Probab. Theory Relat. Fields 94, 525–541 (1993)

Probability Theory and Related Fields

Path integral formulae for heat kernels and their derivatives

J.R. Norris

Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, UK

Received July 27, 1991; in revised form June 1, 1992

Summary. The heat kernel and its derivatives of a vector Laplacian on the sections of a bundle over a compact Riemannian manifold are expressed as products of the scalar heat kernel of the manifold and path integrals over the Brownian bridge. The small-time asymptotics of these integrals are computed.

Mathematics Subject Classification: 58 G 32

1 Introduction

We are concerned with the heat kernel $P_t(x, y)$ of an elliptic second order differential operator

$$\mathscr{L} = \tfrac{1}{2}\operatorname{tr} \nabla^2 + V \tag{1.1}$$

on the sections of a vector bundle F over a compact Riemannian manifold M. Here ∇ is a general connection on F, extended to $F \otimes T^*M$ by the Levi–Civita connection of M, and V is a section of End F. Subordination of the semigroup $P_t = e^{t\mathscr{L}}$ to the scalar heat semigroup of the Laplace–Beltrami operator can be expressed very concretely by means of Wiener integrals over the path space of M:

$$P_t(x, y) = p_t(x, y)\, \mathbb{E}^{x,y,t}(\tau_{0t} e_t)\,. \tag{1.2}$$

Here, $p_t(x, y)$ is the scalar heat kernel, $\mathbb{P}^{x,y,t}$ is the law of the Brownian bridge from x to y in time t, τ_{0t} is parallel translation $F_{x_t} \to F_{x_0}$ according to ∇ along the coordinate process x_t, and e_t, a process in End F over x_t, is defined by the linear covariant equation

$$De_t = e_t V(x_t)\, \partial t, \quad e_0 = \mathrm{id}\,. \tag{1.3}$$

This formula is well known, at least when ∇ is compatible with a metric on F. See for example [9]. It has been used to examine the small time asymptotics of $P_t(x, y)$.

In this paper we obtain formulae of the same type for all the derivatives of $P_t(x, y)$. To do this we develop further a method of Bismut [3] who showed that, for the scalar heat kernel,

$$t\nabla_x p_t(x, y) = p_t(x, y)\,\mathbb{E}^{x,y,t}(\tau_{0t} Q_t) \tag{1.4}$$

for a certain process Q_t in T^*M over x_t defined below. This formula permitted Bismut to show, for y close to x, that as $t \downarrow 0$

$$t\nabla_x \log p_t(x, y) \to \dot{\gamma}(0)$$

where γ is the unique geodesic from x to y in time 1. We have a similar result for the general case.

Here is a sketch of the method. We use the Eells–Elworthy stochastic development to construct, from a single Brownian motion in $\mathbb{R}^n$, a family of processes $x_t(x)$ in M depending smoothly on $x \in M$, each $x_t(x)$ being a Brownian motion starting from x. Then, formally

$$P_t(x, y) = \mathbb{E}[\tau_{0t}(x)\, e_t(x)\, \delta_y(x_t(x))]$$

so

$$\nabla_x P_t(x, y) = \mathbb{E}[\,\nabla_x \tau_{0t}(x)\, e_t(x)\, \delta_y(x_t(x))]\ .$$

We use integration by parts in path space to remove the derivative from the delta function, thus obtaining a formula for the first derivative. This procedure can be repeated as often as one likes.

Another method, which does not require integration by parts, has recently been used by Elworthy [6] to obtain, independently, heat kernel formulae of a similar type.

The formulae we obtain are not in any sense unique and, as the number of derivatives increases, they are not particularly simple. However our method of integrating by parts gives the simplest formulae we can find. Each formula, in common with (1.2) and (1.4), involves an average over the Brownian bridge of certain processes defined by covariant stochastic differential equations. The coefficients of these equations involve only the curvature of ∇, the potential V and their derivatives. Thus the formulae, though complicated, provide some insight as to how the given data in $\mathscr{L}$ affect the heat kernel and its derivatives. Also, the asymptotic result for small time shows that the formulae have some analytic power.

2 Statement of the formulae and small-time asymptotics

Let $T \in (0, \infty)$ and $x \in M$ be given. We are working towards formulae for $\nabla_x P_T(x, y)$, $\nabla_x^2 P_T(x, y)$, etc. Write $\mathbb{P}^x$ for the Wiener measure on the path space making the coordinate process x_t a Brownian motion starting from x. The fibres of F are taken to be isomorphic to a fixed finite-dimensional vector space E. We denote by u_t the horizontal lift of x_t in $\mathrm{GL}(E, F)$ starting from u_0. (We use the same notation u_t for a horizontal lift of x_t in $\mathrm{GL}(\mathbb{R}^n, TM)$.) Then $\tau_{st} := u_s u_t^{-1} : F_{x_t} \to F_{x_s}$ is the parallel translation. The connection ∇ is extended to tensor products in the usual way. The corresponding horizontal lift, for processes z_t in $\mathrm{End}\,F$ over x_t say, is provided by $u_t^{-1} \cdot z_t = u_t^{-1} z_t u_t$, where the dot is used to avoid confusion with a simple composition of linear maps.

We introduce some new bundle valued processes over x_t by covariant stochastic differential equations

$$(Dz_t)_i = \tfrac{1}{2} R(\partial x_t, (z_t)_i)\,\partial x_t\,, \qquad z_0 = \mathrm{id.} \in \mathrm{End}\, TM \tag{2.1}$$

$$X_t = (T-t)\, z_t \tag{2.2}$$

$$(DW_t)_i = R(\partial x_t, (X_t)_i)\,, \qquad W_0 = 0 \in \mathrm{End}\, F \otimes T^*M \tag{2.3}$$

$$(DQ_t)_i = (z_t^*\,\partial x_t)^i\,, \qquad Q_0 = 0 \in T^*M\,. \tag{2.4}$$

In (2.1), $R \in \Gamma(\mathrm{End}\, T^*M \otimes T^*M \otimes T^*M)$ denotes the Riemannian curvature and in (2.3) $R \in \Gamma(\mathrm{End}\, F \otimes T^*M \otimes T^*M)$ is the curvature of ∇. We shall continue to write R for the curvature of ∇ and of the Levi–Civita connection, and any of their tensor products. We write ∂x_t for the Stratonovich differential and $Dz_t := u_t.\partial(u_t^{-1} z_t)$ for the covariant Stratonovich differential. The calculus of these differentials is developed in [9]; see also [5]. The subscript i refers to an orthonormal basis in T^*M, the superscript i to its dual basis in TM, and $z_t^* \in \mathrm{End}\, TM$ is the adjoint of z_t. Since x_t is a Brownian motion, we have

$$R(\partial x_t, \cdot)\,\partial x_t = -\, S(x_t)\,\partial t\,,$$

where $S \in \Gamma(\mathrm{End}\, TM)$ is the Ricci tensor. So (2.1) may be written

$$Dz_t = -\tfrac{1}{2} S(x_t)\, z_t\, \partial t\,. \tag{2.5}$$

We shall consider below the case where x_t is a Brownian motion of speed ε; then (2.1) still provides the process we want whereas (2.5) does not.

For bundles F and F' we write $C^\infty(F, F')$ for the set of smooth maps $g: F \to F'$ with $\pi \circ g = \pi$ where π denotes projection on M. For $V_0 \in C^\infty(F, F)$ and $V_1 \in C^\infty(F, F \otimes T^*M)$ consider the covariant equation in F over x_t

$$Dy_t = V_0(y_t)\,\partial t + V_1(y_t)\,\partial x_t,\; y_0 \in F_x\,. \tag{2.6}$$

Introduce a *derived equation* in $F \otimes T^*M$ over x_t

$$\begin{aligned}(DY_t)_i &= \nabla V_0(y_t)\,(X_t)_i\,\partial t + \nabla V_1(y_t)\,(X_t)_i\,\partial x_t\\ &\quad + DV_0(y_t)\,(Y_t)_i\,\partial t + DV_1(y_t)\,(Y_t)_i\,\partial x_t\\ &\quad + V_1(y_t)\,(DX_t)_i + (DW_t)_i\, y_t\\ Y_0 &= 0 \in F \otimes T^*M\end{aligned} \tag{2.7}$$

where DV denotes the derivative in the fibre and ∇V the covariant derivative with respect to the base point. We call Y_t the *derived process* of y_t. The solution of (2.7) and indeed of (2.6) may explode in a finite time. We introduce in the Appendix a class of *graded linear* equations (2.6) for which we show this cannot happen. The solution y_t of a graded linear equation will be called a *graded exponential.*

We say that y_0 is *independent of* x if, for all frames u_0 over x, $u_0^{-1} y_0$ does not depend on x. The examples we shall consider will involve either $y_0 = 0$ or $y_0 = \mathrm{id}$. (F being some space of endomorphisms), or some combination $y_0 = 0 \oplus \mathrm{id}$. We can now state our fundamental result.

Theorem 2.1 *Let y_t be a graded exponential in F over (t, x_t) with y_0 independent of x, and with derived process Y_t. Then for all functions $g \in C^\infty(F, F)$ of polynomial growth*

$$T\nabla_{x^i}\mathbb{E}^x[\tau_{0T}\, g(y_T)] = \mathbb{E}^x[\tau_{0T}\{(Q_T)_i\, g(y_T) - (W_T)_i\, g(y_T) + Dg(y_T)\,(Y_T)_i\}]\,.$$

The proof is given in Sect. 3.

If we take $F = M \times \mathbb{R}$ with trivial connection and $g = f \circ \pi$ with $f \in C^\infty(M, \mathbb{R})$, then the action of W_T on F is as 0 and $Dg = 0$ so we obtain

$$T\nabla_x \mathbb{E}^x[f(x_T)] = \mathbb{E}^x[\tau_{0T} Q_T f(x_T)] . \tag{2.8}$$

This implies for almost all $y \in M$ that

$$T\nabla_x p_T(x, y) = p_T(x, y)\,\mathbb{E}^{x,y,T}[\tau_{0T} Q_T] . \tag{2.9}$$

This is Bismut's formula [3, Theorem 2.14]. Lemma A.4 from the Appendix shows in particular that the expectation on the right of (2.9) depends continuously on y, so (2.9) in fact holds for *all* $y \in M$.

Our next application of Theorem 2.1 generalizes Bismut's formula to the vector case. Recall that we define e_t in $\operatorname{End} F$ over x_t by

$$De_t = e_t V(x_t)\,\partial t, \qquad e_0 = \mathrm{id.} \in \operatorname{End} F . \tag{2.10}$$

The derived process E_t in $\operatorname{End} F \otimes T^*M$ is then given by

$$(DE_t)_i = (E_t)_i V(x_t)\,\partial t + e_t \nabla V(x_t)\,(X_t)_i\,\partial t + [(DW_t)_i, e_t]$$
$$E_0 = 0 \in \operatorname{End} F \otimes T^*M$$

where $[\,,]$ denotes the commutator of endomorphisms.

Theorem 2.2 *For all $x, y \in M$ and $T \in (0, \infty)$ we have*

$$T\nabla_{x^i} P_T(x, y) = p_T(x, y)\;\mathbb{E}^{x,y,T}[\tau_{0T}\{(Q_T)_i e_T - (W_T)_i e_T + (E_T)_i\}] .$$

Proof. Replace F by $\operatorname{End} F$, y_t by e_t and $g(y)$ by $ef(x)$, with f a section of F, to obtain

$$T\nabla_{x^i} \mathbb{E}^x[\tau_{0T} e_T f(x_T)] = \mathbb{E}^x[\tau_{0T}\{(Q_T)_i\, e_T - (W_T)_i\, e_T + (E_T)_i\} f(x_T)] .$$

Now condition on $x_T = y$ and apply Lemma A.4. □

We can also obtain formulae for higher derivatives. The simplest of these is the second order formula for the scalar case. The process z_t^* satisfies the equation

$$Dz_t^* = -\tfrac{1}{2} z_t^* S(x_t)\,\partial t\,, \qquad z_0^* = \mathrm{id.} \in \operatorname{End} TM \tag{2.11}$$

which is a special case of (2.10). Hence its derived process, which we denote Z_t^*, satisfies

$$(DZ_t^*)_i = -\tfrac{1}{2}(Z_t^*)_i\, S(x_t)\,\partial t - \tfrac{1}{2} z_t^*\;\nabla S(x_t)\,(X_t)_i\,\partial t + [(DW_t)_i, z_t^*]$$
$$Z_0^* = 0 \in \operatorname{End} TM \otimes T^*M .$$

Now Q_t does not satisfy an autonomous covariant stochastic differential equation, so to obtain the derived process $\tilde{Q}_t$ of Q_t we consider the autonomous system (2.11), (2.4) for (z_t^*, Q_t). We find that

$$(D\tilde{Q}_t)_{ij} = \{(Z_t^*)_j \partial x_t + z_t^*(DX_t)_j\}^i - \{Q_t(DW_t)_j\}_i$$
$$\tilde{Q}_0 = 0 \in T^*M \otimes T^*M .$$

Theorem 2.3 *For all $x, y \in M$ and $T \in (0, \infty)$ we have*

$$T^2\,\nabla_x^2 p_T(x, y) = p_T(x, y)\,\mathbb{E}^{x,y,T}[\tau_{0T}\{Q_T \otimes Q_T + Q_T W_T + \tilde{Q}_T\}] .$$

Proof. In Theorem 2.1, take $F = T^*M$, $y_t = (z_t^*, Q_t)$ and $g(y) = Qf(x)$, with $f \in C^\infty(M, \mathbb{R})$ to obtain

$$T\nabla_x \mathbb{E}^x[\tau_{0T} Q_T f(x_T)] = \mathbb{E}^x[\tau_{0T}\{Q_T \otimes Q_T + Q_T W_T + \tilde{Q}_T\} f(x_T)] .$$

By (2.8) the left hand side equals

$$T^2\, \nabla_x^2\, \mathbb{E}^x[f(x_T)] .$$

Now condition on $x_T = y$ and apply Lemma A.4. □

We may rescale the time parameter in our formulae so that all integrals are taken over paths parametrized by the interval [0, 1]. This is appropriate when considering small-time asymptotics. We denote by $\mathbb{P}_\varepsilon^x$ the law of Brownian motion of speed $\varepsilon \in (0, \infty)$ starting from x and by $\mathbb{P}_\varepsilon^{x,y,1}$ the law of the Brownian bridge of speed ε from x to y in time 1. Then we have a rescaled version of Theorem 2.1.

Theorem 2.4 *Let y_t be a graded exponential in F over (t, x_t) with y_0 independent of x, and with derived process Y_t. Then for all functions $g \in C^\infty(F, F)$ of polynomial growth*

$$\nabla_{x^i} \mathbb{E}_\varepsilon^x[\tau_{01} g(y_1)] = \mathbb{E}_\varepsilon^x[\tau_{01}\{(Q_1/\varepsilon)_i\, g(y_1) - (W_1)_i g(y_1) + Dg(y_1)(Y_1)_i\}]. \quad (2.12)$$

This formula is a straightforward deduction from Theorem 2.1. Alternatively, to obtain the formula directly, one can repeat the proof of Theorem 2.1. The only modification required is an extra factor of $\dfrac{1}{\varepsilon}$ in the Girsanov exponential, which leads to the term Q_1/ε in the formula.

Theorem 2.5 *Suppose there is a unique minimal geodesic γ from x to y in time* 1, *and write $\tau_{xy}: F_y \to F_x$ for the parallel translation along γ. Then for all $N = 0, 1, 2, \ldots$ we have*

$$\lim_{t\downarrow 0} \frac{t^N\, \nabla_x^N P_t(x, y)}{p_t(x, y)} = \tau_{xy} \otimes \dot{\gamma}(0)^{\otimes N} .$$

Proof. After rescaling time by a factor of ε, the appropriate equations for e_t and z_t are

$$De_t = \varepsilon e_t\, V(x_t)\partial t, \qquad Dz_t = -\frac{\varepsilon}{2} S(x_t)\, z_t \partial t .$$

So e_t, z_t and all the processes X_t, W_t, Q_t and Y_t defined using z_t, have a dependence on ε, which we now make explicit.

On iterating N times the rescaled integration by parts formula (2.12), we obtain a graded exponential y_t^ε over (t, x_t) with coefficients depending smoothly on ε and a polynomial g (with values in $\operatorname{End} F \otimes (T^*M)^{\otimes N}$) such that, for all $f \in \Gamma(F)$ and all $\varepsilon \in (0, \infty)$

$$\varepsilon^N\, \nabla_x^N \mathbb{E}_\varepsilon^x[\tau_{01} e_t^\varepsilon f(x_1)] = \mathbb{E}_\varepsilon^x[\tau_{01} e_t^\varepsilon f(x_1) \otimes (Q_1^\varepsilon)^{\otimes N}] + \varepsilon \mathbb{E}_\varepsilon^x[\tau_{01} g(\varepsilon, y_1^\varepsilon) f(x_1)] .$$

We condition on $x_1 = y$ and appeal to Lemma A.4 to deduce

$$\frac{\varepsilon^N\, \nabla_x^N P_\varepsilon(x, y)}{p_\varepsilon(x, y)} = \mathbb{E}_\varepsilon^{x,y,1}\,[\tau_{01} e_t^\varepsilon \otimes (\tau_{01} Q_1^\varepsilon)^{\otimes N}] + \varepsilon \mathbb{E}_\varepsilon^{x,y,1}\,[\tau_{01} g(\varepsilon, y_1^\varepsilon)] . \quad (2.13)$$

By Remark A.2, Lemma B.2 and Lemma B.4, we know that for all $\delta > 0$, as $\varepsilon \downarrow 0$,

$$\mathbb{P}_\varepsilon^{x,y,1}[|\tau_{01} e_t^\varepsilon - \tau_{xy}| > \delta] \to 0$$

$$\mathbb{P}_\varepsilon^{x,y,1}[|\tau_{01} Q_1^\varepsilon - \dot{\gamma}(0)| > \delta] \to 0$$

also

$$\sup_{0<\varepsilon\leqq 1} \mathbb{E}_\varepsilon^{x,y,1}[|\tau_{01} \otimes (\tau_{01} Q_1^\varepsilon)^{\otimes N}|^2] < \infty$$

$$\sup_{0<\varepsilon\leqq 1} \mathbb{E}_\varepsilon^{x,y,1}[|\tau_{01} g(\varepsilon, y_1^\varepsilon)|] < \infty .$$

So we can take the limit $\varepsilon \downarrow 0$ in (2.13) to prove the theorem. □

A different probabilistic approach to the behaviour of heat kernels and their derivatives in small time was given by Ben–Arous [1]. He considers only the scalar heat kernel, but investigates the more complex hypoelliptic case. Whereas we compare the derivatives of $P_t(x, y)$ to $p_t(x, y)$ in small time, Ben–Arous' results give asymptotics for the derivatives of a hypoelliptic heat kernel in terms of the solution of a system of auxilliary transport equations. The relative asymptotics we obtain do not require x and y to be non-conjugate along their minimal geodesic and so apply to generic points of the cut-locus, even though the asymptotics of $p_t(x, y)$ itself there are not well understood. The author is grateful to a referee for this comment and for the above reference.

Example 2.6. We consider the special case $F = M \times \mathbb{R}$ and $\mathscr{L} = \frac{1}{2}\Delta + X$ where Δ is the Laplace–Beltrami operator and X is a vector field on M. Then in the decomposition (1.1) we have $\nabla = d + \langle X, \cdot \rangle$, so parallel translation along x_t is given by

$$\tau_{0t} = \exp \int_0^t \langle X(x_t), \partial x_t \rangle .$$

Write $p_t^X(x, y)$ for the heat kernel of $\mathscr{L}$. Theorem 2.5 tells us that if there is a unique geodesic γ from x to y in time 1, then for all $N = 0, 1, 2, \ldots$

$$\lim_{t\downarrow 0} \frac{t^N \nabla_x^N p_t^X(x, y)}{p_t(x, y)} = \dot{\gamma}(0)^{\otimes N} \exp \int_0^t \langle X(\gamma_t), \partial\gamma_t \rangle .$$

The reader may like to check the case where M is $\mathbb{R}^n$ and X is a constant!

3 Proof of Theorem 2.1

We begin by computing formally the derivatives of certain processes defined by stochastic differential equations depending on a parameter. Then we show how Theorem 2.1 follows from the quasi-invariance of Wiener measure in $\mathbb{R}^n$ under certain transformations, where n is the dimension of the manifold M. The basic method goes back to Bismut [2]. The approach followed here is close to Bismut [3].

Let $\tilde{x}_t$ be a Brownian motion in $\mathbb{R}^n$. Replace, if necessary, the bundle F by the tensor product $F \otimes TM$, and ∇ by its product with the Levi–Civita connection. Then we can assume F has the form $F' \otimes TM$ and ∇ respects this product. Fix an initial frame $u_0 \in \mathrm{GL}(E, F)$ also respecting this product structure. The stochastic development x_t of $\bar{x}_t$ in M through u_0 and its horizontal lift u_t starting from u_0 have a global description in terms of the canonical 1-form

$$\theta(V) = u^{-1}\pi^*(V)\,, \quad V \in T_u\,\mathrm{GL}(E, F)$$

and the connection form ω of ∇. In fact u_t is the unique solution of the equations

$$\theta(\partial u_t) = \partial \bar{x}_t, \quad \omega(\partial u_t) = 0$$

starting from u_0.

If y_t is a process in F over x_t, we write $\bar{y}_t$ for the process $u_t^{-1} y_t$ in E: thus, for z_t and W_t defined by (2.1) and (2.3), $\bar{z}_t = u_t^{-1} z_t u_t$ takes values in $\mathrm{End}\,\mathbb{R}^n$ and $\bar{W}_t$ takes values in $\mathrm{End}\,E \otimes (\mathbb{R}^n)^*$. For simplicity we consider first the case where $F = TM$ and ∇ is the Levi–Civita connection. Let $\varepsilon \in \mathbb{R}^n$ be given and recall that $T \in (0, \infty)$ is fixed throughout. Define processes a_t in $\mathbb{R}^n$ and b_t in $\mathrm{End}\,\mathbb{R}^n$ by

$$a_t = -\,\bar{z}_t \varepsilon / T\,, \quad b_t = -\,\bar{W}_t \varepsilon / T \tag{3.1}$$

and let $\gamma(\ell)$ be the horizontal curve in $\mathrm{GL}(M)$ defined by

$$\theta\left(\frac{\partial \gamma}{\partial \ell}\right) = \varepsilon, \quad \omega\left(\frac{\partial \gamma}{\partial \ell}\right) = 0 \tag{3.2}$$

with $\gamma(0) = u_0$. Introduce a perturbed process $\bar{x}_t^\ell$ in $\mathbb{R}^n$, starting from 0, by

$$d\bar{x}_t^\ell = e^{\ell b_t}\, d\bar{x}_t + \ell a_t\, dt\,. \tag{3.3}$$

Here $d\bar{x}_t$ is the Itô differential.

Let u_t^ℓ be the unique solution of the equations

$$\theta(\partial u_t^\ell) = \partial \bar{x}_t^\ell\,, \quad \omega(\partial u_t^\ell) = 0$$

starting from $\gamma(\ell)$. We shall show that

$$T\theta\left(\left.\frac{\partial}{\partial \ell}\right|_{\ell=0} u_t^\ell\right) = \bar{X}_t \varepsilon, \quad T\omega\left(\left.\frac{\partial}{\partial \ell}\right|_{\ell=0} u_t^\ell\right) = \bar{W}_t \varepsilon\,. \tag{3.4}$$

Proof of (3.4) Set $U_t = \left.\dfrac{\partial}{\partial \ell}\right|_{\ell=0} u_t^\ell$, $\theta_t = \theta(U_t)$ and $\omega_t = \omega(U_t)$. We have

$$\left.\frac{\partial}{\partial \ell}\right|_{\ell=0} \theta(\partial u_t^\ell) = b_t\, d\bar{x}_t + a_t\, \partial t$$

$$\left.\frac{\partial}{\partial \ell}\right|_{\ell=0} \omega(\partial u_t^\ell) = 0$$

so

$$d\theta(\partial u_t, U_t) = \partial \theta_t - (b_t\, d\bar{x}_t + a_t\, \partial t)$$

$$d\omega(\partial u_t, U_t) = \partial \omega_t\,.$$

The connection has structure equations

$$d\theta = -\,\omega \wedge \theta$$

$$d\omega = -\,\omega \wedge \omega + \Omega$$

where Ω is the curvature form, so we obtain

$$\partial\theta_t = b_t d\bar{x}_t + a_t \partial t + \omega_t \partial\bar{x}_t$$
$$\partial\omega_t = \Omega(\partial u_t, U_t) = \bar{R}_{u_t}(\partial\bar{x}_t, \theta_t)$$

where $\bar{R}$ is the equivariant representation of the curvature tensor. We know that $\theta_0 = \varepsilon$ and $\omega_0 = 0$, so these equations determine θ_t and ω_t uniquely. Substitution for a_t and b_t by (3.1) and for θ_t and ω_t by the claimed solutions (3.4) yields

$$\partial\bar{X}_t\varepsilon = -\bar{W}_t\varepsilon d\bar{x}_t - \bar{z}_t\varepsilon\partial t + \bar{W}_t\varepsilon\partial\bar{x}_t$$
$$\partial\bar{W}_t\varepsilon = \bar{R}_{u_t}(\partial\bar{x}_t, \bar{X}_t\varepsilon) .$$

The Stratonovich to Itô conversion rule shows

$$\bar{W}_t\varepsilon\partial\bar{x}_t - \bar{W}_t\varepsilon dx_t = \tfrac{1}{2}\partial\bar{W}_t\varepsilon\partial\bar{x}_t = \tfrac{1}{2}\bar{R}_{u_t}(\partial\bar{x}_t, \bar{X}_t\varepsilon)\partial\bar{x}_t .$$

So, in covariant notation, we have to check

$$(DX_t)_i = \tfrac{1}{2}R(\partial x_t, (X_t)_i)\partial x_t - (z_t)_i\partial t$$
$$(DW_t)_i = R(\partial x_t, (X_t)_i)$$

and this follows easily from (2.1), (2.2) and (2.3) . □

If we start with a general connection ∇ on a general bundle F, extended to TM by Levi–Civita, then the perturbation of the Brownian motion x_t remains the same and (3.4) remains true.

Now suppose we define a process y_t in F over x_t by the covariant Eq. (2.6) with y_0 independent of x. Define a perturbed process y_t^ℓ over x_t^ℓ by

$$Dy_t^\ell = V_0(y_t^\ell)\partial t + V_1(y_t^\ell)\partial x_t^\ell, \quad y_0^\ell = y_0(x_0^\ell) .$$

Then $\bar{y}_t^\ell := (u_t^\ell)^{-1} y_t^\ell$ satisfies

$$\partial\bar{y}_t^\ell = \bar{V}_0(u_t^\ell, \bar{y}_t^\ell)\partial t + \bar{V}_1(u_t^\ell, \bar{y}_t^\ell)\partial\bar{x}_t^\ell, \quad \bar{y}_0^\ell = \bar{y}_0$$

where $\bar{V}_0(u, \bar{y}) = u^{-1}V_0(u\bar{y})$ and $\bar{V}_1(u, \bar{y}) = u^{-1}V_1(u\bar{y})u$. We differentiate formally to see that $\bar{Y}_t^0\varepsilon := T\left.\frac{\partial}{\partial\ell}\right|_{\ell=0}\bar{y}_t^\ell$ satisfies the equation

$$\partial\bar{Y}_t^0\varepsilon = Td\bar{V}_0(U_t, \bar{y}_t)\partial t + Td\bar{V}_1(U_t, \bar{y}_t)\partial\bar{x}_t + D\bar{V}_0(u_t, \bar{y}_t)\bar{Y}_t^0\varepsilon\partial t$$
$$+ D\bar{V}_1(u_t, \bar{y}_t)\bar{Y}_t^0\varepsilon\partial\bar{x}_t - \bar{V}_1(u_t, \bar{y}_t)(\bar{z}_t\varepsilon\partial t + \bar{W}_t\varepsilon d\bar{x}_t) .$$

Hence $Y_t^0 := u_t.\bar{Y}_t^0$ satisfies the covariant equation

$$(DY_t^0)_i = \nabla V_0(y_t)(X_t)_i\partial t + \nabla V_1(y_t)(X_t)_i\partial x_t$$
$$+ DV_0(y_t)((Y_t^0)_i + (W_t)_i y_t)\partial t + DV_1(y_t)((Y_t^0)_i + (W_t)_i y_t)\partial x_t$$
$$- V_1(y_t)((z_t)_i\partial t + (W_t)_i dx_t) - (W_t)_i Dy_t .$$

Define Y_t' by $(Y_t')_i = (Y_t^0)_i + (W_t)_i y_t$. Then

$$(DY_t')_i = (DY_t^0)_i + (DW_t)_i y_t + (W_t)_i Dy_t$$

so Y'_t satisfies the Eq. (2.7). We have shown that

$$T\frac{\partial}{\partial\ell}\bigg|_{\ell=0}((u_t^\ell)^{-1}y_t^\ell) + T\omega\left(\frac{\partial}{\partial\ell}\bigg|_{\ell=0}u_t^\ell\right)u_t^{-1}y_t = \bar{Y}_t\varepsilon \tag{3.5}$$

where $\bar{Y}_t = u_t^{-1}Y_t$ and Y_t is given by (2.7).

The process $a_t = -\bar{z}_t\varepsilon/T$ is uniformly bounded on compact time intervals so

$$q_t^\ell := \exp\left\{-\ell\int_0^t\langle a_s, d\bar{x}_s\rangle - \frac{\ell^2}{2}\int_0^t|a_s|^2\,ds\right\}$$

defines a martingale. The process $b_t = -\bar{W}_t\varepsilon/T$ takes values in skew-symmetric matrices, so $e^{\ell b_t}$ is a rotation. Hence under the new probability measure $\mathbb{Q}^\ell$ defined by

$$\frac{d\mathbb{Q}^\ell}{d\mathbb{P}}\bigg|_{\mathscr{F}_t} = q_t^\ell$$

$\bar{x}_t^\ell$ is a Brownian motion. The law of $\bar{x}_t^\ell$ determines the laws of u_t^ℓ and y_t^ℓ but remember that u_t^ℓ starts from $\gamma(\ell)$, so for $g\in C^\infty(F,F)$ of polynomial growth

$$\mathbb{E}[(u_t^\ell)^{-1}g(y_t^\ell)q_t^\ell] = \gamma(\ell)^{-1}\mathbb{E}^{\pi\gamma(\ell)}[\tau_{0t}g(y_t)]$$

and so

$$\frac{\partial}{\partial\ell}\bigg|_{\ell=0}\mathbb{E}[(u_t^\ell)^{-1}g(y_t^\ell)q_t^\ell] = u_0^{-1}\nabla_x\mathbb{E}^x[\tau_{0t}g(y_t)](\varepsilon)\,.$$

Now Lemma A.3 allows us to take the derivative in ℓ inside the expectation and (3.4) and (3.5) enables us to express the derivative in terms of X_t, W_t and Y_t.

We have

$$\begin{aligned}
&\frac{\partial}{\partial\ell}\bigg|_{\ell=0}(u_t^\ell)^{-1}g(y_t^\ell)\\
&\quad = (u_t^{-1}\nabla g(y_t))\,\theta\left(\frac{\partial}{\partial\ell}\bigg|_{\ell=0}u_t^\ell\right) - \omega\left(\frac{\partial}{\partial\ell}\bigg|_{\ell=0}u_t^\ell\right)u_t^{-1}g(y_t)\\
&\qquad + (u_t^{-1}Dg(y_t))\left(\frac{\partial}{\partial\ell}\bigg|_{\ell=0}((u_t^\ell)^{-1}y_t^\ell) + \omega\left(\frac{\partial}{\partial\ell}\bigg|_{\ell=0}u_t^\ell\right)u_t^{-1}y_t\right)
\end{aligned}$$

so

$$\begin{aligned}
&T\frac{\partial}{\partial\ell}\bigg|_{\ell=0}(u_t^\ell)^{-1}g(y_t^\ell)q_t^\ell\\
&\quad = (u_t^{-1}\nabla g(y_t))\bar{X}_t\varepsilon + (u_t^{-1}Dg(y_t))\bar{Y}_t\varepsilon - \bar{W}_t\varepsilon(u_t^{-1}g(y_t)) + (u_t^{-1}g(y_t))\bar{Q}_t\varepsilon\,.
\end{aligned}$$

Let ε run over a basis of $\mathbb{R}^n$ to deduce

$$\nabla_{x^i}\mathbb{E}^x[\tau_{0t}g(y_t)] = \mathbb{E}[\tau_{0t}\{\nabla g(y_t)(X_t)_i + Dg(y_t)(Y_t)_i - (W_t)_i g(y_t) + g(y_t)(Q_t)_i\}]\,.$$

Finally, take $t = T$ and note $X_T = 0$ to prove the theorem. □

Appendix A: Regularity results for graded exponentials

We will obtain some regularity results for a nice class of covariant stochastic differential equations

$$Dy_t = V(y_t)\,\partial x_t\,.$$

Here we are given a semimartingale x_t in M and $V \in C^\infty(F, F \otimes T^*M)$ and the solution y_t is to take values over x_t in the bundle F. We can rewrite this equation using a horizontal lift u_t of x_t in $\mathrm{GL}(E, F)$: set $\bar{y}_t = u_t^{-1} y_t$ and $\bar{V}(u, \bar{y}) = u^{-1} V(u\bar{y}) u$, then

$$\partial \bar{y}_t = \bar{V}(u_t, \bar{y}_t) \partial \bar{x}_t \,. \tag{A.1}$$

These equations possess, for each initial value y_0, a unique solution up to explosion.

Suppose now there is a direct sum decomposition $F = F^1 \oplus \cdots \oplus F^k$, respected by the connection, and that the equation

$$D y_t = V(y_t) \partial x_t \tag{A.2}$$

decomposes as

$$D y_t^j = A^j(\partial x_t) y_t^j + U^j(\partial x_t)(y_t^1, \ldots y_t^{j-1}) \tag{A.3}$$

where, for $j = 1, \ldots, k$, $A^j \in \Gamma(\mathrm{End}\, F^j \otimes T^*M)$ and $U^j(x)(y^1, \ldots, y^{j-1})$ is a polynomial in $y^1, \ldots, y^{j-1}$, depending smoothly on x, with values in $F^j \otimes T^*M$. Then we call (A.2) a *graded linear* equation and we call any solution a *graded exponential* over x_t. The maximum degree d of the polynomials $U^j(x)$, $j = 1, \ldots, k$ is called the *degree* of y_t.

We shall be mainly interested in the case where x_t is a Brownian motion, or a Brownian bridge, and where y_t is a graded exponential over the semimartingale (t, x_t) in $\mathbb{R} \times M$. More specifically we shall be interested in the case where y_t in fact satisfies an equation of the form

$$D y_t = V_0(y_t) \partial t + V_1(y_t) \partial x_t \,.$$

The class of such graded exponentials includes constants, exponentials such as e_t and z_t in Sect. 2, and other processes such as (z_t^*, Q_t) as in Theorem 2.3. Moreover, if y_t is a graded exponential with derived process Y_t, then $(z_t, X_t, W_t, Q_t, y_t, Y_t)$ is also a graded exponential (see Eqs. (2.1), (2.2), (2.3), (2.4) and (2.7)). So Theorem 2.1 and its variant Theorem 2.4 are suitable for iteration.

We return to the case of a general semimartingale x_t in M. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_k)$ we write

$$y_t^{\otimes \alpha} = (y_t^1)^{\otimes \alpha_1} \otimes \cdots \otimes (y_t^k)^{\otimes \alpha_k}$$

and

$$\|\alpha\| = \|\alpha\|_d = \sum_{j=1}^{k} \alpha_j d^j .$$

The following result makes analysis of graded exponentials rather easy.

Lemma A.1 *Let y_t be a graded exponential over x_t of degree d. Then $(y_t^{\otimes \alpha} : \|\alpha\|_d \leqq d^k)$ is a horizontal lift of x_t for some connection.*

Proof. We extend the connection ∇ and the covariant Stratonovich differential D to tensor products in the usual way. Recall (A.3)

$$D y_t^j = A^j(\partial x_t) y_t^j + U^j(\partial x_t)(y_t^1, \ldots, y_t^{j-1})$$

where A^j is linear and U^j is polynomial, of degree at most d. So we have

$$D y_t^j = \tilde{A}^j(\partial x_t)(y_t^{\otimes \alpha} : \|\alpha\| \leqq d^j)$$

where $\tilde{A}^j$ is linear. The product rule shows that

$$D((y_t^j)^{\otimes m}) = A^{j,m}(\partial x_t)\,(y_t^{\otimes \alpha} : \|\alpha\| \leqq md^j)$$

with $A^{j,m}$ linear, and then

$$D(y_t^{\otimes \alpha}) = A^\alpha(\partial x_t)\,(y_t^{\otimes \beta} : \|\beta\| \leqq \|\alpha\|)$$

with A^α linear. Define a new connection on $\bigoplus_{\|\alpha\| \leqq d^k} (F^1)^{\otimes \alpha_1} \otimes \cdots \otimes (F^k)^{\otimes \alpha_k}$ by $\tilde{\nabla}^\alpha = \nabla^\alpha - A^\alpha$, then $(y_t^{\otimes \alpha} : \|\alpha\| \leqq d^k)$ is the corresponding horizontal lift of x_t starting from $y_0^{\otimes \alpha}$. □

Remark A.2 Lemma A.1, in conjunction with Lemma B.2 and Lemma B.4 below, provides L^p-estimates and weak convergence results for graded exponentials over a Brownian bridge x_t of small speed $\varepsilon > 0$. More generally, as discussed above, we are interested in solutions of equations

$$Dy_t = V_0(y_t)\,\partial t + V_1(y_t)\,\partial x_t$$

which are graded exponentials over (t, x_t). Lemma A.1 now provides a reduction to the case of time dependent horizontal lift with drift. As discussed in Remark B.3(i), the L^p-estimates and weak convergence results still apply.

The following result provides justification for the formal differentiation of graded linear equations carried out in Sect. 3. Recall that $\bar{x}_t^\ell$ is defined by (3.3). Let $\gamma : [-1, 1] \to \mathrm{GL}(E, F)$ and $\eta : [-1, 1] \to F$ be given smooth curves with $\pi\gamma = \pi\eta$, and let V_0 and V_1 be coefficients of a graded linear equation.

Lemma A.3 *There exist maps*

$$\varphi : \Omega \times [0, \infty) \times [-1, 1] \to \mathrm{GL}(E, F)$$

$$\psi : \Omega \times [0, \infty) \times [-1, 1] \to F$$

with the following properties:

(i) *For each* $\ell \in [-1, 1]$, $u_t^\ell = \varphi(t, \ell)$ *satisfies the equations*

$$\theta(\partial u_t^\ell) = \partial \bar{x}_t^\ell, \quad \omega(\partial u_t^\ell) = 0, \quad u_0^\ell = \gamma(\ell) \tag{A.4}$$

and $y_t^\ell = \psi(t, \ell)$ *satisfies the graded linear equation over* $x_t^\ell = \pi u_t^\ell$

$$Dy_t^\ell = V_0(y_t^\ell)\,\partial t + V_1(y_t^\ell)\,\partial x_t^\ell, \quad y_0^\ell = \eta(\ell)\,. \tag{A.5}$$

(ii) *Almost surely and for all* $t \in [0, \infty)$, φ_t *and* ψ_t *are smooth in* $\ell \in [-1, 1]$ *and their derivatives satisfy the stochastic differential equations obtained by formal differentiation of* (A.4) *and* (A.5).

(iii) *For any metric on* F, *for all* $t \in [0, \infty)$ *and* $p \in [1, \infty)$ *we have*

$$\sup_{\ell \in [-1,1]} \mathbb{E}\left[|\varphi_t(\ell)|^p + |\psi_t(\ell)|^p + \left|\frac{\partial \varphi_t}{\partial \ell}(\ell)\right|^p + \left|\frac{\partial \psi_t}{\partial \ell}(\ell)\right|^p \right] < \infty\,.$$

Proof. Fix a metric on F and assume that ∇ is compatible with the metric. (The horizontal lift of any other connection $\nabla - A$ satisfies a linear equation $Du_t = A(\partial x_t)\,u_t$ which can be subsumed in (A.5), so we assume this without loss.) By Lemma A.1 we are reduced to the case where (A.5) is replaced by a linear equation

$$Dy_t^\ell = \{A_0(\partial t) + A_1(\partial x_t^\ell)\}\,y_t^\ell\,.$$

Now embed $O(E, F)$ as a (compact) submanifold in some $\mathbb{R}^N$. We can rewrite (A.4) in the form

$$\partial u_t^\ell = U(u_t^\ell)\,\partial \bar{x}_t^\ell \quad U_0^\ell = \gamma(\ell)$$

with $U \in C^\infty(\mathbb{R}^N, \mathbb{R}^N \otimes (\mathbb{R}^n)^*)$ of compact support. Moreover $\bar{y}_t^\ell := (u_t^\ell)^{-1} y_t^\ell$ satisfies

$$\partial \bar{y}_t^\ell = \{\bar{A}_0(u_t^\ell)\,\partial t + \bar{A}_1(u_t^\ell)\,\partial \bar{x}_t^\ell\}\,\bar{y}_t^\ell, \quad \bar{y}_0^\ell = \gamma(\ell)^{-1}\,\eta(\ell)$$

where $\bar{A}(u) = u^{-1} A(x)$. The conclusions we seek now follow from Carverhill and Elworthy [4], for example. □

Lemma A.4 *Let y_t be a graded exponential in F over (t, x_t) and suppose $g \in C^\infty(F, \mathbb{R})$ is of polynomial growth. Then for all $x \in M$ and $T \in (0, \infty)$ the map $y \mapsto \mathbb{E}^{x,y,T}[g(y_T)]$ is continuous.*

Proof. Let U be an open set in M whose closure lies within the domain D of a coordinate chart. Set

$$h(x, y) = \begin{cases} 1 & \text{if } x, y \in D \text{ and } x^i \leqq y^i \text{ for } i = 1, \dots, n \\ 0 & \text{otherwise.} \end{cases}$$

Note that, as $y_n \to y$, for almost all $x \in M$

$$h(x, y_n) \to h(x, y)\,. \tag{A.6}$$

Write the Riemannian volume element as $\rho(x)\,dx^1 \dots dx^n$ within D. Construct a cut-off function $\psi \in C^\infty(M)$ with $1_U \leqq \psi \leqq 1_D$ and define

$$\mathscr{L} f(x) = \frac{\psi(x)}{\rho(x)} \frac{\partial}{\partial x^1} \cdots \frac{\partial}{\partial x^n} f(x), \quad f \in C^\infty(M)\,.$$

Fixing $y \in U$, take a sequence of functions $h_k \in C^\infty(M)$ converging to $h(\cdot, y)$ almost everywhere. Then $f_k := \mathscr{L} h_k \to \delta_y$ in the sense that, for all $\varphi \in C^\infty(M)$,

$$\int_M \varphi(x) f_k(x)\,dx \to \varphi(y)\,.$$

By an integration by parts procedure, similar to that used in the proof of Theorem 2.1, used also in [8] in the case $M = \mathbb{R}^n$, we can find a graded exponential y_t', with values in F' say, and a function $g' \in C^\infty(F' \oplus F', \mathbb{R})$ of polynomial growth, such that, for all $f \in C^\infty(M)$ and all $0 < t \leqq T$

$$T^n \mathbb{E}^x[\mathscr{L} f(x_T) g(y_t)] = \mathbb{E}^x[f(x_T) g'(y_t', y_T')]\,.$$

Hence for the sequence f_k above,

$$\lim_{k \to \infty} T^n\, \mathbb{E}^x[f_k(x_T) g(y_t)] = \mathbb{E}^x[h(x_T) g'(y_t', y_T')]\,.$$

For $0 \leqq t < T$ we have

$$\begin{aligned} \lim_{k \to \infty} \mathbb{E}^x[f_k(x_T) g(y_t)] &= \lim_{k \to \infty} \mathbb{E}^x[(P_{T-t} f_k)(x_t) g(y_t)] \\ &= \mathbb{E}^x[p_{T-t}(x_t, y) g(y_t)] \\ &= p_T(x, y)\, \mathbb{E}^{x,y,T}[g(y_t)]\,. \end{aligned}$$

Hence

$$T^n p_T(x, y)\,\mathbb{E}^{x,y,T}[g(y_t)] = \mathbb{E}^x[h(x_T, y)\,g'(y'_t, y'_T)]\,. \tag{A.7}$$

We know $y'_t \to y'_T$ in $L^p(\mathbb{P}^x)$ for all $p \in [1, \infty)$ by standard estimates, and $y_t \to y_T$ almost surely under $\mathbb{P}^{x,y,T}$. Since (A.7) holds for all g of polynomial growth, we see that $(y_t : 0 \leqq t < T)$ is bounded in $L^p(\mathbb{P}^{x,y,T})$ for all p, hence $y_t \to y_T$ in $L^p(\mathbb{P}^{x,y,T})$ and (A.7) extends to $t = T$. The result now follows from (A.6) and (A.7). □

Appendix B: Estimates on the Brownian bridge

We are interested in the asymptotics as $\varepsilon \downarrow 0$ of parallel translation along a Brownian bridge of speed ε. We shall use the following basic estimates. For all $\beta \in (0, \infty)$, there is a constant $C(M, \beta) \in (0, \infty)$ such that for all $x, y \in M$ and all $0 < t \leqq 1$

$$p_t(x, y) \leqq \frac{C(M, \beta)}{t^{n/2}} \exp\left\{-\frac{d^2(x, y)}{2(1+\beta)t}\right\} \tag{B.1}$$

where d is the Riemannian distance. We need the lower bound only in asymptotic form

$$\liminf_{t \downarrow 0} t \log p_t(x, y) \geqq -\frac{d^2(x, y)}{2}\,. \tag{B.2}$$

For (B.1) see Li and Yau [7, Corollary 3.1] and for (B.2) see Varadhan [10]. The following is a crude version of [7, Theorem 2.3]: there is a constant $C(M) \in (0, \infty)$ such that, for all $x, y, z \in M$ and $t \in (0, \infty)$

$$p_t(z, y) \leqq p_{2t}(x, y) \exp\left\{Ct + \frac{C}{t}\right\}. \tag{B.3}$$

We shall deduce from (B.3) another estimate which is implicit in [3].

Lemma B.1 *There is a constant $C(M) \in (0, \infty)$ such that for all $x, y \in M$ and $t \in (0, \infty)$* we have

$$t\,|\nabla_x \log p_t(x, y)| \leqq Ce^{Ct}.$$

Proof. We use Bismut's formula (2.9)

$$t\,\nabla_x p_t(x, z) = p_t(x, z)\,\mathbb{E}^{x,z,t}[\tau_{0t} Q_t]\,.$$

By integrating against $p_t(z, y)\,dz$ we obtain

$$t\,\nabla_x p_{2t}(x, y) = p_{2t}(x, y)\,\mathbb{E}^{x,y,2t}[\tau_{0t} Q_t]\,.$$

Let x_t be the stochastic development of $\bar{x}_t$ through u_0 and let u_t be the horizontal lift of x_t starting from u_0. Then

$$\tau_{0t} Q_t = \int_0^t \bar{z}_s^*\, d\bar{x}_s$$

where $\bar{z}_t^*$ satisfies

$$\partial \bar{z}_t^* = -\tfrac{1}{2}\bar{z}_t^* \bar{S}(u_t)\,\partial t. \quad \bar{z}_0^* = \mathrm{id}.$$

and $\bar{S}(u) = u.^{-1} S(x)$. Clearly we have

$$|\bar{z}_t^*| \leqq Ce^{Ct}.$$

But by an integration by parts

$$\int_0^t \bar{z}_s^*\, d\bar{x}_s = \bar{z}_t^* \bar{x}_t + \frac{1}{2}\int_0^t \bar{z}_s^* \bar{S}(u_s)\,\bar{x}_s\, ds$$

so

$$|\tau_{0t} Q_t| \leqq Ce^{Ct} \sup_{0 \leqq s \leqq t} |\bar{x}_s| \, .$$

Under $\mathbb{P}^x$, the process $\bar{x}_t$ is a Brownian motion in $\mathbb{R}^n$ so for $\lambda \in [1, \infty)$

$$\mathbb{P}^x\left(\sup_{0 \leqq s \leqq t} |\bar{x}_s| > \lambda\right) \leqq Ce^{-\frac{\lambda^2}{Ct}}.$$

We use the basic estimate (B.3) to deduce

$$\mathbb{P}^{x,y,2t}\left(\sup_{0 \leqq s \leqq t} |\bar{x}_s| > \lambda\right) \leqq e^{Ct + \frac{C}{t} - \frac{\lambda^2}{Ct}}$$

and hence

$$\mathbb{E}^{x,y,2t}\left(\sup_{0 \leqq s \leqq t} |\bar{x}_s|\right) \leqq Ce^{Ct}.$$

The lemma follows. □

Lemma B.2 *For all $x, y \in M$ and all $p \in [1, \infty)$*

$$\sup_{0<\varepsilon\leqq 1} \mathbb{E}_\varepsilon^{x,y,1}\left[\sup_{0 \leqq s \leqq t \leqq 1} \{|\tau_{st}|^p + |\tau_{ts}|^p\}\right] < \infty\, .$$

Proof. The time reversal of a Brownian bridge is a Brownian bridge and parallel translation is invariant under time reversal. For $s \leqq \frac{1}{2} \leqq t$ we have

$$|\tau_{ts}| \leqq |\tau_{t1}|\,|\tau_{1\frac{1}{2}}|\,|\tau_{\frac{1}{2}0}|\,|\tau_{0s}|$$

so it suffices to show

$$\sup_{0<\varepsilon\leqq 1} \mathbb{E}_\varepsilon^{x,y,1}\left[\sup_{0 \leqq t \leqq \frac{1}{2}} \{|\tau_{0t}|^p + |\tau_{t0}|^p\}\right] < \infty\, .$$

Fix an inner product on F and suppose for now that ∇ is compatible with this inner product. For $A \in \Gamma(\operatorname{End} F \otimes T^*M)$ consider the linear equation in $\operatorname{End} F$ over x_t

$$De_t = A(\partial x_t)\,e_t, \quad e_0 = \mathrm{id}. \tag{B.4}$$

We rewrite this equation using a horizontal lift u_t of x_t in $O(E, F)$ as at (A.1)

$$\partial \bar{e}_t = \bar{A}(u_t)\,\bar{e}_t\,\partial \bar{x}_t, \quad \bar{e}_0 = \mathrm{id}.$$

Note that $|e_t| = |\bar{e}_t|$ and by compactness $\bar{A}(u_t)$ is uniformly bounded. Under $\mathbb{P}_\varepsilon^{x,y,1}$ we have

$$\partial \bar{x}_t = \partial b_t + u_t^{-1} \varepsilon \nabla \log p_{\varepsilon(1-t)}(x_t, y)\, \partial t$$

where b_t is a Brownian motion in $\mathbb{R}^n$. The crucial estimate comes from Lemma B.1: for $0 \leqq t \leqq \frac{1}{2}$

$$\varepsilon |\nabla \log p_{\varepsilon(1-t)}(x_t, y)| \leqq C e^{2C}.$$

The usual combination of Gronwall's Lemma and Burkholder–Davis–Gundy inequalities now shows for $p \in [1, \infty)$ that

$$\sup_{0<\varepsilon\leqq 1} \mathbb{E}_\varepsilon^{x,y,1} \left[\sup_{0 \leqq t \leqq \frac{1}{2}} |e_t|^p \right] < \infty .$$

Since e_t^{-1} satisfies the equation

$$D(e_t^{-1}) = -\, e_t^{-1} A(\partial x_t), \quad e_0^{-1} = \mathrm{id}.$$

the same estimate holds for e_t^{-1}. Now $\zeta_{t0} = e_t \tau_{t0}$ is the parallel translation corresponding to the connection $\nabla - A$ and we obtain all connections in this way. □

Remark B.3 (i) The proof of Lemma B.2 applies also to the case of (time-dependent) parallel translation with drift. Replace M by $\mathbb{R} \times M$ and x_t by (t, x_t). Then we find at (B.4) an equation of the form

$$De_t = \{A_0(t, x_t)\, \partial t + A_1(t, x_t)\, \partial x_t\}\, e_t$$

which can be dealt with just as (B.4). The same remark applies to Lemma B.4.

(ii) Consider a family of connections ∇^ε, $\varepsilon \in [0, 1]$, depending smoothly on ε, and write τ_{st}^ε for parallel translation with respect to ∇^ε. We can replace τ_{st} by τ_{st}^ε in Lemma B.2: we find at (B.4) an equation of the form

$$De_t^\varepsilon = A^\varepsilon(\partial x_t)\, e_t^\varepsilon$$

A^ε depending smoothly on ε, and the same proof applies.

We denote a generic element of $C([0, 1], M)$ by ω and a generic element of $C([0, 1], \mathrm{GL}(E, F))$ by $\tilde{\omega}$. Then $x_t(\omega) = \omega_t$ and parallel translations and horizontal lifts are all functions of ω, defined almost surely for each of the Wiener measures considered, also defined when ω is absolutely continuous. Consider a family of connections ∇^ε, as in Remark B.3(ii). For $u \in \mathrm{GL}(E, F)$ with $\pi u = x$ we write μ_ε^u for the law of $(\tau_{t0}^\varepsilon u : 0 \leqq t \leqq 1)$ under $\mathbb{P}_\varepsilon^x$. Then for $F: C([0, 1], \mathrm{GL}(E, F)) \to \mathbb{R}$ continuous and bounded and for $A \subseteq C([0, 1], \mathrm{GL}(E, F))$ closed, we know by Ventcell–Freidlin estimates that

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon^u \left(1_A \exp\left\{ -\frac{F}{\varepsilon} \right\} \right) \leqq -\inf_{\tilde{\omega} \in A} \{ I_u(\tilde{\omega}) + F(\tilde{\omega}) \} \tag{B.5}$$

where $I_u(\tilde{\omega}) = \frac{1}{2} \int_0^1 |\dot{\omega}_s|^2\, ds$ if $\tilde{\omega}$ is the horizontal lift by ∇^0 of some absolutely continuous path ω in M, starting from u, and $I_u(\tilde{\omega}) = \infty$ otherwise.

Lemma B.4 *Suppose there is a unique minimal geodesic γ from x to y in time 1. Then for any metric on F and any $\eta > 0$, as $\varepsilon \downarrow 0$*

$$\mathbb{P}_\varepsilon^{x,y,1}(|\tau_{01}^\varepsilon - \tau_{01}^0(\gamma)| > \eta) \to 0 .$$

Proof. Choose a complete metric (distance function) ρ on $GL(E, F)$. Fix $u_0, u_1 \in GL(E, F)$ with $\pi u_0 = x$ and $\pi u_1 = y$. Then for every $\eta > 0$ there is a $\delta > 0$ such that $|\tau^{\varepsilon}_{01} - \tau^0_{01}(\gamma)| > \eta$ implies either

$$\rho(\tau^{\varepsilon}_{\frac{1}{2}0} u_0, \tau^0_{\frac{1}{2}0}(\gamma) u_0) > \delta \quad \text{or} \quad \rho(\tau^{\varepsilon}_{\frac{1}{2}1} u_1, \tau^0_{\frac{1}{2}1}(\gamma) u_1) > \delta .$$

By time symmetry it then suffices to show

$$\mathbb{P}^{x,y,1}_{\varepsilon}(\rho(u^{\varepsilon}_{\frac{1}{2}}, u^0_{\frac{1}{2}}(\gamma)) \geqq \delta) \to 0$$

where $u^{\varepsilon}_t = \tau^{\varepsilon}_{t0} u_0$. We know that

$$\left.\frac{d\mathbb{P}^{x,y,1}_{\varepsilon}}{d\mathbb{P}^{x}_{\varepsilon}}\right|_{\mathscr{F}_{\frac{1}{2}}} = \frac{p_{\varepsilon/2}(x_{\frac{1}{2}}, y)}{p_{\varepsilon}(x, y)} .$$

Set $F(\tilde{\omega}) = d^2(\pi\tilde{\omega}_{\frac{1}{2}}, y)$ and $A = \{\tilde{\omega} : \rho(\tilde{\omega}_{\frac{1}{2}}, u^0_{\frac{1}{2}}(\gamma)) \geqq \delta\}$. Then by the basic estimate (B.1), for all $\beta \in (0, \infty)$

$$\mathbb{P}^{x,y,1}_{\varepsilon}(\rho(u^{\varepsilon}_{\frac{1}{2}}, u^0_{\frac{1}{2}}(\gamma)) \geqq \delta) \leqq \frac{C(M, \beta)}{(\varepsilon/2)^{n/2}} \frac{\mu^{u_0}_{\varepsilon}\left(1_A \exp\left\{-\dfrac{F}{\varepsilon(1+\beta)}\right\}\right)}{p_{\varepsilon}(x, y)} .$$

By Varadhan's estimate (B.2)

$$\liminf_{\varepsilon \downarrow 0} \varepsilon \log p_{\varepsilon}(x, y) \geqq -\frac{d^2(x, y)}{2}$$

and by (B.5)

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log \mu^{u_0}_{\varepsilon}\left(1_A \exp\left\{-\frac{F}{\varepsilon(1+\beta)}\right\}\right) \leqq -\inf_{\tilde{\omega} \in A} \{I_{u_0}(\tilde{\omega}) + F(\tilde{\omega})/(1+\beta)\}$$

$$\leqq -\inf_{\tilde{\omega} \in A} \{I_{u_0}(\tilde{\omega}) + F(\tilde{\omega})\}/(1+\beta) . \qquad \text{(B.6)}$$

By the usual weak compactness argument, the infimum in (B.6) is attained, so, $\beta \in (0, \infty)$ being arbitrary, it suffices to show that for all $\tilde{\omega} \in A$

$$I_{u_0}(\tilde{\omega}) + d^2(\pi\tilde{\omega}_{\frac{1}{2}}, y) > d^2(x, y)/2 . \qquad \text{(B.7)}$$

Now

$$d^2(x, z) = \inf \frac{1}{2} \int_0^{\frac{1}{2}} |\dot{\omega}_s|^2 \, ds$$

where the infimum is taken over all absolutely continuous paths ω with $\omega_0 = x$ and $\omega_{\frac{1}{2}} = z$. Therefore $I_{u_0}(\tilde{\omega}) \geqq d^2(x, \pi\tilde{\omega}_{\frac{1}{2}})$, so equality in (B.7) would imply

$$d^2(x, \pi\tilde{\omega}_{\frac{1}{2}}) + d^2(\pi\tilde{\omega}_{\frac{1}{2}}, y) \leqq d^2(x, y)/2$$

or

$$\inf \frac{1}{2} \int_0^1 |\dot{\omega}_s|^2 \, ds \leqq d^2(x, y)/2$$

where the infimum is taken over all absolutely continuous paths ω with $\omega_0 = x$, $\omega_{\frac{1}{2}} = \pi\tilde{\omega}_{\frac{1}{2}}$ and $\omega_1 = y$. But we have assumed there is a unique minimal geodesic γ from x to y in time 1 so this would imply $\pi\tilde{\omega}_{\frac{1}{2}} = \gamma_{\frac{1}{2}}$. Moreover equality in (B.7) would then also force $I_{u_0}(\tilde{\omega}) = d^2(x, \gamma_{\frac{1}{2}})$. Since there is a unique minimal geodesic, this would imply $\pi\tilde{\omega}_s = \gamma_s$ for $0 \leqq s \leqq \frac{1}{2}$ and then $\tilde{\omega}_{\frac{1}{2}} = u^0_{\frac{1}{2}}(\gamma)$ which is impossible for $\tilde{\omega} \in A$. □

References

1. Ben–Arous, G.: Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. Ann. Sci. Ec. Norm. Super. **21**, 307–332 (1988)
2. Bismut, J.-M.: Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. Probab. Theory Relat. Fields **56**, 469–505 (1981)
3. Bismut, J.-M.: Large deviations and the Malliavin calculus. (Prog. Math. vol. 45) Boston: Birkhäuser 1984
4. Carverhill, A.P., Elworthy, K.D.: Flows of stochastic dynamical systems: the functional analytic approach. Probab. Theory Relat. Fields **65**, 245–268 (1983)
5. Elworthy, K.D.: Stochastic differential equations on manifolds. (Lond. Math. Soc. Lect. Note Ser., vol. 70) Cambridge: Cambridge University Press 1982
6. Elworthy, K.D.: Stochastic flows on Riemannian manifolds. In: Pinsky, M.A., Wihstutz, V. (eds.) Diffusion processes and related problems in analysis, vol. II. (Prog. Probab., vol. 27, pp. 37–72) Boston: Birkhäuser 1992
7. Li, P., Yau, S.T.: On the parabolic kernel of the Schrödinger operator. Acta Math. **156**, 153–201 (1986)
8. Norris, J.R.: Simplified Malliavin calculus. In: Azéma, J., Yor, M. (eds.) Séminaire de probabilités XX. (Lect. Notes Math., vol. 1204, pp. 101–130) Berlin Heidelberg New York: Springer 1986
9. Norris, J.R.: A complete differential formalism for stochastic calculus in manifolds. (Cambridge preprint)
10. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261–286 (1966)