ASIAN J. MATH.
Vol. 1, No. 1, pp. 17–33, March 1997

002

SHORT TIME BEHAVIOR OF LOGARITHMIC DERIVATIVES OF THE HEAT KERNEL*

DANIEL W. STROOCK† AND JAMES TURETSKY‡

Abstract. Let M be a compact, connected Riemannian manifold. Let $p_t(x,y)$ be the fundamental solution to Cauchy initial value problem for the heat equation $\frac{\partial u}{\partial t} = \frac{1}{2}\Delta_M u$, where Δ_M is the Levi-Civita Laplacian. The purpose of this note is to study the asymptotic behavior of logarithmic derivatives of $\log p_t(\,\cdot\,,y)$ at x as $t \searrow 0$. In particular, we show that a dramatic change takes place when x is at the cut-locus of y.

1. Introduction. Let M be a connected, compact, d-dimensional Riemannian manifold, and use Δ_M to denote the standard (i.e., the Levi-Civita) Laplacian on $C^\infty(M;\mathbb{R})$. Next, consider the Cauchy initial value problem for the associated heat equation:

$$\frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\Delta_M u(t,x) \quad \text{with} \quad \lim_{t\searrow 0} u(t,x) = f(x).$$

By elliptic regularity theory and the strong maximum principle, there is a unique smooth function $t \in (0,\infty) \longmapsto p_t \in C^\infty\big(M \times M;(0,\infty)\big)$ such that

$$u(t,x) = \int_M f(y)p_t(x,y)\lambda_M(dy) \quad \text{for every } f \in C(M),$$

where λ_M stands for the Riemannian measure on M. In fact, because Δ_M is essentially self-adjoint in $L^2(\lambda_M)$, $p_t(x,y) = p_t(y,x)$.

In this paper we will analyze logarithmic derivatives of $p_T(\,\cdot\,,y)$, thereby extending results in [10] about the first and second order logarithmic derivatives. To be more precise, recall Varadhan's formula (cf. [17])

$$\lim_{T\searrow 0} T\log p_T(x,y) = -\frac{\operatorname{dist}(x,y)^2}{2} \quad \text{uniformly in } (x,y) \in M \times M, \tag{1.1}$$

where $\operatorname{dist}(x,y)$ denotes the Riemannian distance between x and y. Our goal is to examine the extent to which the limit in (1.1) can be made to commute with derivatives. Obviously, because the smoothness of $\operatorname{dist}(\,\cdot\,,y)$ breaks down at the cut-locus $\operatorname{Cut}(y)$, one expects that there should be a distinction between $x \notin \operatorname{Cut}(y)$ and $x \in \operatorname{Cut}(y)$. Indeed, this suspicion was confirmed in [10] (cf. Corollary 2.28 there), where it was shown that derivatives of first and second order commute with the limit as long as $x \notin \operatorname{Cut}(y)$, but that, in general, problems occur when $x \in \operatorname{Cut}(y)$. In fact, when $x \in \operatorname{Cut}(y)$, second order logarithmic derivatives of $p_T(\,\cdot\,,y)$ at x can diverge like T^{-2} as $T \searrow 0$ (cf. Theorem 2.34 in [10]). In the present article, we will extend these results to derivatives of all orders. In particular (cf. Theorem 4.1 below), we will show

---

* Received January 8, 1997; accepted for publication (in revised form) February 21, 1997.

† Massachusetts Institute of Technology and Harvard University M.I.T., 2-272, 77 Mass. Ave., Cambridge, MA 02139, USA (dws@math.mit.edu). The author was partly supported by NSF grant 9625782-DMS.

‡ Harvard University, Department of Mathematics, 1 Oxford St., Cambridge, MA 02138, USA (turetsky@math.harvard.edu). The second author was supported by the NSF Graduate Fellowship Award.

that derivatives of all orders commute with the limit in (1.1) when $x \notin \mathrm{Cut}(y)$. Further, when $x \in \mathrm{Cut}(y)$, we show that, in general, nth order logarithmic derivatives of $p_T(\,\cdot\,, y)$ at x can be as bad as T^{-n}, thereby showing that one cannot hope to improve estimates obtained in [16].

Our strategy will be as follows. In §2 we develop explicit formulas for covariant logarithmic derivatives of $p_T(\,\cdot\,, y)$ in terms of integrals with respect to Wiener's measure (cf. (2.11)). These formulas are valid whether or not x is at the cut-locus of y. What they do is express an nth order logarithmic derivative of $p_T(\,\cdot\,, y)$ as a sum of the form

$$\sum_{m=1}^{n} T^{-m} a_m(T).$$

In general, the best that one can say is that the coefficients $a_m(T)$ stay bounded as $T \searrow 0$. Nonetheless, the fact which saves the day is the observation that $a_m(T)$ can be expressed as a finite linear combination of *cumulants* (cf. the Appendix) with respect to a measure obtained by conditioning μ_T.[1] In particular, when $x \notin \mathrm{Cut}(y)$, the theory of large deviations developed in [7] applies and says that this conditioned measure degenerates fast enough to make all the cumulants of each order $m \geq 2$ vanish at order T^m. On the other hand, when $x \in \mathrm{Cut}(y)$, the conditioned measure need not degenerate, with the result that these cumulants may survive as $T \searrow 0$. That is, the degeneration of these conditioned measures is intimately tied to properties of the exponential map at the points in question.

2. Formulas for logarithmic derivatives of the heat kernel.

In this section, we reformulate the differential calculus of §2 in [16] using the idea from §1 in [10] to make it amenable to the theory of large deviations as it was developed in [7]. Our main objective here is to express (cf. (2.12)) the logarithmic derivatives of the heat kernel in terms of conditional expectations on the Wiener space.

2.1. Differential geometric notation.

Because it leads to conceptual as well as notational simplification, we introduce $O(M)$, the bundle of orthonormal frames $\mathfrak{e}$ (cf. [2] for more information on the topic), and the fiber map $\pi : O(M) \longrightarrow M$ with takes the frame $\mathfrak{e}$ into the base point over which it lies. Given $\mathfrak{e} \in O(M)$ with $x = \pi(\mathfrak{e})$, we will identify $\mathfrak{e}$ with the isometric map $\mathbf{v} \in \mathbb{R}^d \longmapsto \mathfrak{e}\mathbf{v} \in T_xM$ which takes $\mathbf{v}$ into the element $\mathfrak{e}\mathbf{v}$ of T_xM whose coordinates in the frame $\mathfrak{e}$ are $\mathbf{v} \in \mathbb{R}^d$. Next, we use the Levi-Civita connection to determine the horizontal subspace $H_{\mathfrak{e}}O(M)$ of $T_{\mathfrak{e}}O(M)$; and, for $\mathbf{v} \in \mathbb{R}^d$, we define the canonical vector field $\mathfrak{E}(\mathbf{v})$ on $O(M)$ so that, at each $\mathfrak{e} \in O(M)$, $\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}$ is the horizontal lift of $\mathfrak{e}\mathbf{v}$ to $\mathfrak{e}$. That is, $\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}$ is the unique element of $H_{\mathfrak{e}}O(M)$ such that $d\pi\mathfrak{E}(\mathbf{v})_{\mathfrak{e}} = \mathfrak{e}\mathbf{v}$. Thus, if F is a smooth, tensor-valued function on M, then $\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}(F \circ \pi)$ is the covariant derivative of F at $\pi(\mathfrak{e})$ in the direction $\mathfrak{e}\mathbf{v}$. In particular, if $f \in C^\infty(M;\mathbb{R})$ and $n \geq 1$, then we define the nth order covariant derivative $\mathfrak{E}^{(n)}f : O(M) \longrightarrow (\mathbb{R}^d)^{\otimes n}$ of f so that

$$\begin{aligned}\left(\mathfrak{E}^{(n)}f \circ \pi, \mathbf{V}\right)_{(\mathbb{R}^d)^{\otimes n}} &= \mathfrak{E}(\mathbf{V})_{\mathfrak{e}} f \circ \pi \equiv \mathfrak{E}(\mathbf{v}^n)_{\mathfrak{e}} \circ \cdots \circ \mathfrak{E}(\mathbf{v}^1) f \circ \pi \\ &= \nabla_{\mathfrak{e}\mathbf{v}^n} \cdots \nabla_{\mathfrak{e}\mathbf{v}^1} f,\end{aligned}$$

---

[1] It is not too surprising that cumulants arise here. Indeed, at a somewhat casual level, $p_T(x, y)$ is the kernel for $e^{\frac{T}{2}\Delta_M}$ and we are looking at $\log p_T(\,\cdot\,, y)$.

when

$$\mathbf{V} = \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n \in (\mathbb{R}^d)^{\otimes n}.$$

Finally, it will be convenient to have introduced three important quantities which are inextricably tied to these considerations. The first of these is the solder form ω which assigns to a vector $\mathfrak{X}_{\mathfrak{e}} \in T_{\mathfrak{e}}O(M)$ the coordinates $\omega(\mathfrak{X}_{\mathfrak{e}})$ of $d\pi\mathfrak{X}_{\mathfrak{e}} \in T_{\pi(\mathfrak{e})}M$ in the frame $\mathfrak{e}$. That is, $d\pi\mathfrak{X}_{\mathfrak{e}} = \mathfrak{e}\omega(\mathfrak{X}_{\mathfrak{e}})$. Second, given $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in (\mathbb{R}^d)^2$, we define the curvature form $\mathfrak{e} \in O(M) \longmapsto \Phi_{\mathfrak{e}}(\boldsymbol{\xi}, \boldsymbol{\eta}) \in o(d)$ so[2] that

$$\mathfrak{e}\Phi_{\mathfrak{e}}(\boldsymbol{\xi}, \boldsymbol{\eta})\boldsymbol{\zeta} = \operatorname{Riem}\bigl(\mathfrak{e}\boldsymbol{\xi}, \mathfrak{e}\boldsymbol{\eta}\bigr)\mathfrak{e}\boldsymbol{\zeta}, \qquad \boldsymbol{\zeta} \in \mathbb{R}^d,$$

where

$$\operatorname{Riem}(X_x, Y_x)Z_x = \nabla_{[X,Y]_x} Z - \bigl[\nabla_X, \nabla_Y\bigr]_x Z$$

is the usual Riemann curvature tensor. Finally, we define the Ricci matrix $\mathfrak{e} \in O(M) \longrightarrow \operatorname{Ric}_{\mathfrak{e}} \in \operatorname{Hom}(\mathbb{R}^d; \mathbb{R}^d)$ by

$$\bigl(\boldsymbol{\zeta}, \operatorname{Ric}_{\mathfrak{e}}\boldsymbol{\xi}\bigr)_{\mathbb{R}^d} = \sum_{i=1}^{d} \bigl(\Phi_{\mathfrak{e}}(\boldsymbol{\zeta}, \boldsymbol{\eta}^i)\boldsymbol{\xi}, \boldsymbol{\eta}^i\bigr)_{\mathbb{R}^d}$$

for any orthonormal basis $(\boldsymbol{\eta}^1, \dots, \boldsymbol{\eta}^d)$ in $\mathbb{R}^d$. Using the symmetries of the Riemann curvature, it is easy to check that $\Phi_{\mathfrak{e}}(\boldsymbol{\xi}, \boldsymbol{\eta}) = -\Phi_{\mathfrak{e}}(\boldsymbol{\eta}, \boldsymbol{\xi})$ and that $\operatorname{Ric}_{\mathfrak{e}}$ is symmetric.

2.2. Calculus on Wiener space. As we said in the introduction, our analysis turns on our ability to represent $p_T(x, y)$ and its derivatives in terms of function space integrals with respect to Wiener's measure. For this reason, we introduce the separable Banach space $\mathfrak{W}$, with the uniform norm $\|\cdot\|_{\mathfrak{W}}$, of continuous paths $\mathbf{w} : [0, 1] \longrightarrow \mathbb{R}^d$ satisfying $\mathbf{w}(0) = \mathbf{0}$. Letting $\mathcal{B}$ be the Borel field over $\mathfrak{W}$, we use μ to denote the standard Wiener measure on $(\mathfrak{W}, \mathcal{B})$. That is, if $\mathcal{B}_t$ is the sub σ-algebra generated by $\mathbf{w} \in \mathfrak{W} \longmapsto \mathbf{w}(\tau) \in \mathbb{R}^d$ as τ runs through $[0, t]$, then, for each $0 \le t_1 < t_2 \le 1$, $\mathbf{w} \in \mathfrak{W} \longmapsto \mathbf{w}(t_2) - \mathbf{w}(t_1) \in \mathbb{R}^d$ under μ is a Gaussian random variable which is independent of $\mathcal{B}_{t_1}$ and has mean $\mathbf{0}$ and covariance $(t_2 - t_1)\mathbf{I}$. Next, for each $T \in (0, 1]$, let μ_T be the distribution of $\mathbf{w} \in \mathfrak{W} \longmapsto T^{\frac{1}{2}}\mathbf{w} \in \mathfrak{W}$ under μ.

For each $T \in (0, 1]$ and $\mathfrak{e} \in O(M)$, there is a μ_T-almost surely unique, $\{\mathcal{B}_t : t \in [0, 1]\}$-progressively measurable map $\mathfrak{F}_{\mathfrak{e},T} : [0, 1] \times \mathfrak{W} \longrightarrow O(M)$ which solves the Stratonovich stochastic differential equation

$$d\mathfrak{F}_{\mathfrak{e},T}(t, \mathbf{w}) = \mathfrak{E}\bigl(\circ d\mathbf{w}(t)\bigr)_{\mathfrak{F}_{\mathfrak{e},T}(t,\mathbf{w})} \qquad \text{with } \mathfrak{F}_{\mathfrak{e},T}(0, \mathbf{w}) = \mathfrak{e}.$$

Moreover, a standard application of Itô calculus leads to the relation

$$\mathbb{E}^{\mu_T}\Bigl[f \circ \pi\bigl(\mathfrak{F}_{\mathfrak{e},T}(1, \mathbf{w})\bigr)\Bigr] = \int_M f(y) p_T\bigl(\pi(\mathfrak{e}), y\bigr)\lambda_M(dy), \quad f \in C(M; \mathbb{R}), \tag{2.1}$$

which is the starting point for our entire program. In fact, starting from (2.1), one can show (cf. [6]) that there is a (weakly) continuous map $(\mathfrak{e}, T, y) \in O(M) \times (0, 1] \times M \longmapsto$

---

[2] $o(d)$ stands for Lie algebra of the orthogonal group $O(d)$, which we identify with the space of skew symmetric matrices. Also, notice we have adopted a different sign convention from the one in [2], where our $\Phi_{\mathfrak{e}}(\boldsymbol{\xi}, \boldsymbol{\eta})$ would be denoted by $-\Phi(\mathfrak{E}(\boldsymbol{\xi})_{\mathfrak{e}}, \mathfrak{E}(\boldsymbol{\eta})_{\mathfrak{e}})$.

$\mu_{T,\mathfrak{e}|y} \in M_1(\mathfrak{W})$ (the space of probability measures on $\mathfrak{W}$) such that

$$\int\limits_{\{\mathbf{w}:\,\pi\circ\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\in\Gamma\}} \Psi(\mathbf{w})\,\mu_T(d\mathbf{w}) = \int_\Gamma \left(\int_{\mathfrak{W}} \Psi(\mathbf{w})\,\mu_{T,\mathfrak{e}|y}(d\mathbf{w})\right) p_T\big(\pi(\mathfrak{e}),y\big)\,\lambda_M(dy) \tag{2.2}$$

for all Borel sets $\Gamma \subseteq M$ and bounded, Borel measurable $\Psi : \mathfrak{W} \longrightarrow \mathbb{R}$. That is, $y \in M \longmapsto \mu_{T,\mathfrak{e}|y} \in M_1(\mathfrak{W})$ is a continuous version of the conditional distribution of μ_T given $\pi \circ \mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w}) = y$. Alternatively, one can interpret (2.2) as saying that

$$p_T\big(\pi(\mathfrak{e}),y\big)\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\big[\Psi(\mathbf{w})\big] = \mathbb{E}^{\mu_T}\Big[\Psi(\mathbf{w})\,\delta_y \circ \pi\big(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\big)\Big], \tag{2.3}$$

where δ_y is the Dirac delta function relative to λ_M.

In order to pass from (2.1) to expressions for derivatives of $p_T(\,\cdot\,,y)$, we adopt the procedure developed in [5], [15], and [16]. To describe this procedure, we need to recall two families of perturbations of the paths $\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w})$: the *backward perturbation*

$$s \in \mathbb{R} \longmapsto \Big(\big[\overleftarrow{\Theta}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s),\,\big[\overleftarrow{\mathfrak{F}}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s)\Big) \in C\big([0,1];\mathbb{R}^d\big) \times C\big([0,1];O(M)\big)$$

for each $\mathbf{v} \in \mathbb{R}^d$, and the *forward perturbation*

$$s \in \mathbb{R} \longmapsto \Big(\big[\overrightarrow{\Theta}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s),\,\big[\overrightarrow{\mathfrak{F}}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s)\Big) \in C\big([0,1];\mathbb{R}^d\big) \times C\big([0,1];O(M)\big)$$

for each smooth $\mathfrak{e} \in O(M) \longmapsto \boldsymbol{\xi}_{\mathfrak{e}} \in C^2\big([0,1];\mathbb{R}^d\big)$ satisfying $\boldsymbol{\xi}_{\mathfrak{e}}(1) = \mathbf{0}$. Namely, given $\mathbf{v} \in \mathbb{R}^d$, determine the family of paths $(s,\mathfrak{e}) \in \mathbb{R} \times O(M) \longmapsto \mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s) \in O(M)$ by[3]

$$\mathfrak{f}'_{\mathfrak{e},\mathbf{v}}(s) \equiv \frac{d}{ds}\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s) = \mathfrak{E}(\mathbf{v})_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s)} \quad \text{with } \mathfrak{f}_{\mathfrak{e},\mathbf{v}}(0) = \mathfrak{e}.$$

Then the backward perturbation has the properties that

B1 $\quad s \rightsquigarrow \Big(\big[\overleftarrow{\Theta}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s),\,\big[\overleftarrow{\mathfrak{F}}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s)\Big)$ is smooth for μ_T-a.e. $\mathbf{w} \in \mathfrak{W}$;

B2 $\quad$ for each $s \in \mathbb{R}$, $\mathbf{w} \rightsquigarrow \Big(\big[\overleftarrow{\Theta}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s),\,\big[\overleftarrow{\mathfrak{F}}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s)\Big)$ has the same distribution under μ_T as $\mathbf{w} \rightsquigarrow \Big(\mathbf{w},\mathfrak{F}_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s)}(\,\cdot\,,\mathbf{w})\Big)$;

B3

$$\omega\Big(\big[\overleftarrow{\mathfrak{F}}_{\mathfrak{e},T,\mathbf{v}}(t,\mathbf{w})\big]'(0)\Big) = A_{\mathfrak{e},T}(t,\mathbf{w})\mathbf{v} \quad \text{where}$$

$$\dot{A}_{\mathfrak{e},T}(t,\mathbf{w}) + \frac{T}{2}\mathrm{Ric}_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})}A_{\mathfrak{e},T}(t,\mathbf{w}) = \mathbf{0} \qquad \text{with } A_{\mathfrak{e},T}(0,\mathbf{w}) = \mathbf{I}.$$

---

[3] We use *prime* to denote differentiation with respect to s and *dot* for differentiation with respect to t.

By combining B2 with B3, we see that, at least for sufficiently smooth functions,

$$\begin{aligned}
&\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}\mathbb{E}^{\mu_T}\Big[G\big(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w})\big)\,f\circ\pi\big(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\big)\Big]\\
&\quad=\mathbb{E}^{\mu_T}\bigg[\frac{d}{ds}G\Big(\big[\overleftarrow{\Theta}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s),\big[\overleftarrow{\mathfrak{F}}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})\big](s)\Big)\Big|_{s=0}\,f\circ\pi\big(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\big)\bigg]\\
&\qquad+\mathbb{E}^{\mu_T}\Big[G\big(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w})\big)\mathfrak{E}\big(A_{\mathfrak{e},T}(1,\mathbf{w})\mathbf{v}\big)_{\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})}f\circ\pi\Big].
\end{aligned}\tag{2.4}$$

The role of the forward perturbation is that it allows us to remove the derivative from f on the right hand side of (2.4). Namely, the forward perturbation has the properties that

F1 $\quad s\rightsquigarrow\Big(\big[\overrightarrow{\Theta}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s),\big[\overrightarrow{\mathfrak{F}}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s)\Big)$ is smooth for μ_T-a.e. $\mathbf{w}\in\mathfrak{W}$;

F2 $\quad$ for each $s\in\mathbb{R}$, the distribution of $\mathbf{w}\rightsquigarrow\Big(\big[\overrightarrow{\Theta}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s),\big[\overrightarrow{\mathfrak{F}}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s)\Big)$ under μ_T is equivalent to that of $\mathbf{w}\rightsquigarrow\big(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w})\big)$;

F3 $$\omega\Big(\big[\overrightarrow{\mathfrak{F}}_{\mathfrak{e},T,\boldsymbol{\xi}}(t,\mathbf{w})\big]'(0)\Big)=A_{\mathfrak{e},T}(t,\mathbf{w})\big(\boldsymbol{\xi}_{\mathfrak{e}}(0)-\boldsymbol{\xi}_{\mathfrak{e}}(t)\big);$$

F4 $\quad$ if $\big[R_{\boldsymbol{\xi}}(\mathfrak{e},T,\mathbf{w})\big](s)$ is the Radon–Nikodym factor such that

$$\begin{aligned}
&\mathbb{E}^{\mu_T}\Big[\big[R_{\boldsymbol{\xi}}(\mathfrak{e},T,\mathbf{w})\big](s)\,G\Big(\big[\overrightarrow{\Theta}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s),\big[\overrightarrow{\mathfrak{F}}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})\big](s)\Big)\Big]\\
&\qquad\qquad=\mathbb{E}^{\mu_T}\big[G\big(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w})\big)\big],
\end{aligned}$$

then $\displaystyle\frac{d}{ds}\big[R_{\boldsymbol{\xi}}(\mathfrak{e},T,\mathbf{w})\big](s)\Big|_{s=0}=\frac{1}{T}\int_0^1\Big(A_{\mathfrak{e},T}(t,\mathbf{w})\dot{\boldsymbol{\xi}}_{\mathfrak{e}}(t),d\mathbf{w}(t)\Big)_{\mathbb{R}^d}.$

In particular, by combining (2.4) with F2–F4, we obtain the formula

$$\begin{aligned}
&\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}\mathbb{E}^{\mu_T}\big[G(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w}))\,f\circ\pi(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w}))\big]\\
&\quad=\mathbb{E}^{\mu_T}\bigg[\bigg(\big[\mathcal{D}^T_{\boldsymbol{\xi}}G\big](\mathbf{w})-\frac{1}{T}\int_0^1\Big(A_{\mathfrak{e},T}(t,\mathbf{w})\dot{\boldsymbol{\xi}}_{\mathfrak{e},\mathbf{v}}(t),d\mathbf{w}(t)\Big)_{\mathbb{R}^d}G(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w}))\bigg)\\
&\qquad\qquad\times f\circ\pi(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w}))\bigg]
\end{aligned}\tag{2.5}$$

when $\boldsymbol{\xi}_{\mathfrak{e}}(0)\equiv\mathbf{v}$, and where

$$\begin{aligned}
\big[\mathcal{D}^T_{\boldsymbol{\xi}}G\big](\mathbf{w})\equiv\frac{d}{ds}\Big(&G\big([\overleftarrow{\Theta}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})](s),[\overleftarrow{\mathfrak{F}}_{\mathfrak{e},T,\mathbf{v}}(\,\cdot\,,\mathbf{w})](s)\big)\\
&-G\big([\overrightarrow{\Theta}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})](s),[\overrightarrow{\mathfrak{F}}_{\mathfrak{e},T,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{w})](s)\big)\Big)\Big|_{s=0}.
\end{aligned}\tag{2.6}$$

Finally, starting from (2.5) and taking (cf. (2.2) and (2.3)) $f=\delta_y$, we arrive at our basic formula:

$$\begin{aligned}
&\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}\mathbb{E}^{\mu_T}\big[G\big(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w})\big)\delta_y\circ\pi\big(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\big)\big]p_T(\pi(\mathfrak{e}),y)^{-1}\\
&\quad=\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\bigg[\big[\mathcal{D}^T_{\boldsymbol{\xi}}G\big](\mathbf{w})-\frac{1}{T}\int_0^1\big(A_{\mathfrak{e},T}(t,\mathbf{w})\dot{\boldsymbol{\xi}}_{\mathfrak{e},\mathbf{v}}(t),d\mathbf{w}(t)\big)_{\mathbb{R}^d}G(\mathbf{w},\mathfrak{F}_{\mathfrak{e},T}(\,\cdot\,,\mathbf{w}))\bigg],
\end{aligned}\tag{2.7}$$

when $\boldsymbol{\xi}_{\mathfrak{e}}(0) \equiv \mathbf{v}$ and $\boldsymbol{\xi}_{\mathfrak{e}}(1) \equiv \mathbf{0}$.

2.3. Computation of logarithmic derivatives. With (2.7) in hand, the only thing between us and our formula for logarithmic derivatives of the $p_T(\,\cdot\,, y)$ is a little combinatorics.

Following [13], for $1 \le m \le n$, we use $\Sigma_{m,n}$ to denote the set of all partitions $\{\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^m\}$ of the set $\mathcal{S}_n = \{1, \dots, n\}$ into m nonempty, *increasingly* ordered subsets $\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^m$, called *blocks*. In other words, $\boldsymbol{\beta}^k \ne \emptyset$, for $1 \le k \le m$; $\boldsymbol{\beta}^j \cap \boldsymbol{\beta}^k = \emptyset$, for $1 \le j < k \le m$; and $\bigcup_{k=1}^{m} \boldsymbol{\beta}^k = \mathcal{S}_n$. Finally, given vectors $\mathbf{v}^1, \dots, \mathbf{v}^n \in \mathbb{R}^d$ and an ordered block $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)$, we put

$$\mathbf{V}^{\boldsymbol{\beta}} = \mathbf{v}^{\beta_1} \otimes \cdots \otimes \mathbf{v}^{\beta_m} \in (\mathbb{R}^d)^{\otimes m}.$$

With this notation, one can use elementary differential calculus and induction to derive the formula (cf. §2.1)

$$\begin{aligned} &\nabla_{\mathfrak{e}\mathbf{v}^n} \cdots \nabla_{\mathfrak{e}\mathbf{v}^1} \log p_T(\,\cdot\,, y) \\ &\quad = \mathfrak{C}(\mathbf{V})_{\mathfrak{e}} \log p_T(\pi(\,\cdot\,), y) \\ &\quad = \sum_{m=1}^{n} \frac{(-1)^{m-1}(m-1)!}{p_T(\pi(\mathfrak{e}), y)^m} \sum_{\{\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^m\} \in \Sigma_{m,n}} \prod_{k=1}^{m} \mathfrak{C}(\mathbf{V}^{\boldsymbol{\beta}^k})_{\mathfrak{e}}\, p_T(\pi(\,\cdot\,), y). \end{aligned} \tag{2.8}$$

In the following statement, if $\boldsymbol{\Xi} = \boldsymbol{\xi}^1 \otimes \cdots \otimes \boldsymbol{\xi}^n$ where $\mathfrak{e} \longmapsto \boldsymbol{\xi}^k_{\mathfrak{e}} \in C^2\big([0,1]; \mathbb{R}^d\big)$ is smooth for each $1 \le k \le n$, then we use induction on $n \ge 1$ to define $B_{\boldsymbol{\Xi}}(\mathfrak{e}, T, \,\cdot\,) : \mathfrak{W} \longrightarrow \mathbb{R}$ so that

$$B_{\boldsymbol{\Xi}}(\mathfrak{e}, T, \mathbf{w}) = \begin{cases} \displaystyle\int_0^1 \big(A_{\mathfrak{e},T}(t, \mathbf{w}) \dot{\boldsymbol{\xi}}^1_{\mathfrak{e}}(t), d\mathbf{w}(t)\big)_{\mathbb{R}^d} & \text{when } n = 1, \\ \big[\mathcal{D}^T_{\boldsymbol{\xi}^n} B_{\boldsymbol{\xi}^1 \otimes \cdots \otimes \boldsymbol{\xi}^{n-1}}(\mathfrak{e}, T, \,\cdot\,)\big](\mathbf{w}) & \text{when } n \ge 2. \end{cases} \tag{2.9}$$

Theorem 2.10. *Let* $n \ge 1$ *and* $\mathbf{v}^1, \dots, \mathbf{v}^n \in \mathbb{R}^d$ *be given, and suppose* $\mathfrak{e} \in O(M) \longmapsto \boldsymbol{\xi}^k_{\mathfrak{e}} \in C^2\big([0,1]; \mathbb{R}^d\big)$, $1 \le k \le n$, *are smooth and chosen so that* $\boldsymbol{\xi}^k_{\mathfrak{e}}(0) = \mathbf{v}^k$ *and* $\boldsymbol{\xi}^k_{\mathfrak{e}}(1) = \mathbf{0}$ *for all* $\mathfrak{e} \in O(M)$. *Then, for every* $p \in [1, \infty)$,

$$\sup_{\substack{\mathfrak{e} \in O(M) \\ T \in (0,1]}} T^n \big\| B_{\boldsymbol{\Xi}}(\mathfrak{e}, T, \mathbf{w}) \big\|_{L^p(\mu_{T,\mathfrak{e}|y})} < \infty. \tag{2.11}$$

and we have the formula

$$\begin{aligned} &\nabla_{\mathfrak{e}\mathbf{v}^n} \cdots \nabla_{\mathfrak{e}\mathbf{v}^1} \log p_T(\,\cdot\,, y) \\ &\quad = \sum_{m=1}^{n} (-T)^{-m} \sum_{\{\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^m\} \in \Sigma_{m,n}} \mathcal{C}^{\mu_{T,\mathfrak{e}|y}}_m \big(B_{\boldsymbol{\Xi}^{\boldsymbol{\beta}^1}}(\mathfrak{e}, T, \mathbf{w}), \dots,\ B_{\boldsymbol{\Xi}^{\boldsymbol{\beta}^m}}(\mathfrak{e}, T, \mathbf{w})\big), \end{aligned} \tag{2.12}$$

where $\boldsymbol{\Xi}^{\boldsymbol{\beta}} \equiv \boldsymbol{\xi}^{\beta_1} \otimes \cdots \otimes \boldsymbol{\xi}^{\beta_k}$ *when* $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)$ *and (cf.* (A.1) *and* (A.2)*)* $\mathcal{C}^{\mu_{T,\mathfrak{e}|y}}_m(X_1, \dots, X_m)$ *is the* m*th cumulant of the random variables* $X_1, \dots, X_m$ *computed with respect to the measure* $\mu_{T,\mathfrak{e}|y}$.

Proof. We defer the proof of (2.11) to §A.2 below and turn immediately to the verification of (2.12). According to our notation of §2.1 with $\mathbf{V} = \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n$,

$$\mathfrak{C}(\mathbf{V})_{\mathfrak{e}} \log p_T(\pi(\,\cdot\,), y) = \nabla_{\mathfrak{e}\mathbf{v}^n} \cdots \nabla_{\mathfrak{e}\mathbf{v}^1} \log p_T(\,\cdot\,, y).$$

By (2.7) with $G \equiv 1$ we see that[4]

$$\frac{\mathfrak{E}(\mathbf{v})p_T(\pi(\,\cdot\,),y)}{p_T(\pi(\,\cdot\,),y)}(\mathfrak{e}) = -T^{-1}\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\big[B_{\boldsymbol{\xi}}(\mathfrak{e},T,\mathbf{w})\big] \quad \text{when } \boldsymbol{\xi}_{\mathfrak{e}}(0) \equiv \mathbf{v}.$$

More generally, starting from here, working by induction, and making repeated application of (2.5), (2.7), and Leibnitz's rule, one finds that

$$\frac{\mathfrak{E}(\mathbf{V})_{\mathfrak{e}}p_T(\pi(\,\cdot\,),y)}{p_T(\pi(\mathfrak{e}),y)} = \mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\left[\sum_{m=1}^{n}(-T)^{-m}\sum_{\{\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m\}\in\Sigma_{m,n}}\prod_{k=1}^{m}B_{\boldsymbol{\Xi}_{\boldsymbol{\beta}^k}}(\mathfrak{e},T,\mathbf{w})\right].$$

Finally, the passage from this to (2.12) is a simple application of the combinatorial fact contained in (A.2). □

3. A little bit of large deviations.

3.1. General setting. We begin by re-formulating the key result from [7] about large deviations in a form which is suitable for the present setting. Since we will be making frequent reference to it, we recall here the *Cameron–Martin* subspace $\mathbf{H}$ of absolutely continuous $\mathbf{h} \in \mathfrak{W}$ for which

$$\|\mathbf{h}\|_{\mathbf{H}} \equiv \left(\int_0^1 \big|\dot{\mathbf{h}}(t)\big|^2\,dt\right)^{\frac{1}{2}} < \infty.$$

Next, to understand the intuition which underlies what follows, keep in mind Feynman's heuristic picture (cf. §4.2 of [14]) in which μ_T is given by the fanciful but compelling formula

$$\mu_T(d\mathbf{w}) = (2\pi)^{-\frac{1}{2}\dim(\mathbf{H})}\exp\left(-\frac{\|\mathbf{w}\|_{\mathbf{H}}^2}{2T}\right)\lambda_{\mathbf{H}}(d\mathbf{w}), \tag{3.1}$$

where $\lambda_{\mathbf{H}}$ is the "Lebesgue measure" on $\mathbf{H}$.

Of course, (3.1) makes essential zero mathematical sense. Nonetheless, as the theory developed in [7] shows, predictions based on (3.1) become increasingly accurate as $T \searrow 0$ and, when interpreted carefully, become correct at "$T = 0$". For example, define $(t,\mathfrak{e},\mathbf{g}) \in [0,1]\times O(M)\times\mathbf{H} \longmapsto \mathfrak{F}_{\mathfrak{e},0}(t,\mathbf{g}) \in O(M)$ so that

$$\dot{\mathfrak{F}}_{\mathfrak{e},0}(t,\mathbf{g}) = \mathfrak{E}\big(\dot{\mathbf{g}}(t)\big)_{\mathfrak{F}_{\mathfrak{e},0}(t,\mathbf{g})} \quad \text{with } \mathfrak{F}_{\mathfrak{e},0}(0,\mathbf{g}) = \mathfrak{e}, \tag{3.2}$$

and set

$$\mathbf{H}(\mathfrak{e},y) = \big\{\mathbf{g}\in\mathbf{H} : \pi\circ\mathfrak{F}_{\mathfrak{e},0}(1,\mathbf{g}) = y\big\}. \tag{3.3}$$

Next, let $\mathcal{K}$ be a compact subset of $O(M)$, and suppose that, for each $T \in (0,1]$, $f_T : \mathcal{K}\times\mathfrak{W} \longrightarrow \mathbb{R}$ is a measurable function. Further, assume that $f_T \restriction \mathcal{K}\times\mathbf{H}$ tends nicely to a smooth limit function $f_0 : \mathcal{K}\times\mathbf{H} \longrightarrow \mathbb{R}$, and define

$$Mf_0(\mathfrak{e},y) = \max\left\{f_0(\mathfrak{e},\mathbf{g}) - \frac{\|\mathbf{g}\|_{\mathbf{H}}^2}{2} : \mathbf{g}\in\mathbf{H}(\mathfrak{e},y)\right\}, \quad (\mathfrak{e},y)\in O(M)\times M. \tag{3.4}$$

---

[4] This is a variant on a beautiful formula proved by J-M Bismut in [1], a fact which accounts for our choice of "B" here.

Then (3.1) should make one believe that, under suitable technical conditions, (cf. (2.3))

$$\mathbb{E}^{\mu_T}\left[\exp\left(\frac{f_T(\mathfrak{e},\mathbf{w})-Mf_0(\mathfrak{e},y)}{T}\right)\delta_y\circ\pi\big(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\big)\right].$$

is asymptotic, uniformly in $\mathfrak{e}\in\mathcal{K}$, to an expression which is independent of $\{f_T : T\in(0,1]\}$, and this belief is precisely what Theorem 4.21 in [7] justifies. Aside from technical conditions, the theory requires that there exists an $\mathfrak{e}\in\mathcal{K}\longmapsto\mathbf{h}_{\mathfrak{e}}\in\mathbf{H}(\mathfrak{e},y)$ such that

$$f_0(\mathfrak{e},\mathbf{h})-\frac{\|\mathbf{h}\|_{\mathbf{H}}^2}{2}<f_0(\mathfrak{e},\mathbf{h}_{\mathfrak{e}})-\frac{\|\mathbf{h}_{\mathfrak{e}}\|_{\mathbf{H}}^2}{2}=Mf_0(\mathfrak{e},y)\tag{3.5}$$

for all $\mathbf{h}\in\mathbf{H}(\mathfrak{e},y)\setminus\{\mathbf{h}_{\mathfrak{e}}\}$, and

$$\frac{d^2}{ds^2}\left(f_0\big(\mathfrak{e},[\Theta(\,\cdot\,)](s)\big)-\frac{\|[\Theta(\,\cdot\,)](s)\|_{\mathbf{H}}^2}{2}\right)\Bigg|_{s=0}<0,\tag{3.6}$$

for any smooth $s\in\mathbb{R}\longmapsto[\Theta(\,\cdot\,)](s)\in\mathbf{H}(\mathfrak{e},y)$ with $[\Theta(\,\cdot\,)](0)=\mathbf{h}_{\mathfrak{e}}$ and $[\Theta(\,\cdot\,)]'(0)\neq\mathbf{0}$.

3.2. The case at hand. In order to describe the applications of these considerations here, we must first discuss the perturbations of the path $\mathfrak{F}_{\mathfrak{e},0}(\,\cdot\,,\mathbf{g})$, which replace the backward and forward perturbations of §2.2 when $T=0$. For this purpose, let $\mathbf{H}(O(M))$ denote the set of absolutely continuous $\mathfrak{f}:[0,1]\longrightarrow O(M)$ such that

$$\dot{\mathfrak{f}}(t)\in H_{\mathfrak{f}(t)}O(M)\quad\text{for almost every } t\in[0,1]\quad\text{and}\quad\int_0^1\big|\omega\big(\dot{\mathfrak{f}}(t)\big)\big|_{\mathbb{R}^d}^2\,dt<\infty.$$

Given an $\mathfrak{e}\in O(M)$, a smooth map $\mathfrak{e}\in O(M)\longmapsto\boldsymbol{\xi}_{\mathfrak{e}}\in\mathbb{R}^d\oplus\mathbf{H}$, and a $\mathbf{g}\in\mathbf{H}$, there exists (cf. Theorem 2.5 in [5]) a unique smooth $s\in\mathbb{R}\longmapsto\big[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{g})\big](s)\in C\big([0,1];O(M)\big)$ such that

$$\text{P1}\qquad\big[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(0,\mathbf{g})\big]'(s)=\mathfrak{E}\big(\boldsymbol{\xi}_{\mathfrak{e}}(0)\big)_{[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(0,\mathbf{g})](s)}\quad\text{and}\quad\big[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{g})\big](0)=\mathfrak{F}_{\mathfrak{e},0}(\,\cdot\,,\mathbf{g});$$

$$\text{P2}\qquad\begin{aligned}&\text{for each } s\in\mathbb{R}:\ \big[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{g})\big](s)\in\mathbf{H}\big(O(M)\big)\\&\text{and }\omega\Big(\big[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(t,\mathbf{g})\big]'(s)\Big)=\boldsymbol{\xi}_{[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(0,\mathbf{g})](s)}(0)-\boldsymbol{\xi}_{[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(0,\mathbf{g})](s)}(t).\end{aligned}$$

Next, set

$$\big[\Theta_{\mathfrak{e},0,\boldsymbol{\xi}}(t,\mathbf{g})\big](s)=\int_0^t\omega\Big(\big[\dot{\mathfrak{F}}_{\mathfrak{e},0,\boldsymbol{\xi}}(\tau,\mathbf{g})\big](s)\Big)\,d\tau,$$

and, for smooth $g:O(M)\times\mathbf{H}\longrightarrow\mathbb{R}$, define

$$\big[\mathcal{D}^0_{\boldsymbol{\xi}}g\big](\mathfrak{e},\mathbf{g})=\frac{d}{ds}g\Big(\big[\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}(0,\mathbf{g})\big](s),\big[\Theta_{\mathfrak{e},0,\boldsymbol{\xi}}(\,\cdot\,,\mathbf{g})\big](s)\Big)\Big|_{s=0}.$$

In particular, when $\boldsymbol{\xi}_{\mathfrak{e}}\equiv\mathbf{h}\in\mathbf{H}$, we write $\mathfrak{F}_{\mathfrak{e},0,\mathbf{h}}$, $\Theta_{\mathfrak{e},0,\mathbf{h}}$, and $\mathcal{D}^0_{\mathbf{h}}$ in place of $\mathfrak{F}_{\mathfrak{e},0,\boldsymbol{\xi}}$, $\Theta_{\mathfrak{e},0,\boldsymbol{\xi}}$, and $\mathcal{D}^0_{\boldsymbol{\xi}}$. Finally, choose and fix a $y\in M$, set $\mathcal{U}(y)=O(M)\backslash\pi^{-1}\big(\mathrm{Cut}(y)\big)$, define $\mathfrak{e}\in\mathcal{U}(y)\longmapsto\boldsymbol{\theta}_{\mathfrak{e}}\in\mathbb{R}^d$ so that $\mathfrak{e}\boldsymbol{\theta}_{\mathfrak{e}}$ is the initial velocity of the unique minimal geodesic on $[0,1]$ running from $\pi(\mathfrak{e})$ to y, and determine $\boldsymbol{\ell}_{\mathfrak{e}}\in\mathbf{H}$ so that $\boldsymbol{\ell}_{\mathfrak{e}}(t)=t\boldsymbol{\theta}_{\mathfrak{e}}$, $t\in[0,1]$. The following statement is proved by the argument given to derive Lemma 2.9 in [10].

Lemma 3.7. *Assume that* $g : O(M) \times \mathbf{H} \longrightarrow \mathbb{R}$ *is a bounded smooth function with the property that*

$$\big[\mathcal{D}^0_{\mathbf{h}} g\big]\big(\mathfrak{e}, \boldsymbol{\ell}_{\mathfrak{e}}\big) = 0 \quad \textit{for all } \mathfrak{e} \in \mathcal{U}(y) \textit{ and } \mathbf{h} \in \mathbf{H}_0 \equiv \big\{\mathbf{h} \in \mathbf{H} : \mathbf{h}(1) = 0\big\}. \tag{3.8}$$

Then, for each compact subset $\mathcal{K}$ *of* $\mathcal{U}(y)$*, there is an* $\epsilon > 0$ *such that* (3.5) *and* (3.6) *hold whenever* $\mathfrak{e} \in \mathcal{K}$, $\mathbf{h}_{\mathfrak{e}} = \boldsymbol{\ell}_{\mathfrak{e}}$, *and* $f_0 = \alpha g$ *for some* $|\alpha| \le \epsilon$.

3.3. The crucial fact. Our application of the considerations in §§3.1–3.2 will be to functions of the form (cf. (2.9))

$$f_T(\mathfrak{e}, \mathbf{w}) = \rho\big(\mathfrak{e}, B_{\boldsymbol{\Xi}}(\mathfrak{e}, T, \mathbf{w})\big) \tag{3.9}$$

where $\rho \in C^\infty_{\mathrm{b}}\big(O(M) \times \mathbb{R}; \mathbb{R}\big)$ and we take

$$B_{\boldsymbol{\Xi}}(\mathfrak{e}, 0, \mathbf{g}) = \begin{cases} \displaystyle\int_0^1 \big(\dot{\boldsymbol{\xi}}^1_{\mathfrak{e}}(t), \dot{\mathbf{g}}(t)\big)_{\mathbb{R}^d}\, dt & \text{when } n = 1, \\ \big[\mathcal{D}^0_{\boldsymbol{\xi}^n} B_{\boldsymbol{\xi}^1 \otimes \cdots \otimes \boldsymbol{\xi}^{n-1}}(\mathfrak{e}, T, \,\cdot\,)\big](\mathbf{g}) & \text{when } n \ge 2. \end{cases} \tag{3.10}$$

Because the functions $B_{\boldsymbol{\Xi}}(\mathfrak{e}, T, \mathbf{w})$ are built out of solutions to stochastic integral equations with smooth coefficients, they satisfy all the technical properties required by [7]. In particular, in order to know that the asymptotic result described in §3.1 applies, Lemma 3.7 tells us that we need only check that we are dealing with a choice of $\boldsymbol{\Xi}$ such that

$$\big[\mathcal{D}^0_{\mathbf{h}} B_{\boldsymbol{\Xi}}\big](0, \mathfrak{e}, \boldsymbol{\ell}_{\mathfrak{e}}) = 0 \quad \text{for all } \mathbf{h} \in \mathbf{H}_0. \tag{3.11}$$

Thus, we must seek such a choice.

To begin, notice that, for any $\mathfrak{e} \in O(M) \longmapsto \boldsymbol{\Xi}_{\mathfrak{e}} = \boldsymbol{\xi}^1_{\mathfrak{e}} \otimes \cdots \otimes \boldsymbol{\xi}^n_{\mathfrak{e}}$,

$$B_{\boldsymbol{\Xi}}(\mathfrak{e}, 0, \mathbf{g}) = \frac{1}{2} \frac{\partial^n}{\partial s_n \cdots \partial s_1} \Big\| \big[\Theta_{\mathfrak{e},0,\boldsymbol{\Xi}}(\,\cdot\,, \mathbf{g})\big](\mathbf{s}) \Big\|^2_{\mathbf{H}} \Big|_{\mathbf{s}=\mathbf{0}}, \tag{3.12}$$

where $\mathbf{s} = (s^1, \dots, s^n)$; and we use induction on $i = 1, \dots, n$ to set $\mathbf{s}_i = (s_{n+1-i}, \dots, s_n)$, $\boldsymbol{\Xi}_i = \boldsymbol{\xi}^{n+1-i} \otimes \cdots \otimes \boldsymbol{\xi}^n$, and define

$$\begin{aligned} \big[\Theta_{\mathfrak{e},0,\boldsymbol{\Xi}_1}(t, \mathbf{g})\big](\mathbf{s}_1) &= \big[\Theta_{\mathfrak{e},0,\boldsymbol{\xi}^n}(t, \mathbf{g})\big](\mathbf{s}_1), \\ \big[\Theta_{\mathfrak{e},0,\boldsymbol{\xi}^{n-i} \otimes \boldsymbol{\Xi}_i}(t, \mathbf{g})\big](\mathbf{s}_{i+1}) &= \big[\Theta_{\mathfrak{e},0,\boldsymbol{\xi}^{n-i}}\big(t, \big[\Theta_{\mathfrak{e},0,\boldsymbol{\Xi}_i}(\,\cdot\,, \mathbf{g})\big](\mathbf{s}_i)\big)\big](s_{n-i}). \end{aligned}$$

Thus,

$$\big[\mathcal{D}^0_{\mathbf{h}} B_{\boldsymbol{\Xi}}\big](0, \mathfrak{e}, \boldsymbol{\ell}_{\mathfrak{e}}) = \frac{1}{2} \frac{\partial^{n+1}}{\partial \sigma \partial s_n \cdots \partial s_1} \Big\| \big[\Theta_{\mathfrak{e},0,\boldsymbol{\Xi} \otimes \mathbf{h}}(\,\cdot\,, \boldsymbol{\ell}_{\mathfrak{e}})\big](\mathbf{s}, \sigma) \Big\|^2_{\mathbf{H}} \Big|_{(\mathbf{s},\sigma)=(\mathbf{0},0)}.$$

In particular, (3.11) will hold if, for sufficiently small $|\mathbf{s}|$,

$$\frac{\partial}{\partial \sigma} \Big\| \big[\Theta_{\mathfrak{e},0,\boldsymbol{\Xi} \otimes \mathbf{h}}(\,\cdot\,, \boldsymbol{\ell}_{\mathfrak{e}})\big](\mathbf{s}, \sigma) \Big\|^2_{\mathbf{H}} \Big|_{\sigma=0} = 0 \quad \text{for all } \mathbf{h} \in \mathbf{H}_0.$$

Given $n \ge 1$ and $\mathbf{V} = \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n \in (\mathbb{R}^d)^{\otimes n}$, we want to make a special choice of $\boldsymbol{\Xi} = \boldsymbol{\xi}^1 \otimes \cdots \otimes \boldsymbol{\xi}^n$ such that $\boldsymbol{\Xi}_{\mathfrak{e}}(0) \equiv \mathbf{V}$. Namely, for each $\mathbf{v} \in \mathbb{R}^d$, define

$\mathfrak{e}\in\mathcal{U}(y)\longmapsto\mathbf{j}^{\mathbf{v}}_{\mathfrak{e}}\in\mathbb{R}^d\oplus\mathbf{H}$ so that (cf. (3.2) and the discussion preceding Lemma 3.7)[5]

$$\mathbf{j}^{\mathbf{v}}_{\mathfrak{e}}(t)=\omega\left(\frac{d}{ds}\mathfrak{F}_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s),0}\big(t,\boldsymbol{\ell}_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s)}\big)\Big|_{s=0}\right),$$

where

$$\mathfrak{f}'_{\mathfrak{e},\mathbf{v}}(s)=\mathfrak{E}(\mathbf{v})_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s)}\ \text{with}\ \mathfrak{f}_{\mathfrak{e},\mathbf{v}}(0)=\mathfrak{e}.$$

Notice that, by uniqueness, for any $\mathfrak{e}\in\mathcal{U}(y)$ and $\mathbf{v}\in\mathbb{R}^d$,

$$\mathfrak{f}_{\mathfrak{e},\mathbf{v}}\restriction[0,s]\subseteq\mathcal{U}(y)\implies\big[\mathfrak{F}_{\mathfrak{e},0,\mathbf{j}^{\mathbf{v}}}(t,\boldsymbol{\ell}_{\mathfrak{e}})\big](s)=\mathfrak{F}_{\mathfrak{e},0}\big(t,\boldsymbol{\ell}_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s)}\big),$$

and therefore

$$\mathfrak{f}_{\mathfrak{e},\mathbf{v}}\restriction[0,s]\subseteq\mathcal{U}(y)\implies\big[\Theta_{\mathfrak{e},0,\mathbf{j}^{\mathbf{v}}}(t,\boldsymbol{\ell}_{\mathfrak{e}})\big](s)=\boldsymbol{\ell}_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}}(s)}(t).$$

Next, for $n\geq 1$ and $\mathbf{V}=\mathbf{v}^1\otimes\cdots\otimes\mathbf{v}^n\in(\mathbb{R}^d)^{\otimes n}$, define $\mathfrak{e}\in\mathcal{U}(y)\longmapsto\mathbf{J}^{\mathbf{V}}_{\mathfrak{e}}\in C^\infty\big([0,1];(\mathbb{R}^d)^{\otimes n}\big)$ so that

$$\mathbf{J}^{\mathbf{V}}_{\mathfrak{e}}(t)=\mathbf{j}^{\mathbf{v}^1}_{\mathfrak{e}}(t)\otimes\cdots\otimes\mathbf{j}^{\mathbf{v}^n}_{\mathfrak{e}}(t);$$

and, for $\mathbf{v}^{n+1}\in\mathbb{R}^d$, use induction to define $\mathfrak{e}\in O(M)\longmapsto\mathfrak{f}_{\mathfrak{e},\mathbf{V}\otimes\mathbf{v}^{n+1}}\in C^\infty\big(\mathbb{R}^{n+1};O(M)\big)$, so that

$$\mathfrak{f}_{\mathfrak{e},\mathbf{V}\otimes\mathbf{v}^{n+1}}\big(\mathbf{s},s^{n+1}\big)=\mathfrak{f}_{\mathfrak{f}_{\mathfrak{e},\mathbf{V}(\mathbf{s})},\mathbf{v}^{n+1}}(s^{n+1})\quad\text{for }(\mathbf{s},s_{n+1})\in\mathbb{R}^n\times\mathbb{R}.$$

Clearly, for each $\mathbf{V}$, there is an upper semi-continuous $\delta_{\mathbf{V}}:\mathcal{U}(y)\longrightarrow(0,\infty)$ such that

$$\mathfrak{e}\in\mathcal{U}(y)\text{ and }|\mathbf{s}|\leq\delta_{\mathbf{V}}(\mathfrak{e})\implies\mathfrak{f}_{\mathfrak{e},\mathbf{V}}(\mathbf{s})\in\mathcal{U}(y).$$

Hence, by induction and uniqueness,

$$\begin{aligned}&\mathfrak{e}\in\mathcal{U}(y)\text{ and }|\mathbf{s}|\leq\delta_{\mathbf{V}}(\mathfrak{e})\implies\\&\qquad\big[\Theta_{\mathfrak{e},0,\mathbf{J}^{\mathbf{V}}}(\,\cdot\,,\boldsymbol{\ell}_{\mathfrak{e}})\big](\mathbf{s})=\boldsymbol{\ell}_{\mathfrak{f}_{\mathfrak{e},\mathbf{v}^n\otimes\cdots\otimes\mathbf{v}^1}(\mathbf{s})}.\end{aligned}\tag{3.13}$$

But, because $\pi\circ\mathfrak{F}_{\mathfrak{e},0}(\,\cdot\,,\boldsymbol{\ell}_{\mathfrak{e}})$ is a minimal geodesic,

$$\frac{d}{d\sigma}\Big\|\big[\Theta_{\mathfrak{e},0,\mathbf{h}}(\,\cdot\,,\boldsymbol{\ell}_{\mathfrak{e}})\big](\sigma)\Big\|^2_{\mathbf{H}}\bigg|_{\sigma=0}=0\quad\text{for any }\mathfrak{e}\in\mathcal{U}(y)\text{ and }\mathbf{h}\in\mathbf{H}_0;$$

and so we have now proved that

$$\begin{aligned}&(\mathfrak{e},\mathbf{h})\in\mathcal{U}(y)\times\mathbf{H}_0\text{ and }|\mathbf{s}|\leq\delta_{\mathbf{V}}(\mathfrak{e})\implies\\&\qquad\frac{\partial}{\partial\sigma}\Big\|\big[\Theta_{\mathfrak{e},0,\mathbf{J}^{\mathbf{V}}\otimes\mathbf{h}}(\,\cdot\,,\boldsymbol{\ell}_{\mathfrak{e}})\big](\mathbf{s},\sigma)\Big\|^2_{\mathbf{H}}\bigg|_{\sigma=0}=0\end{aligned}\tag{3.14}$$

for all $\mathbf{V}\in(\mathbb{R}^d)^{\otimes n}$.

After combining Lemma 3.7 and (3.14) with the results from [7] alluded to in §3.1, we see that, for any $\mathbf{V}\in(\mathbb{R}^d)^{\otimes n}$, $\mathcal{K}\subset\subset\mathcal{U}(y)$, and $\rho\in C^\infty_{\mathrm{b}}(\mathbb{R};\mathbb{R})$, there exists a

---

[5] The notation $\mathbf{j}^{\mathbf{v}}$ here is recognition of the fact that $t\rightsquigarrow\mathfrak{F}_{\mathfrak{e},0}(t,\boldsymbol{\ell}_{\mathfrak{e}})\mathbf{j}^{\mathbf{v}}_{\mathfrak{e}}(t)$ is a Jacobi vector field along the geodesic $\pi\circ\mathfrak{F}_{\mathfrak{e},0}(\,\cdot\,,\boldsymbol{\ell}_{\mathfrak{e}})$.

$\delta_{\mathbf{V}}(\rho,\mathcal{K})>0$ such that, for each $|\alpha|\le\delta_{\mathbf{V}}(\rho,\mathcal{K})$,

$$\exp\left(-\frac{\alpha\rho(0)+\frac{1}{2}\mathrm{dist}(\pi(\mathfrak{e}),y)^2}{T}\right)$$
$$\times\,\mathbb{E}^{\mu_T}\left[\exp\left(\frac{\alpha\rho\big(B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})\big)}{T}\right)\delta_y\circ\pi\big(\mathfrak{F}_{\mathfrak{e},T}(1,\mathbf{w})\big)\right]$$

is asymptotic, as $T\searrow 0$, uniformly in $\mathfrak{e}\in\mathcal{K}$, to an expression which is independent of α. In particular, by taking the ratio of the preceding with the quantity which results from replacing α by 0, one obtains (cf. (2.3))

$$\lim_{T\searrow 0}\sup_{\mathfrak{e}\in\mathcal{K}}\left|\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\left[\exp\left(\frac{\alpha\rho\big(B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})\big)-\alpha\rho(0)}{T}\right)\right]-1\right|=0. \tag{3.15}$$

Lemma 3.16. *For each $K\subset\subset M\setminus\mathrm{Cut}(y)$, $\mathbf{V}\in(\mathbb{R}^d)^{\otimes n}$, and $p\in[1,\infty)$,*

$$\sup_{T\in(0,1]}\max_{\mathfrak{e}\in\pi^{-1}(K)}T^{-1}\big\|B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})\big\|_{L^p(\mu_{T,\mathfrak{e}|y})}<\infty.$$

Proof. Choose $\rho\in C^\infty_{\mathrm{c}}\big((-2,2);[0,1]\big)$ so that $\rho\equiv 1$ on $[-1,1]$. Then, by (3.15), there exists an $\alpha>0$ such that

$$\sup_{T\in(0,1]}\max_{\mathfrak{e}\in\pi^{-1}(K)}e^{\frac{\alpha}{T}}\mu_{T,\mathfrak{e}|y}\Big(\Big\{\mathbf{w}:\ \big|B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})\big|\ge 2\Big\}\Big)$$
$$\le\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\left[\exp\left(\frac{-\alpha\rho\big(B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})\big)+\alpha}{T}\right)\right]<\infty.$$

Hence, in view of (2.11), it suffices for us to check that for some $\alpha>0$,

$$\sup_{\substack{T\in(0,1]\\ \mathfrak{e}\in\pi^{-1}(K)}}\int_{\Gamma(\mathbf{V},\mathfrak{e},T)}\exp\!\left(\frac{\alpha\big|B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})\big|}{T}\right)\mu_{T,\mathfrak{e}|y}(d\mathbf{w})<\infty,$$

where

$$\Gamma(\mathbf{V},\mathfrak{e},T)\equiv\big\{\mathbf{w}:|B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})|\le 2\big\}$$

But, by taking

$$\rho_\pm(t)=t\rho\left(\tfrac{t}{2}\right),$$

we see that (3.15) with $\rho_\pm$ replacing ρ, provides the required estimate. $\square$

4. The main results.

4.1. Outside the cut-locus. We are now ready to prove the main result of this article, which says that when $x\equiv\pi(\mathfrak{e})$ is outside the cut locus of y, derivatives of all orders commute with the limit in (1.1). In fact, a slightly stronger asymptotic statement can be made, and as a consequence, we improve the upper bound of [16] for compact sets outside the cut locus.

Theorem 4.1. *Let $y\in M$ and $K\subset\subset M\setminus\mathrm{Cut}(y)$ be given. Then, for each $n\ge 1$, and $\mathbf{v}^1,\dots,\mathbf{v}^n\in\mathbb{R}^d$:*

$$\sup_{\substack{T\in(0,1]\\ \mathfrak{e}\in\pi^{-1}(K)}}T^{-1}\left|T\nabla_{\mathfrak{e}\mathbf{v}^n}\cdots\nabla_{\mathfrak{e}\mathbf{v}^1}\log p_T(\,\cdot\,,y)+\tfrac{1}{2}\nabla_{\mathfrak{e}\mathbf{v}^n}\cdots\nabla_{\mathfrak{e}\mathbf{v}^1}\mathrm{dist}(\,\cdot\,,y)^2\right|<\infty.$$

In particular, there is a $C_n(K,y)<\infty$ *such that for* $\mathbf{V}=\mathbf{v}^1\otimes\cdots\otimes\mathbf{v}^n$

$$\Big|\nabla_{\mathfrak{e}\mathbf{v}^n}\cdots\nabla_{\mathfrak{e}\mathbf{v}^1}\log p_T(\,\cdot\,,y)\Big|\le\frac{C_n(K,y)}{T}\,\|\mathbf{V}\|_{(\mathbb{R}^d)^{\otimes n}}\quad \textit{for all } (T,\mathfrak{e})\in(0,1]\times\pi^{-1}(K).$$

Proof. Choose $\mathbf{v}^1,\dots,\mathbf{v}^n\in\mathbb{R}^d$, and apply (2.12) with $\boldsymbol{\xi}^k\restriction\pi^{-1}(K)=\mathbf{j}^{\mathbf{v}^k}\restriction\pi^{-1}(K)$ to see that, for all $\mathfrak{e}\in\pi^{-1}(K)$ (cf. §2.1 for the notation),

$$\begin{aligned}&\mathfrak{E}(\mathbf{V})_{\mathfrak{e}}\log p_T(\pi(\,\cdot\,),y)\\&\quad=\sum_{m=1}^{n}(-T)^{-m}\sum_{\{\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m\}\in\Sigma_{m,n}}\mathcal{C}_m^{\mu_{T,\mathfrak{e}|y}}\big(B_{\mathbf{J}\mathbf{V}^{\boldsymbol{\beta}^1}}(\mathfrak{e},T,\mathbf{w}),\dots,B_{\mathbf{J}\mathbf{V}^{\boldsymbol{\beta}^m}}(\mathfrak{e},T,\mathbf{w})\big).\end{aligned}$$

Note that, because

$$\mathcal{C}_1^{\mu_{T,\mathfrak{e}|y}}\big(B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})\big)=\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\big[B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})\big],$$

Lemma 3.16 guarantees that

$$\sup_{T\in(0,1]}\max_{\mathfrak{e}\in\pi^{-1}(K)}\left|\frac{\mathcal{C}_1^{\mu_{T,\mathfrak{e}|y}}\big(B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},T,\mathbf{w})\big)-B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})}{T}\right|<\infty.$$

At the same time, by (3.12) and (3.13), we know that

$$B_{\mathbf{J}\mathbf{V}}(\mathfrak{e},0,\boldsymbol{\ell}_{\mathfrak{e}})=\frac{1}{2}\mathfrak{E}(\mathbf{V})_{\mathfrak{e}}\,\mathrm{dist}\big(\pi(\,\cdot\,),y\big)^2.$$

Thus, all that remains is to check that, when $2\le m\le n$,

$$\sup_{T\in(0,1]}\sup_{\mathfrak{e}\in\pi^{-1}(K)}T^{-m}\mathcal{C}_m^{\mu_{T,\mathfrak{e}|y}}\Big(B_{\mathbf{J}\mathbf{V}^{\boldsymbol{\beta}^1}}(\mathfrak{e},T,\mathbf{w})\big),\dots,B_{\mathbf{J}\mathbf{V}^{\boldsymbol{\beta}^m}}(\mathfrak{e},T,\mathbf{w})\Big)<\infty,$$

which follows from (A.3) and Lemma 3.16, with $p=m$. □

4.2. At the cut-locus. Again let $y\in M$ be fixed, and consider the behavior of the nth order covariant derivative of $\log p_T(\,\cdot\,,y)$ as $T\searrow 0$. When $x\equiv\pi(\mathfrak{e})\in\mathrm{Cut}(y)$, the situation is not so clear. In order to describe what we can say in this situation, we define (cf. [8])

$$M(x,y)=\big\{X_x\in T_xM:\,y=\exp_x(X_x)\text{ and }\mathrm{dist}(x,y)=|X_x|_{T_xM}\big\}$$

and

$$\widehat{M(x,y)}=\left\{(X_x,W_x)\in M(x,y)\times\big(T_xM\setminus\{\mathbf{0}\}\big):\,\frac{d}{ds}\exp_x(X_x+sW_x)\Big|_{s=0}=\mathbf{0}\right\}.$$

Because, $x\in\mathrm{Cut}(y)$, either $M(x,y)$ contains more than one element or $\widehat{M(x,y)}\ne\emptyset$. Finally, for $\mathfrak{e}\in\pi^{-1}(x)$, set

$$M(\mathfrak{e},y)=\big\{\boldsymbol{\theta}\in\mathbb{R}^d:\,\mathfrak{e}\boldsymbol{\theta}\in M(x,y)\big\},$$

and, for $\boldsymbol{\theta}\in M(\mathfrak{e},y)$ define a path $\ell(\boldsymbol{\theta})\in\mathbf{H}(\mathfrak{e},y)$ by $\big[\ell(\boldsymbol{\theta})\big](t)=t\boldsymbol{\theta},\ t\in[0,1]$.

THEOREM 4.2. *Assume that* $M(x,y)$ *contains more than one element, and let* $\widetilde{M(x,y)}\supseteq M(x,y)$ *be a submanifold of* T_xM *with the property that*

$$(X_x,W_x)\in\widehat{M(x,y)}\implies W_x\not\perp T_{X_x}\widetilde{M(x,y)}.$$

Further, assume that $M(x,y)$ *has positive measure when* $\widetilde{M(x,y)}$ *is given the measure determined by the Riemannian structure which it inherits as a submanifold. Given* $\mathfrak{e}\in\pi^{-1}(x)$*, there exists a non-degenerate (i.e., not concentrated at a single point) Borel probability measure* $\lambda_{\mathfrak{e}|y}$ *on* $\mathbb{R}^d$ *which is supported on* $M(\mathfrak{e},y)$ *and for which the asymptotic series*

$$\begin{aligned}&\nabla_{\mathfrak{e}\mathbf{v}^n}\cdots\nabla_{\mathfrak{e}\mathbf{v}^1}\log p_T(\,\cdot\,,y)\\ &\sim\sum_{m=1}^{n}(-T)^{-m}\sum_{\{\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m\}\in\Sigma_{m,n}}\mathcal{C}_m^{\lambda_{\mathfrak{e}|y}}\big(B_{\boldsymbol{\Xi}^{\boldsymbol{\beta}^1}}(\mathfrak{e},0,\ell(\boldsymbol{\theta})),\dots,B_{\boldsymbol{\Xi}^{\boldsymbol{\beta}^m}}(\mathfrak{e},0,\ell(\boldsymbol{\theta}))\big)\end{aligned}\tag{4.3}$$

holds whenever $n\ge 1$, $\mathbf{v}^i\in\mathbb{R}^d$, *and* $\boldsymbol{\xi}^i(t)=(1-t)\mathbf{v}^i$, *for* $i=1,\dots,n$. *In particular, one can choose vectors* $\mathbf{v}^i$ *so that there exist infinitely many* $n\ge 1$ *with*

$$\lim_{T\searrow 0}T^n\nabla_{\mathfrak{e}\mathbf{v}^n}\cdots\nabla_{\mathfrak{e}\mathbf{v}^1}\log p_T(\,\cdot\,,y)\neq 0.\tag{4.4}$$

Proof. Given (2.12), the proof is essentially the same as that of Theorem 2.34 in [10]. In particular, if $\mathbf{v}\in\mathbb{R}^d$ and $\boldsymbol{\xi}(t)=(1-t)\mathbf{v}$, then (cf. (2.9))

$$-B_{\boldsymbol{\xi}}(\mathfrak{e},0,\ell(\boldsymbol{\theta}))=I_{\mathbf{v}}(\boldsymbol{\theta})\equiv(\mathbf{v},\boldsymbol{\theta})_{\mathbb{R}^d}.$$

The only additional ingredient is provided by Lemma A.5, which guarantees that, when $\mathbf{v}\neq\mathbf{0}$ and $\mathbf{v}\not\perp M(\mathfrak{e},y)$, $\mathcal{C}_m^{\lambda_{\mathfrak{e}|y}}\big(I_{\mathbf{v}}(\boldsymbol{\theta}),\dots,I_{\mathbf{v}}(\boldsymbol{\theta})\big)\neq 0$ for infinitely many $m\ge 1$. □

4.3. Final comment. It may be of some interest to see what Theorem 4.1 says about derivatives of $p_T(\,\cdot\,,y)$ itself. For this purpose, let $\mathbf{v}^1,\dots,\mathbf{v}^n\in\mathbb{R}^d$ be given, set $\mathbf{V}=\mathbf{v}^1\otimes\cdots\otimes\mathbf{v}^n$, and note that, by elementary calculus and an easy induction argument:

$$\frac{\mathfrak{C}(\mathbf{V})_{\mathfrak{e}}p_T(\pi(\,\cdot\,),y)}{p_T(\pi(\mathfrak{e}),y)}=\sum_{m=1}^{n}\sum_{\{\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m\}\in\Sigma_{m,n}}\prod_{\ell=1}^{m}\mathfrak{C}(\mathbf{V}^{\boldsymbol{\beta}^\ell})_{\mathfrak{e}}\log p_T\big(\pi(\,\cdot\,),y\big).\tag{4.5}$$

Obviously, if no restrictions are placed on the $\mathbf{v}^k$'s, Theorem 4.1 gives nothing more than

$$\left|\frac{\mathfrak{C}(\mathbf{V})_{\mathfrak{e}}p_T\big(\pi(\,\cdot\,),y\big)}{p_T\big(\pi(\mathfrak{e}),y\big)}\right|\le C\,\|\mathbf{V}\|_{(\mathbb{R}^d)^{\otimes n}}\,T^{-n}\quad\text{when }T\in(0,1],$$

an estimate which, in fact (cf. [16]), holds everywhere, not just outside the cut locus. Our goal here is to show that one can say more when one puts restrictions on the $\mathbf{v}^k$'s and stays away from the cut locus.

Set

$$E(\mathfrak{e},y)=\frac{1}{2}\mathrm{dist}\big(\pi(\mathfrak{e}),y\big)^2\quad\text{and}\quad\big[E(\mathfrak{e},y)\big](\mathbf{V})=\mathfrak{C}(\mathbf{V})_{\mathfrak{e}}E(\,\cdot\,,y).$$

The first part of Theorem 4.1, in conjunction with (4.5), says that, for each $\mathfrak{e}\notin\pi^{-1}\big(\mathrm{Cut}(y)\big)$,

$$\left|\frac{\mathfrak{C}(\mathbf{V})_{\mathfrak{e}}p_T\big(\pi(\,\cdot\,),y\big)}{p_T\big(\pi(\mathfrak{e}),y\big)}\right|\le C(\mathbf{V},\mathfrak{e},y)T^{-m(\mathbf{V},\mathfrak{e},y)}\quad\text{when }T\in(0,1],\tag{4.6}$$

where

$$
m(\mathbf{V},\mathfrak{e},y) \equiv \max \left\{ k : \begin{array}{c} \exists\, 1 \le m \le n,\ \exists\, \{\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m\} \in \Sigma_{m,n}, \text{ and} \\ \exists\, 1 \le \ell_1 < \cdots < \ell_k \le m \ \text{ with} \\ \prod_{j=1}^{k} \big[E(\mathfrak{e},y)\big]\big(\mathbf{V}^{\boldsymbol{\beta}^{\ell_j}}\big) \ne 0 \end{array} \right\}. \tag{4.7}
$$

In order to take advantage of these observations, suppose that $\pi(\mathfrak{e}) \notin \mathrm{Cut}(y)$. Then, by standard (cf. Chapter 9 in [4]) calculations:

$$
\begin{gathered} \pi(\mathfrak{e}) = y \implies \\ \big[E(\mathfrak{e},y)\big](\boldsymbol{\xi}) = 0 \text{ and } \big[E(\mathfrak{e},y)\big](\boldsymbol{\xi}\otimes\boldsymbol{\eta}) = \big(\boldsymbol{\xi},\boldsymbol{\eta}\big)_{\mathbb{R}^d} \quad \text{for all } \boldsymbol{\xi},\boldsymbol{\eta} \in \mathbb{R}^d. \end{gathered} \tag{4.8}
$$

On the other hand, when $\pi(\mathfrak{e}) \ne y$, we will make a special choice of basis. Namely, we take $\mathbf{e}^1(\mathfrak{e},y) \in \mathbb{R}^d$ so that $\mathrm{dist}\big(\pi(\mathfrak{e}),y\big)\,\mathfrak{e}\mathbf{e}^1(\mathfrak{e},y) \in T_{\pi(\mathfrak{e})}M$ is the initial velocity of the unique minimal geodesic from $\pi(\mathfrak{e})$ to y. Then

$$
\frac{\big[E(\mathfrak{e},y)\big](\boldsymbol{\xi})}{\mathrm{dist}\big(\pi(\mathfrak{e}),y\big)} = -\big(\boldsymbol{\xi},\mathbf{e}^1(\mathfrak{e},y)\big)_{\mathbb{R}^d} = \big[E(\mathfrak{e},y)\big]\big(\boldsymbol{\xi}\otimes\mathbf{e}^1(\mathfrak{e},y)\big) \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^d.
$$

Hence, because $\big[E(\mathfrak{e},y)\big](\boldsymbol{\xi}\otimes\boldsymbol{\eta}) = \big[E(\mathfrak{e},y)\big](\boldsymbol{\eta}\otimes\boldsymbol{\xi})$, we can choose $\mathbf{e}^2(\mathfrak{e},y),\dots,\mathbf{e}^d(\mathfrak{e},y) \in \mathbb{R}^d$ so that $\big(\mathbf{e}^1(\mathfrak{e},y),\dots,\mathbf{e}^d(\mathfrak{e},y)\big)$ forms an orthonormal basis and

$$
\begin{aligned} \big[E(\mathfrak{e},y)\big]\big(\mathbf{e}^i(\mathfrak{e},y)\big) &= 0 \quad \text{for } 2 \le i \le d, \\ \big[E(\mathfrak{e},y)\big]\big(\mathbf{e}^i(\mathfrak{e},y)\otimes\mathbf{e}^j(\mathfrak{e},y)\big) &= 0 \quad \text{for } 1 \le i \ne j \le d. \end{aligned} \tag{4.9}
$$

Theorem 4.10. *Let $y \in M$ and $\mathfrak{e} \notin \pi^{-1}\big(\mathrm{Cut}(y)\big)$ be given. If $\pi(\mathfrak{e}) \ne y$, choose the o.n. basis $\big(\mathbf{e}^1(\mathfrak{e},y),\dots,\mathbf{e}^d(\mathfrak{e},y)\big)$ as above; and when $\pi(\mathfrak{e}) = y$, let $\big(\mathbf{e}^1(\mathfrak{e},y),\dots,\mathbf{e}^d(\mathfrak{e},y)\big)$ be an arbitrary o.n. basis in $\mathbb{R}^d$. Next, suppose that $\mathbf{V} = \mathbf{e}^{i_1}(\mathfrak{e},y)\otimes\cdots\otimes\mathbf{e}^{i_n}(\mathfrak{e},y)$, define*

$$
N_k(\mathbf{V},\mathfrak{e},y) = \mathrm{card}\{1 \le m \le n : i_m = k\},
$$

and set[6]

$$
m_1(\mathbf{V},\mathfrak{e},y) = 0 \quad \textit{and} \quad m_2(\mathbf{V},\mathfrak{e},y) = \sum_{k=1}^{d} \left[\frac{N_k(\mathbf{V},\mathfrak{e},y)}{2}\right] \quad \textit{if } \pi(\mathfrak{e}) = y,
$$

$$
m_1(\mathbf{V},\mathfrak{e},y) = N_1(\mathbf{V},\mathfrak{e},y) \quad \textit{and} \quad m_2(\mathbf{V},\mathfrak{e},y) = \sum_{k=2}^{d} \left[\frac{N_k(\mathbf{V},\mathfrak{e},y)}{2}\right] \quad \textit{if } \pi(\mathfrak{e}) \ne y.
$$

Then

$$
\left| \frac{\nabla_{\mathfrak{e}\mathbf{e}^{i_n}} \cdots \nabla_{\mathfrak{e}\mathbf{e}^{i_1}} p_T(\,\cdot\,,y)}{p_T\big(\pi(\mathfrak{e}),y\big)} \right| \le C(\mathbf{V},\mathfrak{e},y)T^{-m(\mathbf{V},\mathfrak{e},y)} \quad \textit{when } T \in (0,1],
$$

where

$$
m(\mathbf{V},\mathfrak{e},y) \le \left[m_1 + m_2 + \left[\frac{n - m_1 - 2m_2}{3}\right] \right](\mathbf{V},\mathfrak{e},y).
$$

---

[6] We use $[t]$ to denote the integer part of $t \in \mathbb{R}$.

In particular, if either $\pi(\mathfrak{e}) = y$ *or* $\pi(\mathfrak{e}) \neq y$ *and* $i_m \neq 1$ *for any* $1 \le m \le n$, *then* $m(\mathbf{V}, \mathfrak{e}, y) \le \left[\frac{n}{2}\right]$.

Proof. In view of (4.6) and our notation convention of §2.1, all that we have to do is check that the $m(\mathbf{V}, \mathfrak{e}, y)$ in (4.7) satisfies the asserted estimate. But, for any $1 \le m \le n$ and $\{\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^m\} \in \Sigma_{m,n}$, (4.7) and (4.8) say that, for any $1 \le \ell \le m$, $\big[E(\mathfrak{e}, y)\big]\big(V^{\boldsymbol{\beta}^\ell}\big) \neq 0$ only if

$$\boldsymbol{\beta}^\ell \in \{(k,k) : 1 \le k \le d\} \text{ or } \left|\boldsymbol{\beta}^\ell\right| \ge 3 \quad \text{when } \pi(\mathfrak{e}) = y,$$
$$\boldsymbol{\beta}^\ell = (1),\ \boldsymbol{\beta}^\ell \in \{(k,k) : 2 \le k \le d\}, \text{ or } \left|\boldsymbol{\beta}^\ell\right| \ge 3 \quad \text{when } \pi(\mathfrak{e}) \neq y.$$

Thus, a partition which achieves the maximum in (4.7) can be chosen from among partitions $\{\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^m\}$ in which

$$\operatorname{card}\big\{\ell : \boldsymbol{\beta}^\ell \in \{(k,k) : 1 \le k \le d\}\big\} = [m_1 + m_2](\mathbf{V}, \mathfrak{e}, y) \quad \text{if } \pi(\mathfrak{e}) = y,$$
$$\operatorname{card}\big\{\ell : \boldsymbol{\beta}^\ell = (1) \text{ or } \boldsymbol{\beta}^\ell \in \{(k,k) : 2 \le k \le d\}\big\} = [m_1 + m_2](\mathbf{V}, \mathfrak{e}, y) \quad \text{if } \pi(\mathfrak{e}) \neq y.$$

Further, among these, the maximum cannot be larger than $m_1 + m_2$ plus the largest number of remaining blocks which have three or more elements. □

Appendix.

A.1. Cumulants.

The purpose of this section is to provide a few elementary facts about cumulants. For more details, see either [11] or [12].

The easiest way to think about cumulants is to consider bounded random variables $X_1, \dots, X_m$ on some probability space $(E, \mathcal{F}, \mu)$ and take the *mth cumulant* to be

$$\text{(A.1)} \qquad \mathcal{C}^\mu_m(X_1, \dots, X_m) = \frac{\partial^m}{\partial\alpha_1 \cdots \partial\alpha_m} \log \mathbb{E}^\mu\left[\exp\left(\sum_{j=1}^m \alpha_j X_j\right)\right]\Bigg|_{\boldsymbol{\alpha}=\mathbf{0}}.$$

However, because we need to consider unbounded random variables, we will need an alternative description of cumulants. Namely (cf. the discussion in §2.3 preceding (2.8)),

$$\text{(A.2)} \qquad \mathcal{C}^\mu_m(X_1, \dots, X_m) = \sum_{k=1}^m (-1)^{k-1}(k-1)! \sum_{\{\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^k\} \in \Sigma_{k,m}} \prod_{j=1}^k \mathbb{E}^\mu\left[\prod_{i \in \boldsymbol{\beta}^j} X_i\right],$$

which is an expression that makes perfectly good sense as soon as each $X_k \in L^m(\mu)$ and agrees with the one in (A.1) when the X_k's are bounded. Moreover, by Hölder's inequality,

$$\left|\prod_{j=1}^k \mathbb{E}^\mu\left[\prod_{i \in \boldsymbol{\beta}^j} X_i\right]\right| \le \prod_{j=1}^k \prod_{i \in \boldsymbol{\beta}^j} \|X_i\|_{L^{|\boldsymbol{\beta}^j|}(\mu)} \le \prod_{i=1}^m \|X_i\|_{L^m(\mu)},$$

and so the multi-linear map

$$(X_1, \dots, X_m) \in L^m(\mu)^m \longmapsto \mathcal{C}^\mu_m\big(X_1, \dots, X_m\big) \in \mathbb{R}$$

is continuous. In particular, for each $m \ge 1$, there is a $c_m \in (0, \infty)$ such that

$$\text{(A.3)} \qquad \big|\mathcal{C}^\mu_m\big(X_1, \dots, X_m\big)\big| \le c_m \prod_{i=1}^m \|X_i\|_{L^m(\mu)}.$$

Finally, notice that, by (A.1), for each $m \geq 2$ and bounded X_i's,

$$\mathcal{C}^{\mu}_m\bigl(X_1,\dots,X_m\bigr)=0 \quad \text{if } X_i \text{ is constant for any } 1\leq i\leq m.$$

Thus, by continuity, multi-linearity, and (A.3), we know that

$$\bigl|\mathcal{C}^{\mu}_m\bigl(X_1,\dots,X_m\bigr)\bigr|\leq c_m\prod_{i=1}^m \|X_i-a_i\|_{L^m(\mu)} \tag{A.4}$$

for any $m\geq 2$ and $(a_1,\dots,a_m)\in\mathbb{R}^m$.

Besides the preceding, we made use in Theorem 4.2 of the fact contained in the following.

LEMMA A.5. *Let X be a bounded*[7]*, non-constant random variable on $(\Omega,\mathcal{F},\mu)$. If the probability measure μ is non-degenerate (i.e., not concentrated at a single point), then $\mathcal{C}^{\mu}_m(X,\dots,X)\neq 0$ for infinitely many $m\geq 1$.*

Proof. Because X is bounded,

$$z\in\mathbb{C}\longmapsto M(z)\equiv\mathbb{E}^{\mu}\Bigl[e^{zX}\Bigr]\in\mathbb{C}$$

is an entire function of exponential type (cf. [3]) which is 1 at the origin; and so, by (A.1), it suffices to show that M has at least one root. But if M never vanished, then the Weierstraß product formula would say that $M(z)=e^{\alpha z}$ for some $\alpha\in\mathbb{C}$, with the conclusion that $\mathbb{R}\ni\mathbb{E}^{\mu}[X]=\alpha$ and $\mathbb{E}^{\mu}[(X-\alpha)^2]=0$. That is, X would have to be constant. □

A.2. An L^p-estimate. Here we will verify the estimate in (2.11). Thus, let $n\geq 1$ and smooth $\mathfrak{e}\in O(M)\longmapsto \boldsymbol{\xi}^i_{\mathfrak{e}}\in C^2\bigl([0,1];\mathbb{R}^d\bigr), 1\leq i\leq n$, be given, set $\boldsymbol{\Xi}=\boldsymbol{\xi}^1\otimes\cdots\otimes\boldsymbol{\xi}^n$, and, for $T\in(0,1]$, define $B_{\boldsymbol{\Xi}}(\mathfrak{e},T,\mathbf{w})$ accordingly, as in (2.9). Arguing as in [16] (cf. Lemma 3.3 and Appendix there), one finds that

$$B_{\boldsymbol{\Xi}}(\mathfrak{e},T,\mathbf{w})=\int_0^1\bigl[\alpha_0(\mathfrak{e},T,\mathbf{w})\bigr](\tau)\,d\tau+\sum_{j=1}^d\int_0^1\bigl[\alpha_j(\mathfrak{e},T,\mathbf{w})\bigr](\tau)\,dw_j(\tau),$$

where, for each $(T,\mathfrak{e})\in(0,1]\times O(M)$ and $0\leq j\leq d$, $(\tau,\mathbf{w})\in[0,1]\times\mathfrak{W}\longmapsto\bigl[\alpha_j(\mathfrak{e},T,\mathbf{w})\bigr](\tau)\in\mathbb{R}$ is a $\{\mathcal{B}_t:\ t\in[0,1]\}$-progressively measurable function and there exists an $\epsilon>0$ such that, for all $R\geq 0$,

$$\sup_{(T,\mathfrak{e})\in[0,1]\times O(M)}\mu_T\left(\sup_{\tau\in[0,1]}\max_{1\leq j\leq d}\bigl|\bigl[\alpha_j(\mathfrak{e},T,\,\cdot\,)\bigr](\tau)\bigr|\geq R^{n-1}\right)\leq 2\exp\bigl(-\epsilon R^2\bigr). \tag{A.6}$$

Now write

$$\int_0^1\bigl[\alpha_0(\mathfrak{e},T,\mathbf{w})\bigr](\tau)\,d\tau=\sum_{\ell=1}^{\infty}2^{-\ell}\Delta_{0,\ell}(\mathfrak{e},T,\mathbf{w})$$
$$\text{where } \Delta_{0,\ell}(\mathfrak{e},T,\mathbf{w})\equiv 2^{\ell}\int_{1-2^{-\ell+1}}^{1-2^{-\ell}}\bigl[\alpha_0(\mathfrak{e},T,\mathbf{w})\bigr](\tau)\,d\tau$$

---

[7] In order to appreciate the importance of boundedness here, consider the case when X is Gaussian.

and, for $1 \le j \le d$,

$$\int_0^1 \big[\alpha_j(\mathfrak{e},T,\mathbf{w})\big](\tau)\,d\tau = \sum_{\ell=1}^{\infty} 2^{-\frac{\ell}{2}}\Delta_{j,\ell}(\mathfrak{e},T,\mathbf{w})$$

$$\text{where } \Delta_{j,\ell}(\mathfrak{e},T,\mathbf{w}) \equiv 2^{\frac{\ell}{2}} \int_{1-2^{-\ell+1}}^{1-2^{-\ell}} \big[\alpha_j(\mathfrak{e},T,\mathbf{w})\big](\tau)\,dw_j(\tau).$$

Obviously, all that we have to do is show that, for all $p \in [1,\infty)$,

$$\text{(A.7)} \qquad \sup_{(T,\mathfrak{e})\in(0,1]\times O(M)} \sup_{\ell\ge 1} \max_{0\le j\le d} T^n \big\|\Delta_\ell(\mathfrak{e},T,\,\cdot\,)\big\|_{L^p(\mu_{T,\mathfrak{e}|y})} < \infty.$$

The key to our proof of (A.7) is contained in the inequality (cf. (A.3) in [16]) which says that, for each $p \in [1,\infty)$, there is a universal $C_p < \infty$ such that

$$\text{(A.8)} \qquad \mathbb{E}^\nu\big[X^pY\big] \le C_p\left(\log\|Y\|_{L^\infty(\nu)} + \mathbb{E}^\nu\big[e^{X^2}\big]\right)^{\frac{p}{2}}$$

for any probability measure ν and any non-negative random variables X and Y such that $\mathbb{E}^\nu[Y] = 1$. In particular, since

$$\begin{aligned}&\mathbb{E}^{\mu_{T,\mathfrak{e}|y}}\Big[\big|\Delta_{j,\ell}(\mathfrak{e},T,\mathbf{w})\big|^{\frac{p}{n}}\Big]\\ &\qquad = \mathbb{E}^{\mu_T}\left[\big|\Delta_{j,\ell}(\mathfrak{e},T,\mathbf{w})\big|^{\frac{p}{n}} \frac{p_{2^{-\ell}T}\big(\pi\circ\mathfrak{F}_{\mathfrak{e},T}(1-2^{-\ell},\mathbf{w}),y\big)}{p_T\big(\pi(\mathfrak{e}),y\big)}\right],\end{aligned}$$

and, by standard heat kernel estimates (cf. [9]), there is a $C < \infty$ such that

$$\log\frac{\|p_{tT}(\,\cdot\,,y)\|_{C(M;\mathbb{R})}}{p_T(x,y)} \le C\left(-\log t + \frac{1}{T}\right) \quad \text{for all } (t,x,y) \in (0,1)\times M\times M,$$

(A.8), with pn replacing p,

$$X(\mathbf{w}) = \big|\Delta_{j,\ell}(\mathfrak{e},T,\mathbf{w})\big|^{\frac{1}{n}},$$

and

$$Y(\mathbf{w}) = \frac{p_{2^{-\ell}T}\big(\pi\circ\mathfrak{F}_{\mathfrak{e},T}(1-2^{-\ell},\mathbf{w}),y\big)}{p_T\big(\pi(\mathfrak{e}),y\big)},$$

will give (A.7) once we check that there exists a $\rho > 0$ for which

$$\text{(A.9)} \qquad \sup_{(T,\mathfrak{e})\in(0,1]\times O(M)} \sup_{\ell\ge 1} \mu_T\Big(\big|\Delta_{j,\ell}(\mathfrak{e},T,\mathbf{w})\big| \ge R^n\Big) \le 2\exp\big(-\rho R^2\big), \quad R \in [0,\infty).$$

Finally, when $j = 0$, (A.9) is an immediate consequence of (A.6). When $1 \le j \le d$, we use the standard exponential estimate for stochastic integrals which says that, for any positive numbers r and $R > 0$ and any progressively measurable $\alpha : [0,1]\times\mathfrak{W} \longrightarrow \mathbb{R}$:

$$\mu_T\left(\int_a^b \alpha(\tau)\,dw_j(\tau) \ge (b-a)^{\frac{1}{2}} rR \;\text{ and }\; \int_a^b \alpha(\tau)^2\,d\tau \le R^2\right) \le \exp\left(-\frac{r^2}{2}\right).$$

Hence, by (A.6), for each $1 \leq j \leq d$,

$$\mu_T\Big(\big|\Delta_{j,\ell}(\mathfrak{e}, T, \mathbf{w})\big| \geq R^n\Big) \leq 2\exp\left(-\frac{R^2}{2}\right) + 2\exp\big(-\epsilon R^2\big),$$

which completes the proof of (2.11).

REFERENCES

[1] J. M. Bismut, *Large Deviations and the Malliavin Calculus*, Progress in Math #45, Birkhäuser, 1984.

[2] R. L. Bishop and R. Crittenden, *Geometry of Manifolds*, Pure and Applied Mathematics #15, Academic Press, 1964.

[3] R. P. Boas, *Entire Functions*, Pure and Applied Mathematics #5, Academic Press, 1954.

[4] M. P. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[5] O. Enchev and D. W. Stroock, *Towards a Riemannian geometry on the path space over a Riemannian Manifold*, J. Funct. Anal., 134 (1995), pp. 392–416.

[6] ———, *Pinned Brownian motion and its perturbations*, Advances in Math., 119 (1996), pp. 127-154.

[7] S. Kusuoka and D. W. Stroock, *Precise asymptotics of certain Wiener functionals*, J. Funct. Anal., 99 (1991), pp. 1–74.

[8] ———, *Asymptotics of certain Wiener functionals with degenerate extrema*, Comm. Pure Appl. Math., 47 (1994), pp. 477–501.

[9] P. Li and S.-T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math., 156 (1986), pp. 153–201.

[10] P. Malliavin and D. W. Stroock, *Short time behavior of the heat kernel and its logarithmic derivatives*, J. Diff. Geom., to appear.

[11] B. Simon, *Functional Integration and Quantum Physics*, Pure and Applied Mathematics # 86, Academic Press, 1979.

[12] T. P. Speed, *Cumulants and partition lattices*, Austral. J. Statist., 25 (1983), pp. 378–388.

[13] R. P. Stanley, *Enumerative Combinatorics, vol. I*, Cambridge Studies in Advanced Mathematics #49, Cambridge University Press, 1996.

[14] D. W. Stroock, *Probability Theory, An Analytic View*, Cambridge University Press, 1994.

[15] ———, *An estimate on the Hessian of the heat kernel*, in Festschrift in Honor of K. Itô, Springer-Verlag, 1995, to appear.

[16] D. W. Stroock and J. Turetsky, *Upper bounds on derivatives of the logarithm of the heat kernel*, to appear.

[17] S. R. Varadhan, *Diffusion processes in a small time interval*, Comm. Pure Appl. Math., 20 (1967), pp. 659–685.