{\int\limits_{S^1}\rho(\theta)\,\lambda(d\theta)}$$

is an element of $C^\infty(M)$ which satisfies $Lu=-f$ and $\int\limits_M u\,dm_L=0$. Furthermore, by lifting to $\tilde{M}$, it is easy to see that u is uniquely determined by these properties; and clearly $\|u\|\leqq K\|f\|$. □

Let M and $\{P_x\colon x\in M\}$ be as in the preceding. The next result refers to the following quantities and hypotheses.

(1) $(B, \|\cdot\|_B)$ is a separable (real) Banach space and $\mathscr{C}\subseteq B$ is a convex set on which there is a complete metric ρ satisfying $\rho(Y_1, Y_2)\leqq\|Y_1-Y_2\|_B$ for all $Y_1, Y_2\in\mathscr{C}$ and for which ρ-balls are convex.

(2) $(E, \mathscr{F})$ is a measurable space and $\{\mathscr{F}_t\colon t\geqq 0\}$ is a non-decreasing family of sub-σ-algebras of $\mathscr{F}$.

(3) $X\colon[0,\infty)\times E\to M$ and $A\colon[0,\infty)\times E\to B$ are $\{\mathscr{F}_t\}$-progressively measurable functions with the properties that for $\xi\in E$: $A(0,\xi)=0$, $A(\cdot,\xi)$ and $X(\cdot,\xi)$ are continuous and $A(t,\xi)/t\in\mathscr{C}$ for $t\in(0,\infty)$.

(4) $\{R_x\colon x\in M\}$ is a measurable family of probability measures on $(E, \mathscr{F})$ such that for every $x\in M$: $P_x=R_x\circ X(\cdot)^{-1}$ and $R_x(A(t+s)-A(s)\in\Delta\,|\,\mathscr{F}_s)=R_{X(s)}(A(t)\in\Delta)$ for all $s, t\in(0,\infty)$ and $\Delta\in\mathscr{B}_B$.

(5) For every $\delta>0$ and $T>0$:

$$\lim_{t\to\infty}\frac{1}{t}\log\sup_{x\in M, s\in[0,T]}R_x(\|A(s)\|_B/t\geqq\delta)=-\infty.$$

(6) For every $L>0$ there is a ρ-compact subset K_L in $\mathscr{C}$ such that

$$\overline{\lim_{t\to\infty}}\frac{1}{t}\log\sup_{x\in M}R_x(A(t)/t\notin K_L)\leqq -L.$$

(1.7) **Theorem.** *Under the hypotheses in Corollary* 1.6, *there is a unique function* $I\colon B\to[0,\infty)\cup\{\infty\}$ *with the properties that*

(1) *I is a ρ-lower semicontinuous function whose level sets $\{Y\in B\colon I(Y)\leqq L\}$, $L\in[0,\infty)$ are ρ-compact subsets of $\mathscr{C}$;*

(2) *for every ρ-closed set $\Gamma\subseteq\mathscr{C}$,*

$$\overline{\lim_{t\to\infty}}\frac{1}{t}\log\sup_{x\in M}R_x(A(t)/t\in\Gamma)\leqq -\inf_\Gamma I;$$

(3) *for every ρ-open set $\Gamma \subseteq \mathscr{C}$*

$$\underline{\lim}_{t\to\infty} \frac{1}{t} \log \inf_{x\in M} R_x(A(t)/t\in\Gamma) \geqq -\inf_{\Gamma} I.$$

In particular, if $\Phi: \mathscr{C} \to \mathbb{R}^1$ is continuous and

$$\sup_{t\geqq 1} \sup_{x\in M} (E^{R_x}[\exp[(1+\varepsilon)\, t\Phi(A(t)/t)]])^{1/t} < \infty$$

for some $\varepsilon > 0$, then

$$\lim_{t\to\infty} \sup_{x\in M} \left| \frac{1}{t} \log E^{R_x}[\exp[t\Phi(A(t)/t)]] - \sup_{C\in\mathscr{C}} (\Phi(C) - I(C)) \right| = 0. \tag{1.8}$$

Proof. Following the scheme used in Sect. 6 of [20] (cf. in particular, the proof of Theorem 6.9 on p. 128), set

$$l(Y,\delta) = -\overline{\lim_{t\to\infty}} \frac{1}{t} \log \left[\inf_{x\in M} R_x\left(\frac{A(t)}{t} \in kB_\rho(Y,\delta)\right)\right], \qquad Y\in\mathscr{C} \quad \text{and} \quad \delta>0,$$

and

$$I(Y) = \sup\{l(Y,\delta): \delta > 0\}, \qquad Y\in\mathscr{C},$$

where $B_\rho(Y, \delta)$ is the ρ-ball in $\mathscr{C}$ of radius δ with center at Y. The derivation of (1), (2), and (3) with this choice of I differs negligibly from the proofs of Theorem 6.9 [20, p. 128] and [21] Theorem 1.7 , p. 843] once it is shown that for every $\delta > 0$ there exist a $K\in[0, \infty)$ and a $\beta_\delta: (0, \infty) \to (0, \infty)$ such that

$$\lim_{t\to\infty} \frac{1}{t} \log(\beta_\delta(t)) = -\infty$$

and, for every measurable $\Gamma \subseteq \mathscr{C}$,

$$R_x\left(\frac{A(t)}{t} \in \Gamma\right) \leqq K R_y\left(\rho\left(\frac{A(t)}{t}, \Gamma\right) < \delta\right) + \beta_\delta(t), \qquad (t,x,y)\in[1,\infty)\times M\times M. \tag{1.9}$$

To prove (1.9), choose ψ as in Corollary (1.6) and define p_ψ accordingly. Then there is a $C < \infty$ such that $p_\psi(x, \cdot) \leqq C p_\psi(y, \cdot)$ for all $x, y\in M$. Set $T = \sup\{t \geqq 0: \psi(t) \neq 0\}$ and, for given $t\in[0. \infty)$ and $\delta > 0$, define

$$\alpha(t,\delta) = \sup_{x\in M} \sup_{s\in[0,T]} R_x\left(\frac{\|A(s)\|_B}{t} \geqq \delta\right).$$

Then, for $(s, x)\in[0, T]\times M$:

$$R_x(A(t)/t\in\Gamma)\leqq R_x\left(\rho\left(\frac{A(t+s)-A(s)}{t}, \Gamma\right)<\delta/2\right)$$
$$+R_x\left(\left\|\frac{A(s)}{t}\right\|_B\geqq\delta/4\right)+R_x\left(\left\|\frac{A(t+s)-A(t)}{t}\right\|_B\geqq\delta/4\right)$$
$$\leqq R_x\left(\rho\left(\frac{A(t+s)-A(t)}{t}, \Gamma\right)<\delta/2\right)+2\alpha(t, \delta/4).$$

Recalling that $\rho(Y_1, Y_2)\leqq\|Y_1-Y_2\|_B$, we see that, for any $(t, x, y)\in(0, T]\times M\times M$:

$$\begin{aligned}R_x(A(t)/t\in\Gamma)&\leqq\int_{(0,\infty)}\psi(s)\,E^{R_x}[R_{X(s)}(\rho(A(t)/t, \Gamma)<\delta/2)]\,ds+2\alpha(t, \delta/4)\\&=\int p_\psi(x, \eta)\,R_\eta(\rho(A(t)/t, \Gamma)<\delta/2)\,m(d\eta)+2\alpha(t, \delta/4)\\&\leqq C\int p_\psi(y, \eta)\,R_\eta(\rho(A(t)/t, \Gamma)<\delta/2)\,m(d\eta)+2\alpha(t, \delta/4)\\&\leqq C\int_{(0,\infty)}\psi(s)\,E^{R_y}[R_{X(s)}(\rho(A(t)/t, \Gamma)<\delta/2)]\,ds+2\alpha(t, \delta/4).\end{aligned}$$

At the same time

$$\begin{aligned}&\int_{(0,\infty)}\psi(s)\,E^{R_y}(R_{X(s)}(\rho(A(t)/t, \Gamma)<\delta/2)]\,ds\\&\quad=\int_{(0,\infty)}\psi(s)\,R_y\left(\rho\left(\frac{A(t+s)-A(s)}{t}, \Gamma\right)<\delta/2\right)ds\\&\quad\leqq R_y(\rho(A(t)/t, \Gamma)<\delta)+2\alpha(t, \delta/4).\end{aligned}$$

From these, we easily pass to (1.9).

As we said, the derivation of the desired conclusions once (1.9) has been established is essentially the same as the argument used in the cited references. However, the following comments may be helpful to the reader who wants to fill in the details. In the first place, the norm $\|\cdot\|_B$ plays no further role; it is only used in the derivation of (1.9). Secondly, it may be helpful to note that the positivity of p_ψ is the key to our ability to avoid the hypothesis made in (6.8) of [20] about the transition probability function. Thirdly, (1.9) plays here the role which the estimate Lemma 2.2 had in [21]. For instance, it is what allows us to replace the "$\overline{\lim}$" by "lim" in the definition of $l(Y, \delta)$ so long as $I(Y)<\infty$ (cf. [21, Lemma 2.2] and [20, Lemma 6.10]). Finally, given (1), (2), and (3), (1.8) is completely standard when Φ is bounded; and the extension required to cover the assumption made here is not difficult. □

In the next corollary, we again refer to the situation described just before the statement of Theorem 1.7. In view of the results obtained in Theorem 1.7, this corollary is nothing more than an application of the inversion formula for the Legendre transform (cf. Theorem (7.15) on p. 135 of [20]).

(1.10) **Corollary.** *Assume that there is a locally convex Hausdorf topology on B whose restriction to $\mathscr{C}$ is the same as the topology determined by the metric ρ, and denote by $\hat{B}$ the dual of B relative to this topology. If, for all $\lambda>0$,*

$$\sup_{t\geqq 1}\sup_{x\in M}(E^{R_x}[\exp[\lambda\|A(t)\|_B]])^{1/t}<\infty,$$

then, for each $\hat{Y}\in\hat{B}$

$$\Lambda(\hat{Y})=\lim_{t\to\infty}\frac{1}{t}\log E^{R_x}[\exp(\langle\hat{Y},A(t)\rangle)]$$

exists uniformly with respect to $x\in M$ and is independent of $x\in M$. Furthermore, the function I in Theorem 1.8 is given by

$$I(Y)=\sup_{\hat{Y}\in\hat{B}}(\langle\hat{Y},Y\rangle-\Lambda(\hat{Y})).$$

We now want to apply these considerations of two specific cases.

(1.11) **Corollary.** *Referring to Corollary 1.6, set $\mathbf{T}_0=\delta_{x(0)}$ and $\mathbf{T}_t=\frac{1}{t}\int_0^t\delta_{x(s)}\,ds$ for $t>0$. Define the function J: $C(M)^*\to[0,\infty)\cup\{\infty\}$ by $J(\mu)=\infty$ if μ is not a probability measure and*

$$J(\mu)=\sup\left\{-\int\frac{Lu}{u+1}\,d\mu:\quad u\in C^\infty(M)^+\right\},$$

if μ is a probability measure. Then for every weakly closed set Γ of $C(M)^*$,

$$\overline{\lim_{t\to\infty}}\log\inf_{x\in M}P_x(\mathbf{T}_t\in\Gamma)\leqq-\inf_\Gamma J;$$

and, for every weakly open set Γ in $C(M)^*$,

$$\underline{\lim_{t\to\infty}}\log\inf_{x\in M}P_x(\mathbf{T}_t\in\Gamma)\geqq-\inf_\Gamma J.$$

In particular, if Φ: $\mathbf{M}_1(M)\to\mathbb{R}^1$ is weakly continuous, then

$$\lim_{t\to\infty}\frac{1}{t}\sup_{x\in M}|\log E^{R_x}[e^{t\Phi(\mathbf{T}_t)}]-\sup_{\mu\in C(M)^*}(\Phi(\mu)-J(\mu))|=0.$$

Proof. We apply Theorem 1.7 and Corollary 1.10 with: $B=C(M)^*$ with the weak* topology, $\hat{B}=C(M)$, $\mathscr{C}=\{\mu\in C(M)^*$: μ is probability measure$\}$, ρ the Lèvy metric on $\mathscr{C}$, $E=\Omega_M$, $X(t)=x(t)$, $A(t)=t\mathbf{T}_t$, and $R_x=P_x$. In checking the hypotheses of Theorem 1.7, take $\|\cdot\|_B$ to be the total variation norm, and note that $\|A(t)\|_B=t$. Also, observe that, since $\mathbf{M}_1(M)$ is compact in the Lèvy metric,

(6) is trivial in this context. For more details, and, in particular, for the identification of

$$\sup_{V\in C(M)} \left(\int V d\mu - \Lambda(V)\right)$$

as $J(\mu_\cdot)$, see Sect. 7 of [20]. □

In our second application we will be working with the following situation.

(1) $\hat{M} = M \times \mathbb{R}^\nu$ for some $\nu \in Z^+$ and $\pi_1\colon \hat{M} \to M$ and $\pi_2 \to \mathbb{R}^\nu$ are the natural projection maps.

(2) $\{\hat{X}_0, \ldots, \hat{X}_d\} \subseteq \Gamma(T\hat{M})$ have the property that $\hat{X}_k(f \circ \pi_1) = (X_k f) \circ \pi_1$ for $f \in C^\infty(M)$ and $\hat{X}_k(f \circ \pi_2) = (Y_k f) \circ \pi_2$ for $f \in C^\infty(\mathbb{R}^\nu)$, where $Y_k = \sum_{j=1}^{\nu} Y_k^j(x) \frac{\partial}{\partial y^j}$ with $Y_k^j \in C^\infty(M)$, $0 \leqq k \leqq d$ and $1 \leqq j \leqq \nu$.

(3) For $(x, \eta) \in M \times \mathbb{R}^\nu$, $\sigma_k(x, \eta) = (Y_k(x), \eta)_{\mathbb{R}^\nu}$, $0 \leqq k \leqq d$ and

$$Q(x, \eta) = \sigma_0(x, \eta) + 1/2 \sum_1^d (X_k(\sigma_k(\cdot, \eta))(x).$$

(4) For $x \in M$, $a(x) = \sum_1^d Y_k(x) \otimes Y_k(x)$ and $a(x, \eta) = (\eta, a(x)\eta)_{\mathbb{R}^\nu}$, $\eta \in \mathbb{R}^\nu$.

(5) For $\eta \in \mathbb{R}^\nu$, $L_\eta = L + \sum_1^d \sigma_k(\cdot, \eta) X_k$ on $C^\infty(M)$ and $J_\eta(\mu)$ is defined on $C(M)^*$

so that

$$J_\eta(\mu) = \sup\left\{-\int_M \frac{L_\eta u}{1+u} d\mu\colon\ u \in C^\infty(M)^+\right\}$$

if μ is a probability measure and $J_\eta(\mu) = \infty$ otherwise.

(6) $\hat{\Omega} = \Omega_{\hat{M}} = \Omega_M \times \Omega_{\mathbb{R}^\nu}$ and for each $\hat{\omega} \in \hat{\Omega}$, $\hat{x}(t, \hat{\omega}) = (x(z, \hat{\omega}), y(t, \hat{\omega})) \in M \times \mathbb{R}^\nu$ for $t \geqq 0$.

(7) $\hat{L} = \sum_1^d \hat{X}_k^2 + \hat{X}_0$ on $C^\infty(\hat{M})$; and, for each $x \in M$, $\hat{P}_x$ on $\hat{\Omega}$ denotes the solution to the martingale problem for $\hat{L}$ starting from $(x, 0) \in M \times \mathbb{R}^\nu$.

(1.12) **Corollary.** *Referring to the preceding and working under the hypotheses in Corollary* 1.6,

$$\Lambda(\eta) = \lim_{t\to\infty} \frac{1}{t} \log E^{\hat{P}_x}[\exp((\eta, y(t))_{\mathbb{R}^\nu})]$$

exists uniformly respect to $x \in M$ *and is independent of* x. *In addition,*

$$\Lambda(\eta) = \sup\left\{\int_M [Q(x, \eta) + \tfrac{1}{2} a(x, \eta)]\, \mu(dx) - J_\eta(\mu)\colon\ \mu \in C(M)^*\right\}. \tag{1.13}$$

Equivalently,

$$(1.14)\qquad \Lambda(\eta)=\inf_{\phi}\sup_{\mu}\left\{\int_M \left[Q(x,\eta)-L\phi(x)+1/2\sum_1^d (X_k\phi(x) - \sigma_k(x,\eta))^2\right]\mu(dx)\right\}$$

where $\phi\in C^\infty(M)$ *and* $\mu\in\mathscr{P}(M)$ *and* $\mathscr{P}(M)=\{\mu\in C(M)^*\colon \mu$ *is a probability measure*$\}$. *In particular,*

$$(1.15)\qquad \Lambda(\eta)\geqq\bar{Q}(\eta)+\alpha(\eta)$$

and

$$(1.16)\qquad \Lambda(\eta)\leqq\bar{Q}(\eta)+\beta(\eta),$$

where

$$\bar{Q}(\eta)=\int_M Q(x,\eta)\,m_L(dx),$$

$$\alpha(\eta)=\inf\left\{\tfrac{1}{2}\sum_1^d\int_M (X_k\phi(x)-\sigma_k(x,\eta))^2 m_L(dx)\colon\ \phi\in C^\infty(M)\right\},$$

$$\beta(\eta)=\sup\left\{\tfrac{1}{2}\int_M\sum_1^d (X_k h(\cdot,\eta)(x)-\sigma_k(x,\eta))^2\mu(dx)\colon\ J_\eta(\mu)<\infty\right\},$$

and we have used $h(\cdot,\eta)$ *to denote the unique* $u\in C^\infty(M)$ *which satisfies* $Lu=Q(\cdot,\eta)-\bar{Q}(\eta)$ *and* $\int_M u\,dm_L=0$. *Finally, set*

$$I(y)=\sup\{(y,\eta)_{\mathbb{R}^\nu}-\Lambda(\eta)\colon\ \eta\in\mathbb{R}^\nu\},\qquad y\in\mathbb{R}^\nu.$$

Then an equivalent expression for $I(y)$ *is*

$$(1.17)\qquad I(y)=\sup_{\phi}\inf_{\mu}\sup_{\eta}\left\{\frac{\left((y,\eta)_{\mathbb{R}^\nu}-\int_M [Q(x,\eta)-L\phi(x)]\,\mu(dx)\right)^2}{2\sum_1^d\int_M (X_k\phi(x)-\sigma_k(x,\eta))^2\mu(dx)}\right\}$$

$\left(\frac{c^2}{0}\right.$ *is* 0 *or* ∞ *according to whether* $c=0$ *or not*$\Big)$ *where* $\phi\in C^\infty(M)$, $\mu\in\mathscr{P}(M)$, *and* $\eta\in S^{\nu-1}$; *and for every* $\Gamma\in\mathscr{B}_{\mathbb{R}^\nu}$

$$-\inf_{\Gamma^0} I\leqq\varliminf_{t\to\infty}\frac{1}{t}\log\inf_{x\in M}\hat{P}_x(y(t)/t\in\Gamma)$$

$$\leqq\varlimsup_{t\to\infty}\frac{1}{t}\log\sup_{x\in M}\hat{P}_x(y(t)/t\in\Gamma)\leqq-\inf_{\bar{\Gamma}} I,$$

and, for $\Phi \in C(\mathbb{R}^\nu)$ *which satisfies* $\lim_{|y|\to\infty} \frac{\Phi(y)}{|y|^2} = 0$

$$\lim_{t\to\infty} \sup_{x\in M} \left| \frac{1}{t} \log E^{\hat{P}_x}[\exp(t\Phi(y(t)/t))] - \sup_{y\in\mathbb{R}^\nu} (\Phi(y) - I(y)) \right| = 0.$$

Proof. The proof involves the use of Corollary 1.11 to evaluate $\Lambda(\eta)$, followed by an application of Theorem 1.7 to the $\mathbb{R}^\nu$-valued additive functional $\{y(t): t \geqq 0\}$.

Define $\bar{y}(t, \eta) = (y(t), \eta)_{\mathbb{R}^\nu} - \int_0^t Q(x(s), \eta)\, ds$ and set

$$R(t, \eta) = \exp\left(\bar{y}(t, \eta) - 1/2 \int_0^t a(x(s), \eta)\, ds\right).$$

Then $R(\cdot, \eta)$ is a non-negative $\hat{P}_x$ martingale and the measure P_x^η on Ω_M defined by

$$P_x^\eta(A) = E^{\hat{P}_x}[R(t, \eta), x(t) \in A]$$

for $A \in \mathscr{B}_t \equiv \sigma(x(s): 0 \leqq s \leqq t)$ and $t \geqq 0$ is the solution to the martingale problem for L_η starting from x. Hence,

$$E^{\hat{P}_x}[\exp((y(t, \eta)_{\mathbb{R}^\nu})] = E^{P_x^\eta}\left[\exp\left(\int_0^t [Q(x(s), \eta) + \tfrac{1}{2} a(x(s), \eta)]\, ds\right)\right].$$

Noting that $\mathrm{Lie}\left(X_0 + \frac{1}{2}\sum_1^d \sigma_k(\cdot, \eta) X_k, X_1, \ldots, X_d\right)(x) = \mathrm{Lie}(X_0, \ldots, X_d)(x)$ at each $x \in M$ and applying Corollary 1.11 to $\{P_x^\eta: x \in M\}$, we conclude that (1.13) holds. The passage from (1.13) to (1.14), and thence to (1.15) and (1.16), is accomplished by the same sort of reasoning as was used in [21, Theorem 1.7]. Similarly, the second expression for $I(y)$ is an easy consequence of (1.14). Of course, here we are taking $\mathscr{C} = B = \mathbb{R}^\nu$ and we are justifying the use of Theorem 1.7 on the basis of the standard estimate

$$\hat{P}_x(|(y(t), \eta)_{\mathbb{R}^\nu}| \geqq \delta) \leqq 2\left[1 - \mathfrak{N}\left(\frac{\delta - \|Q\| t}{\|a\|\, t^{1/2}}\right)\right],$$

where $\mathfrak{N}$ denotes the normal distribution function and $\|Q\|$ and $\|a\|$ are the uniform norm of $Q(\cdot, \eta)$ and $a(\cdot, \eta)$, respectively. This same estimate provides the justification for the final assertion of this corollary. □

(1.18) *Remark.* Given $\eta \in S^{\nu-1}$, let $h(\cdot, \eta)$ be as in the statement of Corollary 1.12 above. Then, by (1.17):

$$I(y) \geqq \inf_\mu \sup_\eta \left\{ \frac{((y, \eta)_{\mathbb{R}^\nu} - \bar{Q}(\eta))^2}{2\sum_1^d \int_M (X_k h(x, \eta) - \sigma_k(x, \eta))^2 \mu(dx)} : \quad \mu \in \mathscr{P}(M) \text{ and } \mu \in S^{\nu-1} \right\}.$$

Hence, if $\bar{Q}\in\mathbb{R}^\nu$ is defined so that $(\bar{Q},\eta)_{\mathbb{R}^\nu}=\bar{Q}(\eta)$, $\eta\in\mathbb{R}^\nu$, then there is a $\gamma\in(0,\infty)$ such that

(1.19) $$I(y)\geqq\gamma|y-\bar{Q}|^2_{\mathbb{R}^\nu},\qquad y\in\mathbb{R}^\nu.$$

At the same time, by (1.17)

$$I(y)\leqq\sup\left\{\frac{(y-\bar{Q},\eta)^2_{\mathbb{R}^\nu}}{2\alpha(\eta)}:\ \eta\in S^{\nu-1}\right\}.$$

Thus, one always has

(1.20) $$I(\bar{Q})=0;$$

and, if

(1.21) $$\alpha\equiv\inf_{\eta\in S^{\nu-1}}\alpha(\eta)>0,$$

then

(1.22) $$I(y)\leqq|y-\bar{Q}|^2/2\alpha,\qquad y\in\mathbb{R}^\nu.$$

Finally, suppose not only that $\alpha=0$ but also that there is an $\eta\in S^{\nu-1}$ and an $f(\cdot,\eta)\in C^\infty(M)$ such that

(1.23) $$X_k f(\cdot,\eta)=\sigma_k(\cdot,\eta),\qquad 1\leqq k\leqq d.$$

Then, after substituting $\phi+f(\cdot,\eta)$ for ϕ in (1.17), one sees that $I(y)=\infty$ for $y\in\mathbb{R}^\nu$ such that

$$(y,\eta)_{\mathbb{R}^\nu}\notin\mathrm{Range}(Q(\cdot,\eta)-Lf(\cdot,\eta)).$$

Although the existence of such η and $f(\cdot,\eta)$ is, in general, not guaranteed by $\alpha=0$, it is when one has $\mathrm{Lie}(X_1,\dots,X_d)(x)=\mathbf{T}_xM$ at every $x\in M$ (cf. Lemma (2.26) in [21]). On the other hand, notice that no such η and $f(\cdot,\eta)$ can exist if $\mathrm{Lie}(\hat{X}_1,\dots,\hat{X}_d)(\hat{x})=\mathbf{T}_{\hat{x}}\hat{M}$ at even one $\hat{x}\in\hat{M}$. Indeed, if they exist and we define $\hat{\theta}=\pi_1^*\,d(f(\cdot,\eta))-\pi_2^*\,\eta\in T^*\hat{M}$, then $\hat{\theta}(\hat{X}_k)\equiv 0$, $1\leqq k\leqq d$. Moreover, $d\hat{\theta}=0$ and therefore $\hat{\theta}(\hat{X})\equiv 0$ for all $\hat{X}\in\mathrm{Lie}(\hat{X}_1,\dots,\hat{X}_d)$. Since $\hat{\theta}$ never vanishes, this shows that $\dim(\mathrm{Lie}(\hat{X}_1,\dots,\hat{X}_d)(\hat{x}))\leqq d+\nu-1$ at every $\hat{x}\in\hat{M}$ as soon as η and $f(\cdot,\eta)$ exist. In particular, we can see from these considerations that $\alpha>0$ if $\mathrm{Lie}(\hat{X}_1,\dots,\hat{X}_d)(\hat{x})=\mathbf{T}_{\hat{x}}\hat{M}$ for every $\hat{x}\in\hat{M}$.

We now see what, when they apply, the results obtained in the appendix say about the situation dealt with in Corollary 1.12. Thus, assume that $\{X_1,\dots,X_d\}$ satisfies (A.1) and that

(1.24) $$X_0=\sum_1^d\alpha_k X_k$$

for some $\alpha_1, \ldots, \alpha_d \in C^\infty(M)$. Then we can rewrite the operators L_η in the form

$$L_\eta = -\tfrac{1}{2}\sum_1^d X_k^* X_k + \sum_1^d (\beta_k + \sigma_k(\cdot, \eta)) X_k$$

for suitably chosen $\beta_1, \ldots, \beta_d \in C^\infty(M)$.

(1.25) **Theorem.** *Under the conditions just described, an equivalent expression for the quantity $\Lambda(\eta)$ in* (1.13) *is*

$$\Lambda(\eta) = \sup\{\int_M (\eta, \tilde{Y}_0(x))_{\mathbb{R}^\nu}\, \mu(dx) + \tfrac{1}{2} D(\mu, \eta) - \tilde{J}_0(\mu):\ \mu \in \mathscr{P}(M)\}, \tag{1.26}$$

where

$$\tilde{Y}_0 \equiv Y_0 - \sum_1^d \alpha_k Y_k$$

$$D(\mu, \eta) \equiv \inf\left\{\sum_1^d \int_M [(Y_k(x), \eta)_{\mathbb{R}^\nu} + \beta_k(x) - X_k \phi(x)]^2 \mu(dx):\ \phi \in C^\infty(M)\right\},$$

$$\tilde{J}_0(\mu) \equiv J_0(\mu) + \tfrac{1}{2} D(\mu, 0).$$

Hence,

$$\lim_{\rho \to \infty} \frac{1}{\rho^2} \Lambda(\rho\eta) = \frac{1}{2} \inf_{\phi \in C^\infty(M)} \left\|\sum_1^d (X_k \phi - (\eta, Y_k)_{\mathbb{R}^\nu})^2\right\|, \qquad \eta \in S^{\nu-1}, \tag{1.27}$$

and an equivalent expression for the quantity $I(y)$ in (1.17) *is*

$$I(y) = \inf_\mu \sup_\eta \left\{\frac{\left(\int_M (\eta, y - \tilde{Y}_0(x))_{\mathbb{R}^\nu}\, \mu(dx)\right)^2}{2D(\mu, \eta)} - \tilde{J}_0(\mu):\ \mu \in \mathscr{P}(M) \text{ and } \eta \in S^{\nu-1}\right\}. \tag{1.28}$$

Proof. The proof is simply a matter of reconciling the notation here with that used in the appendix. □

We conclude this section by summarizing another tack that can be taken in the analysis of the quantity $\Lambda(\eta)$ once one has the sort of regularity result provided by Theorem 1.5.

(1.29) **Theorem.** *We again make the hypotheses used in Corollary* 1.6. *For $\eta \in \mathbb{R}^\nu$ define the operator $\tilde{L}_\eta = L_\eta + Q(\cdot, \eta) + \frac{1}{2} a(\cdot, \eta)$.*

i) *If $f \in C^\infty(M)$ and $g(x, y) \equiv f(x) \cdot \exp((\eta, y)_{\mathbb{R}^\nu})$ for $(x, y) \in M \times \mathbb{R}^\nu$, then*

$$[\hat{L} g](x, y) = (\hat{L}_h f)(x) \cdot \exp((\eta, y)_{\mathbb{R}^\nu}).$$

ii) *$\Lambda(\eta)$ is the largest eigenvalue of $\hat{L}_\eta$, it is simple, and the corresponding eigenfunction can be chosen to be strictly positive.*

iii) *Λ is convex and analytic as a function of $\eta \in \mathbb{R}^\nu$.*
iv) $\bar{Q} = (\operatorname{grad} \Lambda)(0)$ *and*

$$\hat{P}_x\left(\lim_{t\to\infty} \frac{1}{t} y(t)\right) = \bar{Q}) = 1$$

for all $x \in \mathbb{R}^\nu$. Moreover, for each $x \in M$, the distribution of

$$\frac{y(t) - t\bar{Q}}{t^{1/2}}$$

under $\hat{P}_x$ converges to that of an $\mathbb{R}^\nu$-valued normal random variable with mean 0 *and covariance* $(D^2 \Lambda)(0)$ (*the Hessian of Λ at* 0).

Proof. The assertion in i) is a simple computation which, in fact, was already used in the derivation of (1.13). The methods used in Sect. 2 of [2] can be easily adapted to the present situation in order to prove ii), iii), and the strong law assertion in iv) (since, by Corollary 1.6, M is an "invariant control set"). Moreover, the identification of $\bar{Q}$ as grad $\Lambda(0)$ is an easy consequence of (1.15) and (1.16). Finally, the last part of iv) is an easy extension of Corollary (3.2) in [4].

(1.30) *Remark.* In [1] the derivation of the sort of large deviation principle stated at the end of Corollary 1.12 is based on the existence and regularity of $\Lambda(\cdot)$. In particular, iii) above provides more than enough regularity.

2. Applications to Stochastic Flows

Throughout M will denote a compact connected N-dimensional Riemannian manifold and

$$\{X_0, \ldots, X_d\} \subseteq \Gamma(\mathbf{T}M)$$

will be a fixed set of vector fields on M. In addition, $(\Theta, \mathscr{B}_\Theta, \mathscr{W})$ will denote the standard d-dimensional Wiener space (cf. the second paragraph of the proof of Theorem 1.2 and replace $d+1$ there with d here). Finally, $\{\xi_t\colon t \in [0, \infty)\}$ will be used to denote the stochastic flow of diffeomorphisms of M determined, up to a $\mathscr{W}$-null set, by

$$d\xi_t(x) = \sum_{k=1}^{d} X_k(\xi_t(x)) \circ d\theta_k(t) + X_0(\xi_t(x))\, dt \quad \text{with } \xi_0(x) = x \tag{2.1}$$

for $x \in M$. In particular, if P_x denotes the distribution of $\xi_\cdot(x)$ under $\mathscr{W}$, then $\{P_x\colon x \in M\}$ is the Markov family of solutions to the martingale problem for $L = \frac{1}{2} \sum_{k=1}^{d} X_k^2 + X_0$.

For $(t, x) \in [0, \infty) \times M$, let $D\xi_t(x)$: $\mathbf{T}_x M \to \mathbf{T}_{\xi_t(x)} M$ be the derivative of ξ_t at x. The purpose of this section is to study the behavior of the quantities $J_t(x) \equiv \det(D\xi_t(x))$ and $|D\xi_t(x)(v)|$, $v \in \mathbf{T}_x M$, as $t \to \infty$.

It follows from (2.1) that

$$\log J_t(x) = \sum_{k=1}^{d} \int_0^t (\operatorname{div} X_k)(\xi_s(x)) \circ d\theta_k(s) + \int_0^t (\operatorname{div} X_0)(\xi_s(x))\, ds; \tag{2.2}$$

and therefore that, when the hypotheses of Corollary 1.6 are satisfied by the X_k's, the results of Sect. 1 apply directly to the function

$$\Lambda(p) = \lim_{t \to \infty} \frac{1}{t} \log E^{\mathscr{W}}[(J_t(x))^p], \qquad p \in \mathbb{R}^1. \tag{2.3}$$

In order to see we can say about $|D\xi_t(x)(v)|$ it is important to express it in a form to which our results are applicable. For this reason, first think of $\begin{bmatrix} \xi_t \\ D\xi_t \end{bmatrix}$ as a stochastic map $T\xi_t$ from $\mathbf{T}M \to \mathbf{T}M$ (that is, if π: $\mathbf{T}M \to M$ denotes the natural projection of $\mathbf{T}M$ onto M and $v \in \mathbf{T}M$ with $\pi v = x$, then $T\xi_t(v) = \begin{bmatrix} \xi_t(x) \\ D\xi_t(x)(v) \end{bmatrix}$). In this way $\{T\xi_t \colon t \in [0, \infty)\}$ becomes a stochastic flow of diffeomorphisms of $\mathbf{T}M$; and this flow is determined, up to a $\mathscr{W}$-null set, by

$$d(T\xi_t(v)) = \sum_{k=1}^{d} TX_k(T\xi_t(v)) \circ d\theta_k(t) + TX_0(T\xi_t(v))\, dt \quad \text{with } T\xi_0(v) = v. \tag{2.4}$$

In (2.4), TX_k denotes the derivative of X_k thought of as a map on $\mathbf{T}M$. Thus, if we use the Levy-Civita connection ∇ for the Riemannian structure on M to determine the horizontal subspace in $\mathbf{T}_v \mathbf{T}M$ for $v \in \mathbf{T}M$ and we identify the horizontal subspace with $\mathbf{T}_{\pi v} M$ itself, then TX_k is the vector field on $\mathbf{T}M$ for which $X_k(x)$ and $\nabla X_k(x)(v)$ are, respectively the horizontal and vertical components of $TX_k(v)$, $v \in \mathbf{T}_x M$. In particular, if, for $v \in \mathbf{T}_x M$, we set $v_t = T\xi_t(v)$ and $x_t = \pi v_t$, then $x.$ is the path $\xi.(x)$ determined (up to a $\mathscr{W}$-null set) by (2.1), and $v.$ is path in $\mathbf{T}M$ over $x.$ which is determined (up to a $\mathscr{W}$-null set) by

$$\nabla v_t = \sum_{k=1}^{d} \nabla X_k(x_t)(v_t) \circ d\theta_k(t) + \nabla X_0(x_t)(v_t)\, dt \quad \text{with } v_0 = v. \tag{2.5}$$

In other words, (2.4) is equivalent to the conjunction of (2.1) and (2.5).

Because TX_k is linear on each tangent space $\mathbf{T}_x M$, we can use (2.5) to write an autonomous equation for the stochastic process

$$\tilde{\xi}.(v) \equiv \frac{T\xi.(v)}{|T\xi.(v)|}, \qquad \psi \in \mathbf{S}M,$$

on the sphere bundle $\mathbf{S}M \equiv \{v \in \mathbf{T}M: |v|=1\}$. Namely, if $\tilde{X}_k(v)$ denotes the projection of $TX_k(v)$ onto $\mathbf{T}_v \mathbf{S}M$, then

$$(2.6) \qquad d\tilde{\xi}_t(v) = \sum_{k=1}^{d} \tilde{X}_k(\tilde{\xi}_t(v)) \circ d\theta_k(t) + \tilde{X}_0(\tilde{\xi}_t(v))\, dt \quad \text{with } \tilde{\xi}_0(v) = v,$$

for $v \in \mathbf{S}M$. Note that the horizontal component of $\tilde{X}_k(v)$ is the same as that of $TX_k(v)$ (namely, $X_k(x)$ if $\pi v = x$) and that its vertical component is $\nabla X_k(x)(v) - \langle \nabla X_k(x)(v), v\rangle v$ for $v \in \mathbf{S}_x M$. Next, using (2.5) again, we see that

$$(2.7) \qquad \log|D\xi_t(x)(v)| = \sum_{k=1}^{d} \int_0^t g_k(\tilde{\xi}_s(v)) \circ d\theta_k(s) + \int_0^t g_0(\tilde{\xi}_s(v))\, ds,$$

where $g_k \in C^\infty(\mathbf{S}M)$ is given by

$$g_k(v) = \langle \nabla X_k(\pi v)(v), v\rangle, \qquad v \in \mathbf{S}M.$$

Henceforth, without mentioning it again, we will be assuming that

$$(2.8) \qquad \begin{aligned} &\mathrm{Lie}(\tilde{X}_0, \ldots, \tilde{X}_d)(v) = \mathbf{T}_v \mathbf{S}M, \qquad v \in \mathbf{S}M, \\ &\{\Psi(t, v; u):\ t \geqq 0 \ \text{ and } \ u \in C([0, \infty); \mathbb{R}^d)\} \text{ is dense in } \mathbf{S}M, \qquad v \in \mathbf{S}M, \end{aligned}$$

where, in the second part of (2.8), we have used $\Psi(\cdot, v; u)$ to denote the curve satisfying

$$\dot{\Psi}(t, v; u) = \sum_{k=1}^{d} u_k(t)\, \tilde{X}_k(\Psi(t, v; u)) + \tilde{X}_0(\Psi(t, v; u)) \quad \text{with } \Psi(0, v; u) = v.$$

Notice that these are precisely the hypotheses required to apply Corollary 1.6 to the vector fields $\{\tilde{X}_0, \ldots, \tilde{X}_d\}$ on $\mathbf{S}M$. Also, because $[\tilde{X}_k, \tilde{X}_l] = \widetilde{[X_k, X_l]}$ and $\Psi(\cdot, v; u)$ is the lift to $\mathbf{S}M$ of the corresponding curve $\Phi(\cdot, \pi v; u)$ on M determined by $\{X_0, \ldots, X_d\}$, the conditions in (2.8) guarantee that the hypotheses in Corollary 1.6 hold for the vector fields $\{X_0, \ldots, X_d\}$ on M.

Because of (2.8), we know from (2.6) and (2.7) that for each $v \in \mathbf{S}M$ the limit

$$(2.9) \qquad \tilde{\Lambda}(p) = \lim_{t \to \infty} \frac{1}{t} \log E^{\mathscr{W}}[|D\xi_t(\pi v)(v)|^p], \qquad p \in \mathbb{R}^1,$$

exists, is independent of $v \in \mathbf{S}M$, and that the convergence is uniform with respect to v. In addition, just as they do to Λ, the results obtained in Sect. 1 apply to $\tilde{\Lambda}$, although one should notice that the relationship of $\tilde{\Lambda}$ to $\{\tilde{X}_0, \ldots, \tilde{X}_d\}$ is somewhat different from that of Λ to $\{X_0, \ldots, X_d\}$.

As in Corollary 1.6, let m_L and $\tilde{m}_{\tilde{L}}$ be the unique stationary probability measures on M and $\mathbf{S}M$ for the one-point flows $\{\xi_t(x): t \in [0, \infty)\}$ and $\{\tilde{\xi}_t(v): t \in [0, \infty)\}$. (The operator $\tilde{L} \equiv \frac{1}{2}\sum_1^d \tilde{X}_k^2 + \tilde{X}_0$ is the generator of the process $\{\tilde{\xi}_t(v): t \in [0, \infty)\}$.) Clearly $m_L = \tilde{m}_{\tilde{L}} \circ \pi^{-1}$ and, under our hypotheses, we know that each of these measures admits a smooth positive density with respect to Riemannian measure.

According to the *multiplicative ergodic theorem* for stochastic flows of diffeomorphisms (see [7, Theorem 2.1]), there exist *Lyapunov exponents* $\lambda_1 \geqq \ldots \geqq \lambda_N$ such that, for m_L-almost every $x \in M$:

$$\lim_{t \to \infty} (D\xi_t(x)^* D\xi_t(x))^{1/2t} = \Lambda_x (a.s., \mathscr{W}), \tag{2.10}$$

where Λ_x is a $\mathrm{Hom}(\mathbf{T}_x M;\, \mathbf{T}_x M)$-valued random variable having non-random eigenvalues $e^{\lambda_1} \geqq \ldots \geqq e^{\lambda_N}$. In fact, under our hypotheses, Theorem 1.5, allows us to strengthen this statement to the assertion that (2.10) is valid for every $x \in M$.

Applying Theorem 1.29, we see that for all $x \in M$:

$$\lambda_\Sigma \equiv \lambda_1 + \ldots \lambda_N = \lim_{t \to \infty} \frac{1}{t} \log J_t(x) = \Lambda'(0) = \int_M Q(y)\, m_L(dy), \tag{2.11}$$

where $Q(y) = (\operatorname{div} X_0)(y) + \frac{1}{2} \sum_{k=1}^{d} (X_k \operatorname{div} X_k)(y)$; and that for all $v \in \mathbf{S}M$:

$$\lambda_1 = \lim_{t \to \infty} \frac{1}{t} \log |D\xi_t(x)(v)| = \tilde{\Lambda}'(0) = \int_{\mathbf{S}M} R(u)\, \tilde{m}_{\tilde{L}}(du), \tag{2.12}$$

where $R(u) = g_0(u) + \frac{1}{2} \sum_{k=1}^{d} (\tilde{X}_k g_k)(u)$. In both (2.11) and (2.12), the assertion is a $\mathscr{W}$-almost sure statement.

The function Λ was studied previously in [9] and [12]. Because of its connection with the Lyapunov spectrum, $\tilde{\Lambda}$ is called the *Lyapunov moment function*. It has been analyzed for linear systems in [1], [2], [4], and [21]. The formula (2.12) is due originally to A. Carverhill [8].

(2.13) **Theorem.** *$\tilde{\Lambda}(Np) \leqq \Lambda(p)$ if $p \in [-1, 0]$, and $\tilde{\Lambda}(Np) \geqq \Lambda(p)$ if $p \notin (-1, 0)$.*

Proof. Let ν_x denote the unique rotation invariant probability measure on $\mathbf{S}_x M \equiv \{v \in \mathbf{T}_x M\colon |v| = 1\}$. Then, for each $(t, x) \in (0, \infty)$ the random measure $\nu_{\xi_t(x)} \circ \tilde{\xi}_t$ is equivalent to ν_x $\mathscr{W}$-almost surely. In fact by a formula which is essentially due to Furstenberg [13, p. 425],

$$\frac{d(\nu_{\xi_t(x)} \circ \tilde{\xi}_t)}{d\nu_x}(v) = J_t(x) |D\xi_t(x)(v)|^{-N}, \qquad v \in \mathbf{S}_x M.$$

Therefore,

$$\int_{\mathbf{S}_x M} |D\xi_t(x)(v)|^{Np}\, \nu_x(dv) = (J_t(x))^p \int_{\mathbf{S}_x M} \left(\frac{d(\nu_{\xi_t(x)} \circ \tilde{\xi}_t)}{d\nu_x}(v) \right)^{-p} \nu_x(dv).$$

By Jensen's inequality, the right hand side of the preceding is greater or equal or is less than or equal to $(J_t(x))^p$ according to whether $p \notin (-1, 0)$ or $p \in [-1, 0]$. Hence, after applying the operation "$1/t \log E^{\mathscr{W}}[\cdot]$" to both sides, the desired result follows from (2.3) and (2.9). □

(2.14) **Corollary.** i) $\tilde{\Lambda}(-N) = \Lambda(-1) \geqq -\lambda_\Sigma \geqq 0$ *and* $N\tilde{\Lambda}'(0) \geqq \Lambda'(0)$.

ii) *There exists a smooth Riemannian structure on M with respect to which* ξ_t *is* $\mathscr{W}$*-almost surely measure preserving for every* $t \in [0, \infty)$ *if and only if* $\Lambda(\cdot) \equiv 0$ *if and only if* $\tilde{\Lambda}(-N) = 0$ *if and only if* $\lambda_\Sigma = 0$.

iii) *There exists a smooth Riemannian structure on M with respect to which* ξ_t *is* $\mathscr{W}$*-almost surely conformal for every* $t \in [0, \infty)$ *if and only if* $\tilde{\Lambda}(N\cdot) \equiv \Lambda(\cdot)$ *if and only if* $N\tilde{\Lambda}'(0) = \Lambda'(0)$.

iv) *There exists a smooth Riemannian structure on M with respect to which* ξ_t *is* $\mathscr{W}$*-almost surely isometric for every* $t \in [0, \infty)$ *if and only if* $\tilde{\Lambda}(\cdot) \equiv 0$ *if and only if* $N\tilde{\Lambda}'(0) = \Lambda'(0) = 0$.

Proof. It is shown in [3, Corollary 5.1] that $\lambda_\Sigma \leqq 0$, and the rest of i) is a consequence of the Theorem 2.13 and the convexity of Λ.

To prove ii)–iv), observe that, because M is compact (and therefore that all Riemannian structures on M are equivalent), the functions $\tilde{\Lambda}$ and Λ are independent of the particular structure chosen. Now suppose that a Riemannian structure having the asserted property exists. Then, $\mathscr{W}$-almost surely, one has: $J_t(x) \equiv 1$, $|D\xi_t(x)(v)|^N \equiv J_t(x)$, or $|D\xi_t(x)(v)|^N \equiv 1$ for all $x \in M$ and $v \in \mathbf{S}_x M$ according to whether one is dealing with case ii), iii), or iv). Thus, one gets $\Lambda(\cdot) \equiv 0$, $\tilde{\Lambda}(N\cdot) \equiv \Lambda(\cdot)$, or $\tilde{\Lambda}(\cdot) \equiv 0$ respectively; and, in view of i), it is easy to check that each of these imply $\lambda_\Sigma = 0$, $N\tilde{\Lambda}(0) = \Lambda'(0)$, and $N\tilde{\Lambda}(0) = \Lambda'(0) = 0$ respectively. Finally, by [3, Corollary 5.1], $\lambda_\Sigma = 0$ implies that m_L is $\mathscr{W}$-almost surely preserved by ξ_t for every $t \in [0, \infty)$; and so the proof of ii) is completed by adjusting the Riemannian structure on M so that m_L becomes the Riemannian measure. Also, the rest of cases iii) and iv) is covered by [5, Theorem 7.6 and Remark 3]. □

The preceding result deals with the rather rigid situations when either Λ or $\tilde{\Lambda}$ themselves or the relationship between them is degenerate. In the next result we see what can be deduced from information about the asymptotic behavior of $\tilde{\Lambda}$ and Λ.

(2.15) **Theorem.** *Assume that* $\mathrm{Lie}(X_1, \ldots, X_d)(x) = \mathbf{T}_x M$, $x \in M$. *Then* $\lim_{p \to \infty} \frac{1}{p^2} \Lambda(p) = 0$ *if and only if* $\lim_{p \to -\infty} \frac{1}{p^2} \Lambda(p) = 0$ *if and only if there is a uniformly positive* $h \in C^\infty(M)$ *such that* $\mathrm{div}(hX_k) \equiv 0$ *for all* $1 \leqq k \leqq d$, *in which case*

$$\lim_{p \to \pm\infty} \frac{1}{p} \Lambda(p) = \pm \sup\left\{ \frac{\pm \mathrm{div}(hX_0)(x)}{h(x)} :\ x \in M \right\}. \tag{2.16}$$

Now assume that

$$\mathrm{Lie}(\tilde{X}_1, \ldots, \tilde{X}_d)(v) = \mathbf{TS}_v M, \qquad v \in \mathbf{S}M.$$

Then $\lim_{p\to\infty}\frac{1}{p^2}\tilde{\Lambda}(p)=0$ *if and only if* $\lim_{p\to-\infty}\frac{1}{p^2}\tilde{\Lambda}(p)=0$ *if and only if there is a smooth Riemannian structure with metric* ρ *on* M *such that* X_k *is an infinitesimal* ρ*-isometry for each* $1\leqq k\leqq d$; *in which case*

$$\lim_{p\to\pm\infty}\frac{1}{p}\tilde{\Lambda}(p)=\pm\sup\{\pm(g_0(v)-\tilde{X}_0 f(v)):\ v\in \mathbf{S}M\}, \tag{2.17}$$

where $f\equiv -\frac{1}{2}\log\rho(\cdot,\cdot)$ *and* g_0 *is the function appearing in* (2.7).

Proof. Note that the Remark (1.18) allows us to conclude that there is an $f\in C^\infty(M)$ satisfying

$$X_k f=\operatorname{div} X_k, \qquad 1\leqq k\leqq d$$

if either

$$\lim_{p\to\infty}\frac{1}{p^2}\Lambda(p)=0 \quad\text{or}\quad \lim_{p\to-\infty}\frac{1}{p^2}\Lambda(p)=0,$$

from which we get the required function by setting $h=e^{-f}$. Conversely, if such an h exists, then, substituting $\phi-\log h$ for ϕ in (1.14) and reverting to the equivalent formulation (1.13), yields

$$\Lambda(p)=\sup\left\{\int_M \frac{\operatorname{div}(hX_0)}{h}\,d\mu-J_0(\mu):\ \mu\in\mathscr{P}(M)\right\},$$

where the $J_0(\mu)$ is the element of $[0,\infty]$ described in the paragraph preceding Corollary 1.12 (and should not be confused with the matrix $J_t(x)$ for $t=0$ of this section). In particular, since $J_0(\mu)<\infty$ whenever μ is smooth, this proves (2.16) holds and therefore that $\lim_{|p|\to\infty}\frac{1}{p^2}\Lambda(p)=0$ when h exists.

The proof of the analogous statements for $\mathbf{S}M$ is precisely the same, only now one must notice that the existence of an $f\in C^\infty(\mathbf{S}M)$ satisfying $\tilde{X}_k f(v)=g_k(v)=\langle \nabla X_k(\pi(v))(v),v\rangle$ for all $v\in\mathbf{S}M$ is equivalent to the existence of a Riemannian metric ρ on M for which the X_k's, $1\leqq k\leqq d$, are infinitesimal isometries and that, up to an additive constant, $f=-\frac{1}{2}\log\rho(\cdot,\cdot)$ on $\mathbf{S}M$. To see this equivalence, first suppose that such a metric ρ exists and set $f=-\frac{1}{2}\log\rho(\cdot,\cdot)$ on $\mathbf{S}M$. Then, using ρ to denote $\rho(\cdot,\cdot)$, we see that for each $1\leqq k\leqq d$,

$$0=D\rho(v)(\mathbf{T}X_k(v))=(2\langle\nabla X_k(\pi v)(v),v\rangle-2\tilde{X}_k f(v))\,\rho(v), \qquad v\in\mathbf{S}M;$$

and so this choice of f works. Conversely, suppose that f exists and define $\rho\colon \mathbf{T}M\to[0,\infty)$ by

$$\rho(v)=\begin{cases}|v|^2\exp\left(-2f\left(\frac{v}{|v|}\right)\right) & \text{if } |v|\neq 0\\ 0 & \text{if } |v|=0.\end{cases}$$

By the preceding calculations, we see that $D\rho(\cdot)(\mathbf{T}X_k)(\cdot)\equiv 0$. Hence, all that we have to do is check that for each $x\in M$ the function $v\in\mathbf{T}_xM\mapsto\rho(v)$ is quadratic. To this end, set

$$\mathscr{L}_x=\{\nabla X(x):\ X\in\mathrm{Lie}(X_1,\ldots,X_d)\ \text{ and }\ X(x)=0\}.$$

Clearly $\mathscr{L}_x$ is a Lie subalgebra of $\mathrm{gl}(\mathbf{T}_xM)$. Let G_x denote the corresponding connected subgroup of $\mathrm{GL}(\mathbf{T}_xM)$. Because ρ on $\mathbf{T}_xM$ is constant along all G_x-orbits, G_x must be conjugate to a subgroup of $\mathrm{SO}(\mathbf{T}_xM)$. Hence there is a $Q_x\in\mathrm{GL}(\mathbf{T}_xM)$ such that $\langle Q_x\mathscr{A}v, Q_xv\rangle=0$ for all $\mathscr{A}\in\mathscr{L}_x$. We will show that $\rho(\cdot)$ is a constant multiple of $\langle Q_x\cdot, Q_x\cdot\rangle$. To this end, set

$$h(v)=\log\rho(v)-\log\langle Q_xv, Q_xv\rangle,\qquad v\in\mathbf{S}_xM,$$

and note that it suffices to check that $Dh(v)(u)=0$ for $u\in\mathbf{T}_xM$ satisfying $\langle u, v\rangle=0$. But, because $\mathrm{Lie}(\tilde{X}_1,\ldots,\tilde{X}_d)(v)$ has full rank, for each such u there exists a $\mathscr{A}\in\mathscr{L}_x$ such that $u=\mathscr{A}v-\langle\mathscr{A}v, v\rangle v$; and so

$$\begin{aligned}Dh(v)(u)&=\frac{D\rho(v)(\mathscr{A}v-\langle\mathscr{A}v,v\rangle v)}{\rho(v)}-2\frac{\langle Q_xv, Q_x(\mathscr{A}v-\langle\mathscr{A}v,v\rangle v\rangle}{\langle Q_xv, Q_xv\rangle}\\&=-2\langle\mathscr{A}v,v\rangle+2\langle\mathscr{A}v,v\rangle=0.\quad\square\end{aligned}$$

(2.18) *Remark.* Note that the condition $\mathrm{div}(hX_k)=0$, $1\leqq k\leqq d$, is equivalent to the statement that the measure having density h with respect to the Riemann measure is invariant under the flows generated by $\{X_1,\ldots,X_d\}$. In particular, if, in addition, $\mathrm{div}(hX_0)=0$, then the conclusion drawn in ii) of Corollary 2.14 follows from the support theorem for diffusions. The result in iv) of Corollary 2.14 is related similarly to the second part of Theorem 2.15. Finally, in the case when $X_0=0$ and $X_k=\mathrm{grad}\,f_k$, $1\leqq k\leqq d$, where $(f_1,\ldots,f_d)$: $M\to\mathbb{R}^d$ is an isometric embedding, Chappell in [9] evaluates λ_Σ and $\lim_{p\to\infty}\frac{1}{p^2}\Lambda(p)$ in terms of the mean curvature of M as a submanifold of $\mathbb{R}^d$.

3. The Two Point Motion near the Diagonal

Let $\{\xi_t:\ t\in[0,\infty)\}$ be the stochastic flow of diffeomorphisms given by (2.1); and, for $(x,y)\in\hat{M}\equiv M^2\backslash\Delta$ (Δ denotes the diagonal in M^2), consider the *two-point motion* $\{(\xi_t(x),\xi_t(y)):\ t\in[0,\infty)\}$. Because, $\mathscr{W}$-almost surely, $\xi_t(x)\neq\xi_t(y)$ for any $t\in[0,\infty)$, we can think of two-point motion as a conservative Markov process on $\hat{M}$, in which case the problem of determining when $\mathrm{dist}(\xi_t(x),\xi_t(y))\to 0$ as $t\to\infty$ becomes a question of the transience or recurrence of the two-point motion on $\hat{M}$. In order to study this question, we first write a stochastic differential equation for the two-point motion. Namely, for a vector field X on M define the vector field $X^{(2)}\in\Gamma(\mathbf{T}\hat{M})$ so that

(3.1) $$X^{(2)}(x,y)=(X(x),X(y))\in\mathbf{T}_xM\times\mathbf{T}_yM\simeq\mathbf{T}_{(x,y)}\hat{M},\qquad (x,y)\in\hat{M}.$$

It is then clear that, for $z=(x, y)\in\hat{M}$, the two-point motion $\eta_t(z)\equiv(\xi_t(x), \xi_t(y))$ is determined by

(3.2) $$d\eta_t(z)=\sum_{k=1}^{d} X_k^{(2)}(\eta_t(z))\circ d\theta_k(t)+X_0^{(2)}(\eta_t(z))\,dt \quad \text{with } \eta_0(z)=z.$$

In particular, the law of the two-point is determined by the operator

(3.3) $$L^{(2)}=\tfrac{1}{2}\sum_{k=1}^{d}(X_k^{(2)})^2+X_0^{(2)}$$

on $C_0^\infty(\hat{M})$. It should be noted that no matter how one chooses the X_k's, the operator $L^{(2)}$ degenerates at Δ. Hence, there is no possibility that $L^{(2)}$ is uniformly elliptic (or even uniformly subelliptic) on $\hat{M}$. On the other hand, it is possible for $L^{(2)}$ to be elliptic on $\hat{M}$.

Define the map Φ: $\mathbf{T}M\to M\times M$ by

(3.4) $$\Phi(v)=(\pi v, \exp_{\pi v}(v)).$$

That is, $\Phi(v)$ is the pair consisting of the initial and end points of the geodesic which starts from πv with velocity v and runs for a unit length of time. Because M is compact, there is a positive δ_0 (the injectivity radius) such that Φ is diffeomorphic from $\{v\in\mathbf{T}M\colon 0<|v|<\delta_0\}$ onto $\hat{M}_{\delta_0}\equiv\{(x, y)\in M^2\colon 0<\operatorname{dist}(x, y)<\delta_0\}$. In particular, we can use Φ to transfer the natural polar coordinate system on $\{v\in\mathbf{T}M\colon |v|\neq 0\}$ to $\hat{M}_{\delta_0}$:

$$(r, \theta)\in(0, \delta_0)\times\mathbf{S}M\mapsto\Phi(r\theta)\in\hat{M}_{\delta_0}.$$

This polar coordinate system will play a role in our analysis of the two point motion when we compare the behavior of $\eta_t(z)$ near Δ to that of the linearized motion $D\xi_t(\pi v)(v)$ near the zero-section in $\mathbf{T}M$.

Given a vector field X on M, define the vector field $\Phi^* X^{(2)}$ on $\Phi^{-1}(\hat{M}_{\delta_0})$ by

$$\Phi^* X^{(2)}=(D\Phi(v))^{-1}X^{(2)}(\Phi(v)).$$

(That is, $\phi^* X^{(2)}$ is the vector field on $\Phi^{-1}(\hat{M}_{\delta_0})$ obtained as the pullback under Φ of $X^{(2)}$.) In the following, recall the decomposition, described between (2.4) and (2.5), of the tangent space $\mathbf{T}_v\mathbf{T}M$ into horizontal and vertical subspaces.

(3.5) **Lemma.** *For $r\in(0, \delta_0)$ and $\theta\in\mathbf{S}M$, let $W(t)$, $t\in[0, r]$, be the Jacobi field along the geodesic $x(t)=\exp_{\pi\theta}(t\theta)$ which satisfies $W(0)=X(x(0))$ and $W(r)=X(x(r))$. Then*

$$\Phi^* X^{(2)}(r\theta)=\left(X(\pi\theta), r\frac{\nabla W}{dt}(0)\right).$$

Proof. Let $x=x(0)$ and $y=x(r)$. Using $s\in\mathbb{R}^1\mapsto\psi_s$ to denote the flow of diffeomorphisms generated by X, set

$$v_s=\Phi^{-1}(\psi_s(x), \psi_s(y)).$$

Then $r\theta = v_0$ and

$$\Phi^* X^{(2)}(r\theta) = \frac{d}{ds}(v_s)|_{s=0}.$$

The horizontal component of $\Phi^* X^{(2)}(r\theta)$ is given by

$$\frac{d}{ds}(\pi \circ v_s)(0) = \frac{d}{ds}\psi_s(x)|_{s=0} = X(x),$$

and the vertical component is $\frac{\nabla}{ds} v_s|_{s=0}$. Next, define

$$\alpha(s,t) = \exp_{\psi_s(x)}\left(\frac{t}{r} v_s\right) \quad \text{for } (s,t)\in(-\varepsilon,\varepsilon)\times[0,r].$$

Clearly, $\alpha(0,\cdot) = x(\cdot)$ while $t\in[0,r]\mapsto\alpha(s,t)$ is a geodesic for each $s\in(-\varepsilon,\varepsilon)$. From this it is clear that

$$t\in[0,r]\mapsto W(t) \equiv \frac{\partial\alpha}{\partial s}(0,t)$$

is the Jacobi field along $x(t)$, $t\in[0,r]$, satisfying $W(0) = X(x)$ and $W(r) = X(y)$. Moreover, $\frac{\partial\alpha}{\partial t}(s,0) = \frac{1}{r} v_s$; and so

$$\frac{\nabla W}{dt}(0) = \frac{\nabla}{\partial t}\frac{\partial\alpha}{\partial s}(0,0) = \frac{\nabla}{\partial s}\frac{\partial\alpha}{\partial t}(0,0) = \frac{1}{r}\frac{\nabla}{ds} v_s|_{s=0}. \quad \square$$

For information about Jacobi fields, the reader might want to consult [10, pp. 14–16].

Give $(0,\delta_0)\times \mathbf{S}M$ the Riemannian structure it inherits as a product, and define $\pi_2: (0,\delta_0)\times \mathbf{S}M \to M$ so that $\pi_2(r,\theta) = \pi\theta$.

(3.6) **Lemma.** *Define* $(r,\theta)\times \mathbf{S}M \mapsto H(r,\theta)\in \mathbf{T}_{\pi\theta}M$ *by the equation*

$$\Phi^* X^{(2)}(r\theta) = (X(\pi\theta), r\nabla X(\pi\theta)(\theta) + r^2 H(r,\theta)).$$

Then, H is a smooth section of the pullback bundle $\pi_2^* \mathbf{T}M$; and, for each $0<\delta_1<\delta_0$, any order covariant derivative of H is bounded on $(0,\delta_1]\times \mathbf{S}M$.

Proof. Notice first that $\Phi^* X^{(2)}$ has a (unique) smooth extension to $\mathbf{B}M(\delta_0) \equiv \{v\in \mathbf{T}M: |v|<\delta_0\}$. Define $G(v)$ for $v\in \mathbf{B}M(\delta_0)$ so that $G(v) + \nabla X(\pi v)(v)$ is the vertical component of $\Phi^* X^{(2)}(v)$. It is then clear that G is smooth on $\mathbf{B}M(\delta_0)$ and that $r^2 H(r,\theta) = G(r\theta)$ for $(r,\theta)\in(0,\delta_0)\times \mathbf{S}M$. Thus, it remains only to check that $\frac{1}{r^2}G(r\theta)$ and each of its derivatives remain bounded as $r\searrow 0$.

For fixed $\theta\in \mathbf{S}M$ let $W(\cdot)$ denote the Jacobi field along the geodesic $x(\cdot)$ issuing from $\pi\theta$ in the direction θ. Also, use $P_t: \mathbf{T}_{x(t)}M \to \mathbf{T}_{x(0)}$ to denote parallel translation along $x(\cdot)$ from $x(t)$ to $x(0)$. The Jacobi field equation together with the smoothness of X imply the existence of a $K<\infty$ and an $\varepsilon>0$, which are independent of θ, such that

$$\left|P_t W(t)-W(0)-t\frac{\nabla W}{dt}(0)\right|\leqq K t^2$$

and

$$|P_t X(x(t))-X(t)-t\nabla X(\pi\theta)(\theta)|\leqq K t^2$$

so long as $0<t<\varepsilon$. Combined with Lemma 3.5, these inequalities lead to $|G(r\theta)|\leqq 2Kr^2$ for $(r,\theta)\in(0,\varepsilon\wedge\delta_0)\times\mathbf{S}M$. In particular, if 0_x denotes the zero vector in $\mathbf{T}_x M$, then not only $G(0_x)=0$ but also $DG(0_x)(0,\theta)=0$, where $DG(\cdot)(0,\theta)$ denotes the total derivative of G in the vertical direction $(0,\theta)\in\mathbf{T}_\theta\mathbf{T}M$. Hence, by Taylor's formula:

$$H(r,\theta)=\int_0^1 (1-s)\,D^2G(sr\theta)((0,\theta),(0,\theta))\,ds;$$

and obviously the required result follows from this. □

Since $\Phi^* X^{(2)}$ is just $X^{(2)}$ written in terms of v rather than (x, y), we will, from now on, use $X^{(2)}$ in place of $\Phi^* X^{(2)}$. Further, we will consider both $X^{(2)}$ and $\mathbf{T}X$ in terms of the polar coordinate system (r,θ). Finally, identifying $\mathbf{T}_{(r,\theta)}((0,\delta_0)\times\mathbf{S}M)$ with $\mathbf{T}_r(0,\delta_0)+\mathbf{T}_\theta\mathbf{S}M$, we have

$$\text{(3.7)}\qquad \mathbf{T}X(r,\theta)=\langle\nabla X(\pi\theta)(\theta),\theta\rangle\, r\frac{\partial}{\partial r}+\tilde{X}(\theta)$$

and

$$\text{(3.8)}\qquad X^{(2)}(r,\theta)=(\langle\nabla X(\pi\theta)(\theta),\theta\rangle r+\langle H(r,\theta),\theta\rangle r^2)\frac{\partial}{\partial r}+(\tilde{X}(\theta)+r\tilde{H}(r,\theta)),$$

where the horizontal and vertical components of $\tilde{H}(r,\theta)$ are $0_{\pi\theta}$ and $H(r,\theta)-\langle H(r,\theta),\theta\rangle\theta$, respectively.

(3.9) **Proposition.** *For each $0<\delta_1<\delta_0$ and $X\in\Gamma(\mathbf{T}M)$ there exists a $K<\infty$ with the properties that*

$$|(X^{(2)}-\mathbf{T}X)(\phi\otimes\psi)(r,\theta)|\leqq K(r|\phi(r)|+r^2|\phi'(r)|)\,\|\psi\|_{C^1(\mathbf{S}M)}$$

and

$$|((X^{(2)})^2-(\mathbf{T}X)^2)(\phi\otimes\psi)(r,\theta)|\leqq K(r|\phi(r)|+r^2|\phi'(r)|+r^3|\phi''(r)|)\,\|\psi\|_{C^2(\mathbf{S}M)}$$

for all $\phi\in C^2((0,\delta_1))$, $\psi\in C^2(\mathbf{S}M)$, and $(r,\theta)\in(0,\delta_1)\times\mathbf{S}M$ (where $(\phi\otimes\psi)(r,\theta)=\phi(r)\,\psi(\theta)$).

Proof. We will prove the first estimate; the second one follows in a similar fashion from the bounds on H and its first covariant derivatives.

Set $g(\theta)=\langle \nabla X(\pi\theta)(\theta), \theta\rangle$. Clearly $g\in C^\infty(\mathbf{S}M)$ and, from (3.7) and (3.8),

$$\mathbf{T}X\ (\phi\otimes\psi)(r,\theta)=r\phi'(r)\,g(\theta)\,\psi(\theta)+\phi(r)(\tilde{X}\psi)(\theta)$$

while

$$X^{(2)}(\phi\otimes\psi)(r,\theta)=\mathbf{T}X(\phi\otimes\psi)(r,\theta)+r^2\phi'(r)\langle H(r,\theta),\theta\rangle\,\psi(\theta)+r\phi(r)(\tilde{H}(r,\cdot)\psi)(\theta).$$

Clearly the desired estimate follows from these. □

(3.10) **Corollary.** *Set*

$$\mathbf{T}L=\tfrac{1}{2}\sum_{k=1}^{d}(\mathbf{T}X_k)^2+\mathbf{T}X_0.$$

(That is $\mathbf{T}L$ *is the operator on* $C^\infty(\mathbf{T}M)$ *which determines the distribution of the derivative process* $\{D\xi_t(\pi v)(v)\colon t\in[0,\infty)\}$.*) Then for each* $0<\delta_1<\delta_0$ *there is a* $K<\infty$ *such that*

$$|(L^{(2)}-\mathbf{T}L)(\phi\otimes\psi)(r,\theta)|\leqq K(r|\phi(r)|+r^2|\phi'(r)|+r^3|\phi''(r)|)\,\|\psi\|_{C^2(\mathbf{S}M)}$$

for all $\phi\in C^2((0,\delta_0))$, $\psi\in C^2(\mathbf{S}M)$, *and* $(r,\theta)\in(0,\delta_1)\times\mathbf{S}M$.

(3.11) *Remark.* When $\phi(r)=\log r$ and $\psi\equiv 1$, the estimates in (3.9), and therefore in (3.10), can be improved by replacing the term $r|\phi(r)|$ on the right hand side by r.

Up to this point the discussion in this section has not used any non-degeneracy hypotheses about the vector fields X_k. However, in order to construct suitable functions ϕ and ψ to put into the preceding corollary, we will from now on impose the conditions stated in (2.8). Recalling the notation introduced following (2.6) and (2.7), set

$$\tilde{Q}(\theta)=g_0(\theta)+\tfrac{1}{2}\sum_{k=1}^{d}(\tilde{X}_k g_k)(\theta)\quad\text{and}\quad \tilde{a}(\theta)=\sum_{k=1}^{d}(g_k(\theta))^2$$

for $\theta\in\mathbf{S}M$. For $p\in\mathbb{R}^1$, set

$$\tilde{L}_p=\tilde{L}+p\sum_{k=1}^{d}g_k\tilde{X}_k.$$

Then, for any $\psi\in C^2(\mathbf{S}M)$,

$$\mathbf{T}L(r^p\psi(\theta))=r^p(\tilde{L}_p+p\tilde{Q}+\tfrac{1}{2}p^2\tilde{a})\,\psi(\theta) \tag{3.12}$$

and

$$\mathbf{T}L(r^p(\log r)\,\psi(\theta))=r^p(\log r)(\tilde{L}_p+p\tilde{Q}+\tfrac{1}{2}p^2\tilde{a})\,\psi(\theta)+r^p\left(\sum_{k=1}^{d}g_k\tilde{X}_k+\tilde{Q}+p\tilde{a}\right)\psi(\theta). \tag{3.13}$$

(The operator $\tilde{L}=\frac{1}{2}\sum_{1}^{d}\tilde{X}_k^2+\tilde{X}_0$ is the one associated with the process $\{\tilde{\xi}_t\colon t\in[0,\infty)\}$ on $\mathbf{S}M$.) Both of these expressions are direct calculations from the formula (3.7) with $X=X_k$. Also, notice that (3.12) is precisely i) of Theorem 1.29 when M is replaced by $\mathbf{S}M$. In fact, Eqs. (2.6) and (2.7) show that we are in the situation described just before Corollary 1.12 and that our $\mathbf{T}L$ here is the same as the $\hat{L}$ there when one uses coordinates (r,θ) instead of $(\theta,\log r)$. Hence, by Theorem 1.29, for each $p\in\mathbb{R}^1$ there is a strictly positive $\phi_p\in C^\infty(\mathbf{S}M)$ such that

$$(3.14)\qquad (\tilde{L}_p+p\tilde{Q}+p^2\tilde{a})\,\phi_p=\tilde{\Lambda}(p)\,\phi_p.$$

(The smoothness of ϕ_p is guaranteed by the conditions in (2.8).) In addition, the eigenvalue $\tilde{\Lambda}(p)$ is simple for each p; and, therefore, by analytic perturbation theory (cf. [14, p. 365]), the map $p\in\mathbb{R}^1\mapsto\phi_p\in C^\infty(\mathbf{S}M)$ can be chosen to be analytic. Writing ϕ'_p to denote $\frac{\partial}{\partial p}\phi_p$, we obtain from (3.14) the equation

$$(3.15)\qquad (\tilde{L}_p+p\tilde{Q}+\tfrac{1}{2}p^2\tilde{a})\,\phi'_p+\left(\sum_{k=1}^{d} g_k\tilde{X}_k+\tilde{Q}+p\tilde{a}\right)\phi_p=\tilde{\Lambda}(p)\,\phi'_p+\tilde{\Lambda}'(p)\,\phi_p.$$

After combining Eqs. (3.12) through (3.15), we arrive at

$$(3.16)\qquad (\mathbf{T}L-\tilde{\Lambda}(p))(r^p\phi_p(\theta))=0$$

and

$$(3.17)\qquad (\mathbf{T}L-\tilde{\Lambda}(p))(r^p(\log r)\,\phi_p(\theta)+r^p\phi'_p(\theta))=r^p\tilde{\Lambda}'(p)\,\phi_p(\theta).$$

(3.18) **Theorem.** *Assume that the conditions in* (2.8) *hold and that* $\tilde{\Lambda}$ *does not vanish identically. Then for each choice of* $-\infty<a<b<\infty$ *there is a* $\delta>0$ *and a* $K<\infty$ *for which the following assertions hold.*

i) *For each* $p\in[a,b]$ *there exist* $\phi_p^\pm\in C^\infty((0,\delta)\times\mathbf{S}M)$ *such that*

$$(L^{(2)}-\tilde{\Lambda}(p))\,\phi_p^+\geqq 0\geqq(L^{(2)}-\tilde{\Lambda}(p))\,\phi_p^-\quad on\quad (0,\delta)\times\mathbf{S}M$$

and

$$\frac{1}{K}r^p\leqq\phi_p^\pm(r,\theta)\leqq K r^p,\qquad (r,\theta)\in(0,\delta)\times\mathbf{S}M.$$

ii) *There exist* $\psi^\pm\in C^\infty((0,\delta)\times\mathbf{S}M)$ *such that*

$$L^{(2)}\psi^+\geqq\tilde{\Lambda}'(0)\geqq L^{(2)}\psi^-\quad on\quad (0,\delta)\times\mathbf{S}M$$

and

$$|\psi^\pm(r,\theta)-\log r|\leqq K,\qquad (r,\theta)\in(0,\delta)\times\mathbf{S}M.$$

iii) *For each* $p\in[a,b]$ *there exists a* $\zeta_p\in C^\infty((0,\delta)\times\mathbf{S}M)$ *such that*

$$(L^{(2)}-\tilde{\Lambda}(p))\,\zeta_p+\tilde{\Lambda}'(p)\,\phi_p^+\geqq 0\quad on\quad (0,\delta)\times\mathbf{S}M$$

and

$$\frac{1}{K} r^p |\log r| \leqq \zeta_p(r,\theta) \leqq K r^p |\log r|, \qquad (r,\theta)\in(0,\delta)\times \mathbf{S}M.$$

Proof. Without loss in generality, we assume that $a<0<b$ and that $\phi_0\equiv 1$. The analytic dependence of ϕ_p on p together with Corollary 3.10 and Remark 3.11 then imply the existence of $\delta>0$ and $K<\infty$ such that

$$\frac{1}{K} \leqq \phi_p \leqq K \quad \text{and} \quad \|\phi'_p\| \leqq K,$$

$$|(L^{(2)}-\mathbf{T}L)(r^p\phi_p(\theta))| \leqq K r^{p+1}, \qquad (r,\theta)\in(0,\delta)\times\mathbf{S}M,$$

$$|(L^{(2)}-\mathbf{T}L)(\log r+\phi'_0(\theta))| \leqq K r, \qquad (r,\theta)\in(0,\delta)\times\mathbf{S}M,$$

$$|(L^{(2)}-\mathbf{T}L)\left(r^p\left(\log\frac{1}{r}\right)\phi_p(\theta)-r^p\phi'_p(\theta)\right)| \leqq K r^p|\log r|, \qquad (r,\theta)\in(0,\delta)\times\mathbf{S}M,$$

for all $p\in[a, b]$. Since $\tilde{\Lambda}$ is non-constant on every non-empty open interval of $\mathbb{R}^1$, there exists an $\varepsilon>0$ such that for each $p\in[a, b]$ there exists a $q\in[p+\frac{1}{2}, p+\frac{3}{4}]$ for which $|\tilde{\Lambda}(q)-\tilde{\Lambda}(p)|\geqq\varepsilon$. Set $\delta_1=\min\left(\delta, \frac{\varepsilon}{2(K^2+K^4)}\right)$. Next, define $k=k(p, q)\in\mathbb{R}^1$ so that $k(\tilde{\Lambda}(q)-\tilde{\Lambda}(p))>0$ and

$$|k|=\frac{K^2\delta_1^{p+1-q}}{\varepsilon-K^2\delta_1}.$$

Finally, set

$$\phi_p^{\pm}(r,\theta)=r^p\phi_p(\theta)\pm k r^q\phi_q(\theta).$$

Clearly $\phi_p^{\pm}\in C^\infty((0,\delta_1)\times\mathbf{S}M)$ and

$$(L^{(2)}-\tilde{\Lambda}(p))\,\phi_p^{\pm}(r,\theta)=\pm k(\tilde{\Lambda}(q)-\tilde{\Lambda}(p))\,r^q\phi_q(\theta)+R_1+R_2,$$

where $|R_1|\leqq K r^{p+1}$ and $|R_2|\leqq|k|\,K r^{q+1}$ for $r\in(0,\delta_1)$. It follows from the upper and lower bounds on ϕ_p and ϕ_q together with the choice of δ_1 and k that the requirements of i) are satisfied by this choice of $\phi_p^{\pm}$ when the K there is twice the K here.

Next, for ε, δ_1, q corresponding to $p=0$, and k as in the preceding, set

$$\phi^{\pm}(r,\theta)=\log r+\phi'_0(\theta)\pm k r^q\phi_q(\theta).$$

Then the calculations like those made above show that $\psi^{\pm}$ can be used to obtain a proof of ii).

To prove iii), take

$$\zeta_p(r,\theta)=r^p\left(\log\frac{1}{r}\right)\phi_p(\theta)-r^p\phi'_p(\theta)+l r^q\phi_q(\theta),$$

where l will be chosen below. Then

$$(L^{(2)}-\tilde{\Lambda}(p))\,\zeta_p(r,\theta)+\tilde{\Lambda}'(p)\,\phi_p^+(\theta)=(l(\tilde{\Lambda}(q)-\tilde{\Lambda}(p))+k\tilde{\Lambda}'(p))\,r^q\phi_q(\theta)+R_1+R_2,$$

where $|R_1|\leqq K r^{p+1}|\log r|$ and $|R_2|\leqq|l|\,K r^{q+1}$ for $0<r<\delta_1$. For any $0<\delta_2 \leqq\min(\delta_1, e^{-4})$ we have

$$r^{p+1}|\log r|\leqq r^{q+\frac{1}{4}}|\log r|\leqq r^q|\log\delta_2|\,\delta_2^{1/4}, \qquad r\in(0,\delta_2].$$

Thus

$$(L^{(2)}-\tilde{\Lambda}(p))\,\zeta_p(r,\theta)+\tilde{\Lambda}'(p)\,\phi_p^+(r,\theta)\geqq 0$$

so long as $r\in(0,\delta_2)$, $l(\tilde{\Lambda}(q)-\tilde{\Lambda}(p))>0$, and

$$|l|(\varepsilon-K^2\delta_2)\geqq K^2\delta_2^{1/4}\log\frac{1}{\delta_2}-k\tilde{\Lambda}'(p).$$

Since both $|k(p,q)|$ and $|\tilde{\Lambda}'(p)|$ stay bounded as p ranges over $[a,b]$, there exists a $K_1<\infty$ such that for all $\delta_2\leqq\min(\delta_1, e^{-4})$ and $p\in[a,b]$ the preceding will hold for an $l\in[-K_1, K_1]$. Finally, if δ_2 is chosen so that

$$\frac{1}{|\log\delta_2|}+K_1\,\delta_2^{1/2}<\frac{1}{2K^2},$$

then one can check that ζ_p satisfies the required upper and lower bounds on $(0,\delta_2)\times SM$. □

We now want to use the results in Theorem 3.18 to analyze the behavior of the two-point motion $\{\eta_t(z)\colon t\in[0,\infty)\}$ near the diagonal Δ. On the basis of what the derivative process $\{D\xi_t(\pi v)(v)\colon t\in[0,\infty)\}$ predicts, we should expect that the two-point process on $\hat{M}$ is recurrent when the greatest Lyapunov exponent λ_1 is strictly positive. Indeed, when $\lambda_1>0$ we know that, $\mathscr{W}$-almost surely, $|D\xi_t(\pi v)(v)|$ grows exponentially fast as $t\to\infty$; and so it seems reasonable to guess that the paths $\xi_t(x)$ and $\xi_t(y)$ tend to stay away from one another. In order to test this intuition, recall that $\tilde{\Lambda}'(0)=\lambda_1$. Thus, when $\lambda_1>0$, the convexity of $\tilde{\Lambda}$ combined with the facts that $\tilde{\Lambda}(-N)\geqq 0$ and $\tilde{\Lambda}(0)=0$ imply both that $\tilde{\Lambda}$ has a negative absolute minimum value $-\beta$ which is achieved at some point $-p_0\in(-N,0)$ and that $\tilde{\Lambda}(-\gamma)=0$ for some (unique) $\gamma\in(0,N]$. For $\alpha\in[0,\beta]$, define $p(\alpha)=-\max\{p\in\mathbb{R}\colon \tilde{\Lambda}(p)=-\alpha\}$; in other words, $p(\alpha)$ is the unique $p\in[0, p_0]$ which satisfies $\tilde{\Lambda}(-p)=-\alpha$.

(3.19) **Theorem.** *Given $z=(x,y)\in\hat{M}$ and $R>0$, define*

$$\tau_R(z)=\inf\{t\geqq 0\colon\ \operatorname{dist}(\xi_t(x),\xi_t(y))=R\}.$$

Assuming that the conditions in (2.8) *hold and that* $\lambda_1 > 0$, *there exist* $\delta > 0$ *and* $K < \infty$ *such that whenever* $0 < \mathrm{dist}(x, y) < R < \delta$:

i) $$\mathscr{W}(\tau_R(z) < \infty) = 1,$$

ii) $$\frac{1}{\lambda_1}\log\left(\frac{R}{\mathrm{dist}(x, y)}\right) - K \leqq E^{\mathscr{W}}[\tau_R(z)] \leqq \frac{1}{\lambda_1}\log\left(\frac{R}{\mathrm{dist}(x, y)}\right) + K,$$

iii) $$\frac{1}{K}\left(\frac{R}{\mathrm{dist}(x, y)}\right)^{p(\alpha)} \leqq E^{\mathscr{W}}[\exp(\alpha\tau_R(z))] \leqq K\left(\frac{R}{\mathrm{dist}(x, y)}\right)^{p(\alpha)}, \quad \alpha \in [0, \beta].$$

Moreover, there exists a $k \in (0, 1)$ *such that whenever* $0 < \varepsilon < \mathrm{dist}(x, y) < kR < k\delta$:

iv) $$E^{\mathscr{W}}[\tau_R(z)\exp(\beta\tau_R(z)] = \infty,$$

v) $$\frac{1}{K}\left(\frac{\varepsilon}{\mathrm{dist}(x, y)}\right)^{\gamma} \leqq \mathscr{W}(\tau_\varepsilon(z) < \tau_R(z)) \leqq K\left(\frac{\varepsilon}{\mathrm{dist}(x, y)}\right)^{\gamma}$$

Proof. The authors take this opportunity to thank R.J. Williams for assistance with a preliminary version of this theorem.

Let $\{(r_t, \theta_t): t \in [0, \infty)\}$ be the polar coordinate expression for two-point motion $\{\eta_{t \wedge \tau_{\delta_0(z)}}: t \in [0, \infty)\}$ stopped at time $\tau_{\delta_0}(z)$; and let $\phi_p^{\pm}$, $\psi^{\pm}$, $K < \infty$ and $\delta > 0$ be the quantities in Theorem 3.18 corresponding to $[a, b] = [-\gamma, 0]$.

Setting $f(t, r, \theta) = \exp(\beta t)\,\phi^{-}_{-p(\beta)}(r, \theta)$, we have, from i) in Theorem 3.18, that

$$\left(\frac{\partial}{\partial t} + L^{(2)}\right) f \leqq 0$$

on $\mathbb{R}^1 \times (0, \delta) \times SM$; and so, by Itô's formula, for any $0 < \varepsilon < \mathrm{dist}(x, y) < R < \delta$, we have that for all bounded stopping times τ which are dominated by $\tau_R(z) \wedge \tau_\varepsilon(z)$

$$E^{\mathscr{W}}[f(\tau, r_\tau, \theta_\tau)] \leqq f(0, r_0, \theta_0)$$

where $r_0 = \mathrm{dist}(x, y)$ and $\theta_0 \in S_x M$ satisfies $y = \exp_x(r_0\,\theta_0)$. Hence, by the upper and lower bounds on $\phi^{-}_{-p(\beta)}$, we see that

$$E^{\mathscr{W}}[\exp(\beta t), \tau_\varepsilon(z) \wedge \tau_R(z) > t] \leqq K^2\left(\frac{R}{\mathrm{dist}(x, y)}\right)^{p(\beta)} \tag{3.20}$$

and

$$E^{\mathscr{W}}[\exp(\beta\tau_\varepsilon(z)), \tau_\varepsilon(z) < \tau_R(z)] \leqq K^2\left(\frac{\varepsilon}{\mathrm{dist}(x, y)}\right)^{p(\beta)} \tag{3.21}$$

Together, (3.20) and (3.21) imply that $\mathscr{W}(\tau_\varepsilon(z) \wedge \tau_R(z) < \infty) = 1$ and that $\mathscr{W}(\tau_\varepsilon(z) < \tau_R(z)) \to 0$ as $\varepsilon \searrow 0$. Hence i) has been proved.

To prove ii), we use ii) in Theorem 3.18 and Itô's formula to show that

$$\pm E^{\mathscr{W}}[\psi^{\pm}(r_\tau, \theta_\tau) - \lambda_1\tau] \geqq \pm\psi^{\pm}(r_0, \theta_0)$$

for any bounded stopping time τ which is dominated by $\tau_\varepsilon(z)\wedge\tau_R(z)$. Hence, using (3.20) and (3.21), we get

$$\pm E^{\mathscr{W}}[\psi^{\pm}(r_{\tau_R(z)},\theta_{\tau_R(z)})-\lambda_1\tau_R(z)]\geqq\pm\psi^{\pm}(r_0,\theta_0)$$

upon letting $t\to\infty$ and $\varepsilon\searrow 0$. Thus, by the upper and lower bounds on ψ^+ and ψ^-, we now arrive at the estimates in ii). A similar argument using $\pm e^{\alpha t}\phi^{\pm}_{-p(\alpha)}$ yields iii) when $\alpha<\beta$; and the case when $\alpha=\beta$ results from this by the monotone converge theorem.

To prove iv), let $\alpha<\beta$ be given. Then $\tilde{\Lambda}'(-p(\alpha))>0$ and, arguing as we did above, we can use i) and iii) of Theorem 3.18 to conclude that

$$\begin{aligned}&\tilde{\Lambda}'(-p(\alpha))\,E^{\mathscr{W}}[\tau_R(z)\exp(\alpha\tau_R(z))\,\phi^{+}_{-p(\alpha)}(r_{\tau_R(z)},\theta_{\tau_R(z)})]\\&\quad+E^{\mathscr{W}}[\exp(\alpha\tau_R(z))\,\zeta_{-p(\alpha)}(r_{\tau_R(z)},\theta_{\tau_R(z)})]\geqq\zeta_{-p(\alpha)}(r_0,\theta_0).\end{aligned}$$

Using the bounds on $\phi^{+}_{-p(\alpha)}$ and $\zeta_{-p(\alpha)}$ together with the estimate on $E^{\mathscr{W}}[\exp(\alpha\tau_R(z))]$ already obtained, we see that

$$\begin{aligned}&\tilde{\Lambda}'(-p(\alpha))\,KR^{-p(\alpha)}E^{\mathscr{W}}[\tau_R(z)\exp(\alpha\tau_R(z))]\\&\quad\geqq\tilde{\Lambda}'(-p(\alpha))\,E^{\mathscr{W}}[\tau_R(z)\exp(\alpha\tau_R(z))\,\phi^{+}_{-p(\alpha)}(r_{\tau_R(z)},\theta_{\tau_R(z)})]\\&\quad\geqq\zeta_{-p(\alpha)}(r_0,\theta_0)-E^{\mathscr{W}}[\exp(\alpha\tau_R(z))\,\zeta_{-p(\alpha)}(r_{\tau_R(z)},\theta_{\tau_R(z)})]\\&\quad\geqq\frac{1}{K}\left(\log\frac{1}{\operatorname{dist}(x,y)}\right)(\operatorname{dist}(x,y))^{-p(\alpha)}-K^2\left(\frac{R}{\operatorname{dist}(x,y)}\right)^{p(\alpha)}K\left(\log\frac{1}{R}\right)R^{-p(\alpha)}.\end{aligned}$$

Hence

$$\begin{aligned}E^{\mathscr{W}}[\tau_R(z)\exp(\alpha\tau_R(z))]\geqq&(K^2\tilde{\Lambda}'(-p(\alpha)))^{-1}\left(\frac{R}{\operatorname{dist}(x,y)}\right)^{p(\alpha)}\\&\times\left[\left(\log\frac{1}{\operatorname{dist}(x,y)}\right)-K^4\left(\log\frac{1}{R}\right)\right].\end{aligned}$$

Because $p(\alpha)\to p(\beta)$ and $\tilde{\Lambda}'(-p(\alpha))\to 0$ as $\alpha\to\beta$, we now see that iv) holds with $k=\exp(-K^4)$.

Finally, to prove v), we apply the same sort of reasoning to $\phi^{+}_{-\gamma}$ to obtain

$$\begin{aligned}&\mathscr{W}(\tau_\varepsilon(z)<\tau_R(z))\,E^{\mathscr{W}}[\phi^{+}_{-\gamma}(r_{\tau_\varepsilon(z)},\theta_{\tau_\varepsilon(z)})\,|\,\tau_\varepsilon(z)<\tau_R(z)]\\&\quad+\mathscr{W}(\tau_\varepsilon(z)\geqq\tau_R(z))\,E^{\mathscr{W}}[\phi^{+}_{-\gamma}(r_{\tau_R(z)},\theta_{\tau_R(z)})\,|\,\tau_\varepsilon(z)\geqq\tau_R(z)]\geqq\phi^{+}_{-\gamma}(r_0,\theta_0).\end{aligned}$$

Using the upper and lower bounds on $\phi^{+}_{-\gamma}$ and then rearranging terms in the preceding, we arrive at

$$E^{\mathscr{W}}(\tau_\varepsilon(z)<\tau_R(z))\geqq\frac{K^{-2}(\operatorname{dist}(x,y))^{-\gamma}-R^{-\gamma}}{\varepsilon^{-\gamma}-R^{-\gamma}}; \tag{3.22}$$

and a similar argument applied to $-\phi^-_{-\gamma}$ yields

$$E^{\mathscr{W}}(\tau_\varepsilon(z)<\tau_R(z)) \leqq \frac{K^2(\operatorname{dist}(x,y))^{-\gamma}-R^{-\gamma}}{\varepsilon^{-\gamma}-R^{-\gamma}}. \tag{3.23}$$

Clearly (3.22) and (3.23) together imply the existence of a $k \in (0,1)$ for which v) holds. □

(3.24) **Corollary.** *Under the conditions of Theorem* 3.19, *there exist* $\delta>0$, $K<\infty$, *and* $k\in(0,1)$ *for which*

$$\frac{1}{K}\left(\frac{\varepsilon}{\operatorname{dist}(x,y)}\right)^\gamma \leqq E^{\mathscr{W}}\left[\int_0^{\tau_R(z)} \chi_{(0,\varepsilon]}(\operatorname{dist}(\xi_t(x),\xi_t(y))\,dt\right] \leqq K\left(\frac{\varepsilon}{\operatorname{dist}(x,y)}\right)^\gamma,$$

whenever $0<\varepsilon<\operatorname{dist}(x,y)<kR<k\delta$.

Proof. Let δ, K, and k be chosen as in Theorem 3.19 and choose K_1 so that $\log K_1 = 2\lambda_1 K$. Then

$$\begin{aligned} &E^{\mathscr{W}}\left[\int_0^{\tau_R(z)} \chi_{(0,\varepsilon]}(r_t)\,dt\right] \\ &\geqq \mathscr{W}(\tau_{\varepsilon/K_1}(z)<t_R(z))\,E^{\mathscr{W}}[\inf\{t>0: r_{t+\tau_{\varepsilon/K_1}}(z)=\varepsilon\}\,|\,\tau_{\varepsilon/K_1}(z)<\tau_R(z)] \\ &\frac{1}{K}\left(\frac{\varepsilon}{K_1\operatorname{dist}(x,y)}\right)^\gamma\left(\frac{1}{\lambda_1}(\log K_1)-K\right) = K_1^\gamma\left(\frac{\varepsilon}{\operatorname{dist}(x,y)}\right)^\gamma, \end{aligned}$$

where we have used ii) and v) of Theorem 3.19. Clearly this proves the left hand side of the inequality when K is replaced by $K_1^\gamma = \exp(2\lambda_1\gamma K)$.

To prove the right hand side, set

$$f(\varepsilon) = \sup_{\operatorname{dist}(x,y)=\varepsilon} E^{\mathscr{W}}\left[\int_0^{\tau_R(z)} \chi_{(0,\varepsilon]}(r_t)\,dt\right]$$

and let $K_2=(2K)^{1/\gamma}$. If $\operatorname{dist}(x,y)<\dfrac{k}{K_2}R$, then $K_2\varepsilon<kR$ and so

$$\begin{aligned} f(\varepsilon) &\leqq \sup_{\operatorname{dist}(x,y)=\varepsilon} E^{\mathscr{W}}[\tau_{K_2\varepsilon}(z)] + \sup_{\operatorname{dist}(x,y)=K_2\varepsilon} \mathscr{W}(\tau_\varepsilon(z)<\tau_R(z)) f(\varepsilon) \\ &\leqq \frac{1}{\lambda_1}\log K_2 + K(1+K_2^{-\gamma} f(\varepsilon)), \end{aligned}$$

where we have again used ii) and v) of Theorem 3.19. Hence, $f(\varepsilon) \leqq 2\left(\dfrac{1}{\lambda_1}\log K_2 + K\right)$ and therefore

$$E^{\mathscr{W}}\left[\int_0^{\tau_R(z)} \chi_{(0,\varepsilon]}(r_t)\,dt\right] \leqq \mathscr{W}(\tau_\varepsilon(z)<\tau_R(z)) f(\varepsilon) \leqq 2K\left(\frac{\varepsilon}{\operatorname{dist}(x,y)}\right)^\gamma\left(\frac{1}{\lambda_1}\log K_2+K\right). \quad \square$$

(3.25) **Corollary.** *Under the same assumptions,*

$$\mathscr{W}(\overline{\lim_{t\to\infty}}\ \mathrm{dist}(\xi_t(x), \xi_t(y))>0)=1$$

for all $(x, y)\in\hat{M}$.

Proof. This is an immediate consequence of i) of Theorem 3.19. □

(3.26) *Remark.* The local stable manifold theorem for stochastic flows of diffeomorphisms (cf. [7, Theorem 2.1.1]) implies, under much weaker conditions than those in (2.8), that if $\lambda_1<0$ then for all $\varepsilon>0$ and m_L-almost every $x\in M$ there is a $\delta>0$ such that

$$\mathscr{W}(\mathrm{dist}(\xi_t(x), \xi_t(y))\to 0 \text{ as } t\to 0)\geqq 1-\varepsilon$$

whenever $\mathrm{dist}(x, y)<\delta$. However, the local stable manifold theorem does not apply to cases when $\lambda_1\geqq 0$, and this fact has motivated the analysis contained in the present section. In this connection, it should be clear that the functions constructed in Theorem 3.19 can also be used to study the cases when $\lambda_1=0$ or $\lambda_1<0$.

(3.27) *Remark.* It should be clear that the results obtained in Theorem 3.19 and its corollaries are equally valid for the hitting times of the process $\{|D\xi_t(\pi v)(v)|\ t\in[0, \infty)\}$. In this case, Theorem 3.18 is irrelevant and we replace $L^{(2)}$ by $\mathbf{T}L$ and $\phi_p^{\pm}$ by ϕ_p in the proof of Theorem 3.19; similar changes must be made to $\psi_p^{\pm}$ and ζ_p.

4. More on the Two Point Motion

We continue our discussion of the two point mention, only here we will not restrict our attention to what happens in a neighborhood of the diagonal.

(4.1) **Proposition.** *Either there is a non-empty compact* $C\subset\hat{M}$ *which is invariant under the two-point motion or for every* $\delta>0$ *there is an* $\alpha_\delta>0$ *such that*

$$\sup_{z=(x,y)\in\hat{M}:\ \mathrm{dist}(x,y)\geqq\delta} E^{\mathscr{W}}[\exp(\alpha_\delta\,\tau_\delta(z))]<\infty, \tag{4.2}$$

where $\tau_s(z)$ *is defined as in Theorem* 3.19.

Proof. Given $\delta>0$ and $z=(x, y)\in\hat{M}$, define $\sigma_\delta(z)=\inf\{t\geqq 0:\ \mathrm{dist}(\xi_t(x), \xi_t(y))<\delta\}$. It is then clear that $z\mapsto\sigma_\delta(z)$ is an upper semi-continuous function and therefore that $C(\delta)\equiv\{z\in\hat{M}:\ \mathscr{W}(\sigma_\delta(z)<\infty)=0\}$ is closed. Moreover, since $C(\delta)\subseteq\{z\in\hat{M}:\ \mathrm{dist}(x, y)\geqq\delta/2\}$, we see that $C(\delta)$ is compact in $\hat{M}$. To see that $C(\delta)$ is invariant under the two-point motion, let $\zeta_\delta(z)$ denote the first exit time of $\eta_\cdot(z)=(\xi_\cdot(x), \xi_\cdot(y))$ from $C(\delta)$. If $\mathscr{W}(\zeta_\delta(z)<\infty)>0$ for some $z\in C(\delta)$, then we would have that

$$\mathscr{W}(\sigma_\delta(z)<\infty)=\mathscr{W}(\sigma_\delta(\eta_{\zeta_\delta(z)}(z))<\infty \quad\text{and}\quad \zeta_\delta(z)<\infty)>0,$$

which clearly cannot be. Hence, for each $\delta>0$ the set $C(\delta)$ is a compact subset of $\hat{M}$ and is invariant under the two-point motion.

To complete the proof, we need only show that if $C(\delta)=\emptyset$ for some $\delta>0$ then there is an $\alpha_\delta>0$ for which (4.2) holds. But, again by upper semi-continuity, if $C(\delta)$ is empty, then there exist $T\in(0,\infty)$ and $\varepsilon\in(0,1)$ such that $\mathscr{W}(\sigma_\delta(z)\leqq T)\geqq\varepsilon$ for every $z=(x,y)\in\hat{M}$ with $\operatorname{dist}(x,y)\geqq\delta$. Hence, by induction $\mathscr{W}(\sigma_\delta(z)<nT)\leqq(1-\varepsilon)^n$ for all $n\in\mathbf{Z}^+$; and so the existence of α_δ is assured. □

Obviously, if the $C\neq\emptyset$ of Proposition 4.1 exists, then the two-point motion admits an invariant measure which is concentrated on C and there is no hope that the two-point motion started in C will ever approach the diagonal. Thus, from now on, we will be assuming that no such C exists. Note that, by the support theorem for diffusions, the non-existence of C is equivalent to the statement that for each $z\in\hat{M}$ and $\delta>0$ there is control $u\in C([0,\infty);\mathbb{R}^d)$ such that the path $\boldsymbol{\Psi}(\cdot,z;u)$ determined by

$$\dot{\boldsymbol{\Psi}}(t,z;u)=\sum_{k=1}^{d} u_k(t)\,X_k^{(2)}(\boldsymbol{\Psi}(t,z;u))+X_0^{(2)}(\boldsymbol{\Psi}(t,z;u)) \quad \text{with } \boldsymbol{\Psi}(0,z;u)=z$$

gets to within a distance δ of Δ. Thus, by Proposition 4.1, the existence of such controls means that for each $\delta>0$ (4.2) holds for some $\alpha_\delta>0$.

We now want to apply Khas'minskii's construction (cf. [15]) of invariant measures to the two-point motion. For $0<\delta_1<\delta_2$, set $U_i=\{(x,y)\in\hat{M}: \operatorname{dist}(x,y)<\delta_i\}$ and let $\Gamma_i=\partial U_i$. Using induction, define

$$\begin{aligned}\sigma_0(z)&=\inf\{t\geqq 0:\ \eta_t(z)\in\Gamma_1\},\\ \sigma_n'(z)&=\inf\{t\geqq\sigma_n(z):\ \eta_t(z)\in\Gamma_2\},\\ \sigma_{n+1}(z)&=\inf\{t\geqq\sigma_n'(z):\ \eta_t(x)\in\Gamma_1\}.\end{aligned}$$

Under the assumptions made in (2.8), $\lambda_1>0$ and (4.2) imply that the times $\sigma_n(z)$ and $\sigma_n'(z)$ are almost surely finite so long as δ_2 is less than the δ appearing in Theorem 3.19. In fact, one then has that

$$\sup_{z\in\hat{M}} E^{\mathscr{W}}[\sigma_1(z)-\sigma_0(z)]<\infty.$$

For $n\geqq 0$ and $z\in\Gamma_1$, set $Z_n(z)=\eta_{\sigma_n}(z)$. It is then easy to see that $\{Z_n:\ n\geqq 0\}$ is a Markov chain with transition function $\Pi(z,\cdot)=\mathscr{W}\circ(Z_1(z))^{-1}$. In order to make sure that this Markov chain is ergodic, we will assume that

$$\text{(4.3)}\qquad \operatorname{Lie}(\mathbf{T}X_1,\ldots,\mathbf{T}X_d)(v)=\mathbf{T}_v\mathbf{T}M,\quad v\in\mathbf{T}M \quad\text{with } |v|\neq 0.$$

It should be noted that (4.3) is a considerably stronger assumption than that made in (2.8). In the first place, it places conditions on the entire derivative process $D\xi_t(\pi v)(v)$, not just the angular part $\dfrac{D\xi_t(\pi v)(v)}{|D\xi_t(\pi v)(v)|}$. Secondly, and perhaps more seriously, the vector field $\mathbf{T}X_0$ is excluded from the generators of the Lie algebra.

(4.4) **Lemma.** *Assume that* (4.3) *holds. Then there exists a* $\bar{\delta}>0$ *such that*

$$\mathrm{Lie}(X_1^{(2)}, \ldots, X_d^{(2)})(z)=\mathbf{T}_z\hat{M}$$

for all $z=(x, y)\in\hat{M}$ *with* $\mathrm{dist}(x, y)<\bar{\delta}$. *If, in addition,* $\lambda_1>0$, *then* $\bar{\delta}>0$ *can be chosen so that whenever* $0<\delta_2<\bar{\delta}$ *the transition function* $\Pi(z, \cdot)$ *for the Markov chain* $\{Z_n(z)\colon n\geqq 0\}$ *admits a unique invariant measure* $\tilde{\mu}$. *In fact,* $\Pi(z, \cdot)$ *is equivalent to* $\tilde{\mu}$ *for each* $z\in\Gamma_1$.

Proof. To check the assertion about the Lie algebra generated by the $X_k^{(2)}$'s, note that (4.3) is equivalent to the statement that for any $\theta\in\mathbf{S}M$, $u\in\mathbf{T}_\theta\,\mathbf{S}M$, and $s\in\mathbb{R}^1$ with $s^2+|u|^2\neq 0$, there is a $V\in\mathrm{Lie}(X_1, \ldots, X_d)$ for which

$$\langle u, \tilde{V}(\theta)\rangle+s\langle\nabla V(\pi\theta)(\theta), \theta\rangle\neq 0.$$

Since

$$\{(\theta, u, s)\colon\ \theta\in\mathbf{S}M, u\in\mathbf{T}_\theta\,\mathbf{S}M, s\in\mathbb{R}^1, \text{and } s^2+|u|^2=1\}$$

is compact, it follows that there is an $m\in\mathbf{Z}^+$ and vector fields $V_1, \ldots, V_m\in\mathrm{Lie}(X_1, \ldots, X_d)$ such that

$$\sum_{l=1}^{m}|\langle u, \tilde{V}_l(\theta)\rangle+s\langle\nabla V_l(\pi\theta)(\theta), \theta\rangle|\geqq\varepsilon$$

for some $\varepsilon>0$ and all $\theta\in\mathbf{S}M$, $u\in\mathbf{T}_\theta\,\mathbf{S}M$, and $s\in\mathbb{R}^1$ with $s^2+|u|^2=1$. The required result about $\mathrm{Lie}(X_1^{(2)}, \ldots, X_d^{(2)})$ now follows by applying Lemma 3.6 to each of the vector fields $V_1^{(2)}, \ldots, V_m^{(2)}$.

To prove the assertion when $\lambda_1>0$, first note that, for every Borel set $A\subset\Gamma_1$, the function

$$z\mapsto\mathcal{W}(\sigma_0(z)\in A)$$

is $L^{(2)}$-harmonic in $\hat{M}\setminus\Gamma_1$ and therefore is smooth on that part of $\hat{M}\setminus\Gamma_1$ where $\mathrm{Lie}(X_1^{(2)}, \ldots, X_d^{(2)})$ has full rank. Combining this with Bony's strong maximum principle (cf. [6, Theorem 3.1]) and noting that $\mathbf{S}M$ is connected since M itself is, one sees that for any given A either $\Pi(\cdot, A)\equiv 0$ or it is uniformly positive on Γ_1. From these observations and the standard ergodic theory for Markov chains, the existence of a $\tilde{\mu}$ with the required properties is immediate. □

(4.6) **Theorem.** *Assume that* (4.2) *and* (4.3) *hold and that* $\lambda_1>0$. *Then there exists a unique probability measure* μ *on* $\hat{M}$ *such that*

$$\mathcal{W}\left(\frac{1}{T}\int_0^T f(\eta_t(z))\,dt\to\int_{\hat{M}} f\,d\mu \quad as \quad t\to\infty\right)=1$$

for every bounded measurable $f\colon\hat{M}\to\mathbb{R}^1$ *and* $z\in\hat{M}$. *In particular,* μ *is the unique probability measure on* $\hat{M}$ *for which two-point motion is invariant. Moreover, if* $\gamma\in(0, N]$ *is the number defined just before Theorem* 3.19, *then*

$$\frac{1}{K}\varepsilon^\gamma\leqq\mu((x, y)\in\hat{M}\colon\ \mathrm{dist}(x, y)<\varepsilon)\leqq K\varepsilon^\gamma$$

and

$$\mathscr{W}\left(\frac{1}{K}\varepsilon^{\gamma} \leqq \lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_{(0,\varepsilon]}(\operatorname{dist}(\xi_t(x),\xi_t(y))\,dt \leqq K\varepsilon^{\gamma}\right)=1$$

for some $K<\infty$ and all sufficiently small $\varepsilon>0$.

Proof. The passage from the $\tilde{\mu}$ of Lemma 4.5 to the existence of μ is carried out in [15]: one simply takes μ to be the normalization of the measure

$$A\in\mathscr{B}_{\hat{M}}\mapsto \int_{\Gamma_1} E^{\mathscr{W}}\left[\int_0^{\sigma_1(z)} \chi_A(\eta_t(z))\,dt\right]\tilde{\mu}(dz).$$

In addition, the asserted ergodic properties follow immediately from the equivalence of the transition functions $\Pi(z,\cdot)$ with the measure $\tilde{\mu}$. Thus, all that remains is to use the results obtained in Corollary 3.24 to obtain the asserted estimates on μ. □

(4.7) *Remark.* There are several comments which should be made about the preceding. In the first place, the condition (4.3) is certainly more than one needs to get the conclusions in Theorem 4.6. Secondly, the ergodicity asserted in Theorem 4.6 is not a consequence of Doeblin's theorem since, by Theorem 3.19,

$$E^{\mathscr{W}}[\tau_\delta(z)]\to\infty$$

as $z\to\Delta$; and therefore it must be true that

$$\sup_{z\in\hat{M}} \mathscr{W}(\operatorname{dist}(\xi_t(x),\xi_t(y))<\delta)=1$$

for every $t>0$. Finally, it is interesting to compare the invariant measure μ just constructed for the two-point motion to the measure m_L which is invariant for the one-point motion. Indeed, only when $\gamma=N$, in which case Corollary 2.14 says that m_L is $\mathscr{W}$-almost surely preserved by the flow ξ_t, is $\mu=m_L^2|_{\hat{M}}$. When $0<\gamma<N$, the measure μ will have more mass near Δ than will m_L^2.

Appendix

The purpose of this appendix is to derive a somewhat more closed expression for the Donsker-Varadhan rate function $J(\mu)$ when the diffusion generator is a tame perturbation of a hypoelliptic symmetric operator (see Theorem A.8 below). Similar results, in a slightly different context, were investigated previously by R. Pinsky in [19]. We begin by recalling some basic facts about the symmetric case.

Let M be a compact manifold and m a positive smooth probability measure on M. Supose that $\{X_1,\ldots,X_d\}\subseteq\Gamma(\mathbf{T}M)$ are given and assume that

(A.1) $$\operatorname{Lie}(X_1,\ldots,X_d)(x)=\mathbf{T}_xM,\qquad x\in M.$$

Define $\mathscr{L}^0 = -\frac{1}{2}\sum_{k=1}^{d} X_k^* X_k$ on $C^\infty(M)$, where X^* is used to denote the formal adjoint of $X \in \Gamma(\mathbf{T}M)$ relative to m; and denote by $P^0(t, x, \cdot)$ the transition probability function for the diffusion on M determined by $\mathscr{L}^0$. Then, by Hörmander's Theorem, $P^0(t, x, dy) = p^0(t, x, y)\, m(dy)$, where $p^0 \in C^\infty((0, \infty) \times M \times M)$. In addition, because $\mathscr{L}^0$ is symmetric in $L^2(m)$, $p^0(t, x, y) = p^0(t, y, x)$ for all $(t, x, y) \in (0, \infty) \times M \times M$; and therefore the semigroup $\{P_t^0 : t > 0\}$ on $C(M)$ given by

$$P_t^0\, \phi(x) = \int_M \phi(y)\, P^0(t, x, dy)$$

has the property that $\|P_t^2 \phi\|_{L^2(m)} \leqq \|\phi\|_{L^2(m)}$, $t > 0$. In particular, there is a unique strongly continuous contraction semigroup $\{\overline{P_t^0} : t > 0\}$ on $L^2(m)$ such that $\overline{P_t^0}|_{C(M)} = P_t^0$, $t > 0$. Use $\overline{\mathscr{L}^0}$ to denote the generator of $\{\overline{P_t^0} : t > 0\}$. Clearly, $\overline{\mathscr{L}^0}$ is non-positive and self-adjoint. Set

$$\mathscr{E}(\phi, \phi) = \|(-\overline{\mathscr{L}^0})^{1/2} \phi\|_{L^2(m)}^2$$

for $\phi \in \mathrm{Dom}(\mathscr{E}) \equiv \mathrm{Dom}((-\overline{\mathscr{L}^0})^{1/2})$ and $\mathscr{E}(\phi, \phi) = \infty$ for $\phi \in L^2(m) \backslash \mathrm{Dom}(\mathscr{E})$. Finally, given $X \in \Gamma(\mathbf{T}M)$ let $\bar{X}$ denote the minimal closure of X as an operator on $L^2(m)$ (i.e. $\bar{X} = (X^*)^*$).

(A.2) **Lemma.** $\mathrm{Dom}(\mathscr{E}) = \bigcap_1^d \mathrm{Dom}(\bar{X}_k)$. *In particular, for every* $\phi \in \mathrm{Dom}(\mathscr{E})$:

$$\mathscr{E}(\phi, \phi) = \frac{1}{2}\sum_1^d \int_M (\bar{X}_k \phi)^2\, dm$$

and there is a sequence $\{\phi_n\}_1^\infty \subseteq C^\infty(M)$ *such that*

$$\lim_{n \to \infty} \int_M \left[(\phi_n - \phi)^2 + \sum_1^d (X_k \phi_n - \bar{X}_k \phi)^2\right] dm = 0.$$

Finally, if $g \in \mathrm{Dom}(\mathscr{E})$, *then* $\bar{X}_k(g^2) = 2g\bar{X}_k g$, $1 \leqq k \leqq d$, *in the sense that*

$$\int_M g^2 X_k^* \psi\, dm = \int_M \psi(2g\bar{X}_k g)\, dm$$

for all $\psi \in C^\infty(M)$.

Proof. First note that, for each $t > 0$, $\overline{P_t^0}$ maps $L^2(m)$ into $C^\infty(M)$ and therefore that $\overline{\mathscr{L}^0}$ is the minimal closure in $L^2(m)$ of $\mathscr{L}^0$ on $C^\infty(M)$. Thus, since $\mathrm{Graph}((-\overline{\mathscr{L}^0})^{1/2}|_{\mathrm{Dom}(\mathscr{L}^0)})$ is dense in $\mathrm{Graph}((-\overline{\mathscr{L}^0})^{1/2})$, we see that forevery $\phi \in \mathrm{Dom}(\mathscr{E})$ there is a sequence $\{\phi_n\}_1^\infty \subseteq C^\infty(M)$ such that

$$\lim_{n \to \infty} \|\phi_n - \phi\|_{L^2(m)} + \|(-\overline{\mathscr{L}^0})^{1/2} \phi_n - (-\overline{\mathscr{L}^0})^{1/2} \phi\|_{L^2(m)} = 0.$$

At the same time,

$$E(\phi,\phi)=-\int_M \phi\mathscr{L}^0\phi\,dm=\tfrac{1}{2}\sum_1^d \int_M (X_k\phi)^2\,dm,\qquad \phi\in C^\infty(M).$$

Combining these, we arrive at the stated properties of $\mathscr{E}$.

To prove the final assertion, simply approximate $g\in\mathrm{Dom}(\mathscr{E})$ in the Graph$((-\overline{\mathscr{L}^0})^{1/2})$-norm by g_n's from $C^\infty(M)$. □

We now recall a result due originally to Donsker and Varadhan [11] (cf. Theorem (7.44) of [20]). Namely, define $\mathscr{I}^0(\mu)$ for $\mu\in C(M)^*$ so that

$$\mathscr{I}^0(\mu)=\sup\left\{-\int_M \frac{\mathscr{L}^0 u}{1+u}\,d\mu:\ u\in C^\infty(M)^+\right\}$$

if $\mu\in\mathscr{P}(M)(\equiv\{\mu\in C(M)^*: \mu$ is a probability measure$\})$ and $\mathscr{I}^0(\mu)=\infty$ otherwise.

(A.3) **Lemma.** *For any $\mu\in\mathscr{P}(M)$, $\mathscr{I}^0(\mu)<\infty$ if and only if $\mu\ll m$ and $f^{1/2}\in\mathrm{Dom}(\mathscr{E})$, where $f\equiv\frac{d\mu}{dm}$, in which case $\mathscr{I}^0(\mu)=\mathscr{E}(f^{1/2},f^{1/2})$.*

Next, suppose that

$$\Sigma=\begin{bmatrix}\sigma_1\\ \vdots\\ \sigma_d\end{bmatrix}\in(C^\infty(M))^d$$

and define

$$\mathscr{L}^\Sigma=\mathscr{L}^0+\sum_1^d \sigma_k X_k$$

on $C^\infty(M)$. At the same time, define $\mathscr{I}^\Sigma(\mu)$ for $\mu\in C(M)^*$ so that

$$\mathscr{I}^\Sigma(\mu)=\sup\left\{-\int_M \frac{\mathscr{L}^\Sigma u}{1+u}\,d\mu:\ u\in C^\infty(M)^+\right\}$$

if $\mu\in\mathscr{P}(M)$ and $\mathscr{I}^\Sigma(\mu)=\infty$ otherwise.

(A.4) **Lemma.** *For every $\mu\in C(M)^*$,*

$$\tfrac{1}{2}\mathscr{I}^0(\mu)-\sum_1^d\|\sigma_k\|^2\leqq\mathscr{I}^\Sigma(\mu)\leqq 2\mathscr{I}^0(\mu)+\sum_1^d\|\sigma_k\|^2.$$

In particular, if $\mu\in\mathscr{P}(M)$, then $\mathscr{I}^\Sigma(\mu)<\infty$ if and only if $\mu\ll m$ and $f^{1/2}\in\mathrm{Dom}(\mathscr{E})$, where $f\equiv\frac{d\mu}{dm}$.

Proof. Note that, by writing $1+u$ as $ce^{-\phi}$, the expressions for $\mathscr{I}^0(\mu)$ and $\mathscr{I}^\Sigma(\mu)$ become

$$\mathscr{I}^0(\mu)=\sup\left\{\int_M \left[\mathscr{L}^0\phi-\tfrac{1}{2}\sum_1^d (X_k\phi)^2\right]d\mu:\ \phi\in C^\infty(M)\right\}$$

and

$$\mathscr{I}^\Sigma(\mu)=\sup\left\{\int_M \left[\mathscr{L}^\Sigma\phi-\tfrac{1}{2}\sum_1^d (X_k\phi)^2\right]d\mu:\ \phi\in C^\infty(M)\right\}$$

so long as $\mu\in\mathscr{P}(M)$. Hence, for $\mu\in\mathscr{P}(M)$,

$$\begin{aligned}\mathscr{I}^\Sigma(\mu)&=\sup_\phi\left\{\int_M \left[\mathscr{L}^0\phi-\tfrac{1}{4}\sum_1^d (X_k\phi)^2\right]d\mu-\sum_1^d\int_M [(\sigma_k-\tfrac{1}{2}X_k\phi)^2-|\sigma_k|^2]\,d\mu\right\}\\ &\leqq\sup_\phi\left\{\int_M \left[\mathscr{L}^0\phi-\tfrac{1}{4}\sum_1^d (X_k\phi)^2\right]d\mu\right\}+\sum_1^d \|\sigma_k\|^2\\ &=2\mathscr{I}^0(\mu)+\sum_1^d \|\sigma_k\|^2.\end{aligned}$$

The lower bound on $\mathscr{I}^\Sigma(\mu)$ in terms of $\mathscr{I}^0(\mu)$ is obtained in a similar fashion. □

Given $\mu\in\mathscr{P}(M)$, let $\|\cdot\|_\mu$ denote the $L^2(\mu;\ \mathbb{R}^d)$-norm and set $\mathscr{H}(\mu)$ equal to the closure in $L^2(\mu;\ \mathbb{R}^d)$ of $\{\mathbf{X}\phi:\ \phi\in C^\infty(M)\}$, where

$$\mathbf{X}\phi\equiv\begin{bmatrix}X_1\phi\\ \vdots\\ X_d\phi\end{bmatrix}.$$

Also, denote by $\Sigma(\mu)$ the orthogonal projection in $L^2(\mu;\ \mathbb{R}^d)$ of Σ onto $\mathscr{H}(\mu)$. We may and will assume that

$$\Sigma(\mu)=\begin{bmatrix}\sigma_1^\mu\\ \vdots\\ \sigma_d^\mu\end{bmatrix},$$

where the σ_k^μ's are $\mathscr{B}_M$-measurable from M to $\mathbb{R}^1$.

(A.5) **Lemma.** *The map $\mu\in\mathscr{P}(M)\mapsto P^\Sigma(\mu)\equiv\|\Sigma(\mu)\|_\mu^2$ is weak* lower semi-continuous and convex.*

Proof. Simply observe that

$$P^\Sigma(\mu)=\|\Sigma\|_\mu^2-\|\Sigma-\Sigma(\mu)\|_\mu^2$$

and that

$$\|\Sigma-\Sigma(\mu)\|_\mu^2=\inf\{\|\Sigma-\mathbf{X}\phi\|_\mu^2:\ \phi\in C^\infty(M)\}.\quad\square$$

(A.6) **Lemma.** *Let $\mu \in \mathscr{P}(M)$ and assume that $d\mu = f\,dm$ where $f^{1/2} \in \mathrm{Dom}(\mathscr{E})$. Then*

$$\mathscr{I}^{\Sigma}(\mu) = \mathscr{I}^{0}(\mu) + \tfrac{1}{2} P^{\Sigma}(\mu) - (f^{1/2}\Sigma(\mu), \bar{\mathbf{X}} f^{1/2})_m.$$

Proof. Choose $\{\psi_n\}_1^{\infty} \subseteq C^{\infty}(M)$ so that $\|\Sigma(\mu) - \mathbf{X}\psi_n\|_{\mu} \leqq \frac{1}{n}$ for $n \geqq 1$.

From each $n \geqq 1$,

$$\mathscr{I}^{\Sigma}(\mu) = \sup\{\int_M \mathscr{L}^{\Sigma}(\phi + \psi_n)\,d\mu - \tfrac{1}{2}\|\mathbf{X}(\phi+\psi_n)\|_{\mu}^2 \colon \phi \in C^{\infty}(M)\};$$

and for each $\phi \in C^{\infty}(M)$,

$$\begin{aligned}
&\int_M \mathscr{L}^{\Sigma}(\phi+\psi_n)\,d\mu - \tfrac{1}{2}\|\mathbf{X}(\phi+\psi_n)\|_{\mu}^2 \\
&\quad = \int_M \mathscr{L}^{0}\phi\,d\mu - \tfrac{1}{2}\sum \|\mathbf{X}\phi\|_{\mu}^2 - (\mathbf{X}\phi, \mathbf{X}\psi_n)_{\mu} - \tfrac{1}{2}\|\mathbf{X}\psi_n\|_{\mu}^2 \\
&\qquad + (\Sigma(\mu), \mathbf{X}\phi)_{\phi} + (\Sigma(\mu), \mathbf{X}\psi_n)_{\mu} + \int_M \mathscr{L}^{0}\psi_n\,d\mu \\
&\quad = \int_M \mathscr{L}^{0}\phi\,d\mu - \tfrac{1}{2}\|\mathbf{X}\phi\|_{\mu}^2 + \tfrac{1}{2}P^{\Sigma}(\mu) + \int_M \mathscr{L}^{0}\psi_n\,d\mu + \mathscr{O}\left(\frac{1}{n}\right)(1 + \|\mathbf{X}\phi\|_{\mu}).
\end{aligned}$$

At the same time, by Lemma A.2,

$$\begin{aligned}
\int_M \mathscr{L}^{0}\psi_n\,d\mu &= \int_M f\mathscr{L}^{0}\psi_n\,dm = -\sum_1^d (f^{1/2}X_k\psi_n)(\bar{X}_k f^{1/2})\,dm \\
&= -\sum_1^d \int_M f^{1/2}\sigma_k^{\mu}\bar{X}_k f^{1/2}\,dm + \sum_1^d \int_M f^{1/2}(\sigma_k^{\mu} - X_k\psi_n)\bar{X}_k f^{1/2}\,dm \\
&= -\sum_1^d \int_M f^{1/2}\sigma_k^{\mu}\bar{X}_k f^{1/2}\,dm + \mathscr{O}\left(\frac{1}{n}\right).
\end{aligned}$$

Combined with the preceding, this now shows that

$$\begin{aligned}
&\int_M \mathscr{L}^{\Sigma}(\phi+\psi_n)\,d\mu - \tfrac{1}{2}\|\mathbf{X}(\phi+\psi_n)\|_{\mu}^2 \\
&\quad = \int_M \mathscr{L}^{0}\phi\,d\mu - \tfrac{1}{2}\|\mathbf{X}\phi\|_{\mu}^2 + \tfrac{1}{2}P^{\Sigma}(\mu) - (f^{1/2}\Sigma(\mu), \bar{\mathbf{X}}f^{1/2})_m + \mathscr{O}\left(\frac{1}{n}\right)(1 + \|\mathbf{X}\phi\|_{\mu});
\end{aligned}$$

and therefore

$$\begin{aligned}
&\left(1 - \frac{1}{n}\right)\mathscr{I}^{0}(\mu) + \tfrac{1}{2}P^{\Sigma}(\mu) - (f^{1/2}\Sigma(\mu), \bar{\mathbf{X}}f^{1/2})_m - \mathscr{O}\left(\frac{1}{n}\right) \leqq \mathscr{I}^{\Sigma}(\mu) \\
&\quad \leqq \left(1 + \frac{1}{n}\right)\mathscr{I}^{0}(\mu) + \tfrac{1}{2}P^{\Sigma}(\mu) - (f^{1/2}\Sigma(\mu), \bar{\mathbf{X}}f^{1/2})_m + \mathscr{O}\left(\frac{1}{n}\right),
\end{aligned}$$

from which the asserted result is immediate. □

The authors thank S.R.S. Varadhan for the elegant proof of the following lemma.

(A.7) **Lemma.** *Assume that $\mu\in\mathscr{P}(M)$ and that $d\mu=f\,dm$ where $f^{1/2}\in\mathrm{Dom}(\mathscr{E})$ and $f\geqq\varepsilon$ for some $\varepsilon>0$. Then there exists a sequence $\{g_n\}_1^\infty\subseteq C^\infty(M)$ such that*

$$\left\|\mathbf{X}g_n-\frac{\bar{\mathbf{X}}f^{1/2}}{f^{1/2}}\right\|_\mu\to 0$$

as $n\to\infty$.

Proof. Choose $\{u_n\}\subseteq C^\infty(M)^+$ so that

$$-\int_M\frac{\mathscr{L}^0u_n}{1+u_n}\,d\mu\to\mathscr{E}(f^{1/2},f^{1/2}),$$

and set $g_n=\log(1+u_n)$. Then

$$\begin{aligned}\left\|\frac{\bar{\mathbf{X}}f^{1/2}}{f^{1/2}}\right\|_\mu^2&=2\mathscr{E}(f^{1/2},f^{1/2})\\&=\lim_{n\to\infty}-2\int_M\frac{\mathscr{L}^0u_n}{1+u_n}\,d\mu=\lim_{n\to\infty}\left(\bar{\mathbf{X}}\left(\frac{f}{1+u_n}\right),\mathbf{X}u_n\right)_m\\&=\lim_{n\to\infty}\left[(\bar{\mathbf{X}}f,\mathbf{X}g_n)_m-\|\mathbf{X}g_n\|_\mu^2\right]\\&=\lim_{n\to\infty}\left[2\left(\frac{\bar{\mathbf{X}}f^{1/2}}{f^{1/2}},\mathbf{X}g_n\right)_\mu-\|\mathbf{X}g_n\|_\mu^2\right],\end{aligned}$$

where each of these steps is trivial when f is smooth and can be justified in general by an easy approximation procedure. □

(A.8) **Theorem.** *For every $\mu\in\mathscr{P}(M)$,*

$$\mathscr{I}^\Sigma(\mu)=\mathscr{I}^0(\mu)+\tfrac{1}{2}P^\Sigma(\mu)-\tfrac{1}{2}\int_M\mathbf{X}^*\cdot\Sigma\,d\mu,$$

where $\mathbf{X}^*\cdot\Sigma\equiv\sum_1^d X_k^*\,\sigma_k$.

Proof. By Lemma A.3, it suffices to treat the case when $d\mu=f\,dm$ with $f^{1/2}\in\mathrm{Dom}(\mathscr{E})$. In addition, because both sides of (A.9) are weak* lower semi-continuous and convex, we need only prove (A.9) when, in addition, $f\geqq\varepsilon$ for some $\varepsilon>0$, since otherwise we can replace μ by $(1-\varepsilon)\mu+\varepsilon m$ and let $\varepsilon\searrow 0$. Thus, we proceed under these assumptions.

In view of Lemma A.6, all that has to be shown is that

$$\sum_1^d\int_M f^{1/2}\sigma_k^\mu\bar{X}_kf^{1/2}\,dm=\tfrac{1}{2}\int_M\mathbf{X}^*\cdot\Sigma\,d\mu.$$

To this end, choose $\{g_n\}_1^\infty$ as in Lemma A.7. Then

$$\sum_1^d \int_M f^{1/2}\sigma_k^\mu \bar{X}_k f^{1/2}\,dm = \sum_1^d \int_M \sigma_k^\mu \frac{\bar{X}_k f^{1/2}}{f^{1/2}}\,d\mu$$

$$= \lim_{n\to\infty} (\Sigma(\mu), \mathbf{X} g_n)_\mu = \lim_{n\to\infty} (\Sigma, \mathbf{X} g_n)_\mu$$

$$= \left(\Sigma, \frac{\bar{\mathbf{X}} f^{1/2}}{f^{1/2}}\right)_\mu = \sum_1^d \int_M \sigma_k f^{1/2} \bar{X} f^{1/2}\,dm$$

$$= \tfrac{1}{2}\sum_1^d \int_M f X_k^* \sigma_k\,dm = \tfrac{1}{2}\int_M \mathbf{X}^*\cdot\Sigma\,d\mu. \quad \square$$

References

1. Arnold, L., Kliemann, W.: Large deviations of linear stochastic differential equations. In: Engelbert, H.J. Schmidt (ed.) Proceedings of the Fifth IFIP Working Conference on Stochastic Differential Systems, Eisenach 1986 (Lect. Notes Control Inf. Sci. Vol. 96) Berlin Heidelberg New York: Springer 1987
2. Arnold, L., Oeljeklaus, E., Pardoux, E.: Almost sure and moment stability for linear Itô equations. In: Arnold, L., Wihstutz, V. (eds.) Lyapunov exponents (Lect. Notes Math., Vol. 1186, pp. 129–159) Berlin Heidelberg New York: Springer 1986
3. Baxendale, P.: The Lyapunov spectrum of a stochastic flow of diffeomorphisms. In: Arnold, L., Wihstutz, V. (eds.) Lyapunov exponents. (Lect. Notes Math., Vol. 1186, pp. 322–337) Berlin Heidelberg New York: Springer 1986
4. Baxendale, P.: Moment stability and large deviations for linear stochastic differential equations. In: Ikeda, N. (ed) Proceedings of the Taniguchi Symposium on Probabilistic Methods in Mathematical Physics. Katata and Kyoto 1985, pp. 31–54. Tokyo: Kinokuniya 1987
5. Baxendale, P.: Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms. Probab. Theor. Rel. Fields., in press
6. Bony, J.-M.: Principe du maximum, unicité du problem de Cauchy et inéquality de Harnack pour les operateurs elliptiques dégeneres. Ann. Inst. Fourier **19**, 277–304 (1969)
7. Carverhill, A.: Flows of stochastic dynamical systems: ergodic theory. Stochastics **14**, 273–317 (1985)
8. Carverhill, A.: A formula for the Lyapunov numbers of a stochastic flow. Applications to a perturbation theorem. Stochastics **14**, 209–226 (1985)
9. Chappell, M.J.: Bounds for average Lyapunov exponents of gradient stochastic systems. In: Arnold, L., Wihstutz, V. (eds.) Lyapunov exponents. (Lect. Notes Math., Vol. 1186, pp. 308–321) Berlin Heidelberg New York: Springer 1986
10. Cheeger, J., Ebin, D.G.: Comparison theorems in Riemanian geometry. Amsterdam New York: North Holland/American Elsevier 1975
11. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, I. Commun. Pure. Appl. Math. **28**, 1–47 (1975)
12. Elworthy, K.D., Stroock, D.W.: Large deviation theory for mean exponents of stochastic flows. In: Albeverio, S., Blanchard, P., Streit, L. (eds.) Stochastic processes, mathematics and physics (Lect. Notes Math., Vol. 1158, pp. 72–80) Berlin Heidelberg New York: 1986
13. Furstenberg, H.: Noncommuting random products. Trans. Am. Math. Soc. **108**, 377–428 (1963)
14. Kato, T.: Perturbation theory of linear operators, Berlin Heidelberg New York: Springer 1966
15. Khas'menskii, R.K.: Ergodic properties and stabilization of the solution to the Cauchy problem for parabolic equations. Theor. Probab. Appl. **5**, 179–196 (1960)
16. Kusuoka, S., Stroock, D.W.: Applications of the Malliavin calculus, I. In: Itô, K. (ed.) Stochastic analysis, pp. 271–307. Proceedings, Taniguchi Int. Symp. at Katata and Kyoto in 1982. Tokyo Amsterdam: Kinokuniya/North Holland 1984

17. Kusuoka, S., Stroock, D.W.: Applications of the Malliavin calculus, II. J. Fac. Sci. Univ. Tokyo, Sect. IA Math. **32**, 1–76 (1985)
18. Ledrappier, F., Young, L.-S.: Entropy formula for random transformations. Probab. Theor. Rel. Fields **80**, 217–240 (1989)
19. Pinsky, R.: On evaluating the Donsker-Varadhan I-functional. Ann. Probab. **13**, 342–362 (1985)
20. Stroock, D.W.: An introduction of the theory of large deviations. (Universitext Series) New York Heidelberg Berlin: Springer 1984
21. Stroock, D.W., On the rate at which a homogeneous diffusion approaches a limit, an application of the large deviation theory of certain stochastic integrals. Ann. Probab. **14**, 840–859 (1986)

Received May 11, 1987; received in revised form June 10, 1988