Upper Bounds on Derivatives of the Logarithm of the Heat Kernel

Daniel W. Stroock and James Turetsky

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 $\frac{\partial u}{\partial t} = \frac{1}{2} \Delta u$, where Δ is the Levi-Civita Laplacian. The purpose of this note is to show that for every $n \geq 1$ the n-th derivative of $\log p_t(\cdot, y)$ at x is bounded above by a constant times $\left(\frac{1}{t} + \frac{\operatorname{dist}(x,y)^2}{t^2} \right)^{\frac{n}{2}}$ for $t \in (0, 1]$.

1. Introduction.

Let M be a compact, connected d-dimensional Riemannian manifold and $O(M)$ the associated bundle of orthonormal frames with the fiber map $\pi : O(M) \longrightarrow M$. Recall that, for each $\mathbf{v} \in \mathbb{R}^d$, there is a globally defined *canonical* vector field $\mathfrak{E}(\mathbf{v})$ on $O(M)$ which is the unique horizontal lift to $\mathfrak{e}$ of the vector $\mathfrak{e}\mathbf{v} \in T_{\pi(\mathfrak{e})}M$. In other words, the vector $d\pi\mathfrak{E}(\mathbf{v})_\mathfrak{e} = \mathfrak{e}\mathbf{v}$ in $T_{\pi(\mathfrak{e})}M$ has $\mathbf{v}$ as its coordinates with respect to the frame $\mathfrak{e}$. Next, given $n \geq 1$, we define $\mathfrak{E}^{(n)}F : O(M) \longrightarrow (\mathbb{R}^d)^{\otimes n}$ for $F \in C^n(O(M); \mathbb{R})$ so that

$$
\left(\mathfrak{E}^{(n)} F(\mathfrak{e}), \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n \right)_{(\mathbb{R}^d)^{\otimes n}} = \mathfrak{E} \left(\mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n \right)_\mathfrak{e} F
$$
$$
\equiv \left[\mathfrak{E} \left(\mathbf{v}^n \right) \circ \cdots \circ \mathfrak{E} \left(\mathbf{v}^1 \right) F \right](\mathfrak{e});
$$

and, when $f \in C^n(M; \mathbb{R})$, we set $\mathfrak{E}^{(n)}f \equiv \mathfrak{E}^{(n)}(f \circ \pi)$.

Our purpose here is to derive estimates on logarithmic derivatives of the heat kernel. To be precise, let Δ denote the Laplacian for the Levi-Civita connection, and consider the Cauchy initial value problem for the associated heat equation:

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

By elliptic regularity theory and the strong maximum principle, one knows that there is a unique smooth function

$$(t, x, y) \in (0, \infty) \times M \times M \longmapsto p_t(x, y) \in (0, \infty)$$

with the property that, for every $f \in C(M)$,

$$u_f(t, x) = \int_M f(y) p_t(x, y) \, \lambda_M(dy), \qquad (t, x) \in (0, \infty) \times M,$$

where λ_M stands for the normalized Riemann measure on M. In addition, because Δ is essentially self-adjoint on $L^2(\lambda_M)$, one knows that $p_t(x, y) = p_t(y, x)$. The estimates which we will derive say that, for each $n \geq 1$, there is a $C_n < \infty$ such that

$$(1.1) \qquad \left| \left[\mathfrak{C}^{(n)} \log p_T(\,\cdot\,, y) \right](x) \right|_{(\mathbb{R}^d)^{\otimes n}} \leq C_n \left(\frac{1}{T} + \frac{\mathrm{dist}(x, y)^2}{T^2} \right)^{\frac{n}{2}}$$

$$\text{for} \quad (T, x, y) \in (0, 1] \times M \times M,$$

where $\mathrm{dist}(x, y)$ denotes the Riemannian distance between x and y. At least for $n \in \{1, 2\}$, the results in [6] show that (1.1) is sharp.

Using methods of stochastic control theory, the problem of finding the upper bounds on derivatives of the logarithm of the heat kernel was studied by S.-J. Sheu (cf. [8]) in the following general set-up. He worked on $\mathbb{R}^d$ and considered a uniformly elliptic operators L given by

$$Lf(x) = \frac{1}{2} \sum_{i,j=1}^{d} a^{i,j}(x) \frac{\partial f(x)}{\partial x_i \, \partial x_j} + \sum_{i=1}^{d} b^j(x) \frac{\partial f(x)}{\partial x_j}, \qquad f \in C_0^2\left(\mathbb{R}^d\right),$$

where the coefficients $a^{i,j}$ and b^j are all bounded smooth functions and the matrix $\left((a^{i,j}(x)) \right)$ is symmetric with all its eigenvalues lying between two fixed positive numbers. Let $p_T^L(x, y)$ denote the fundamental solution for the Cauchy initial value problem

$$\frac{\partial u}{\partial t}(t, x) = Lu(t, x), \qquad \text{with} \quad \lim_{t \searrow 0} u(t, x) = f(x);$$

and $E_L \colon (0, \infty) \times M \times M \longrightarrow \mathbb{R}$ is the action functional given by:

$$E_L(T, x, y) = \frac{1}{2} \inf \left\{ \int_0^T \left(\dot{\phi}(t) - b \circ \phi(t), \, a^{-1} \circ \phi(t) \left(\dot{\phi}(t) - b \circ \phi(t) \right) \right)_{\mathbb{R}^d} dt : \right.$$

$$\left. \phi \in C^1\left([0, T]; \mathbb{R}^d\right), \, \phi(0) = x, \text{ and } \phi(T) = y \right\}.$$

Sheu's estimate can be summarized as the statement that for $T \in (0, \infty)$:

$$(1.2) \quad \left| \frac{\partial^n \log p_T^L(\,\cdot\,, y)}{\partial x^\alpha} \right| \leq \begin{cases} C_1 T^{-\frac{1}{2}} \left(1 + E_L(T, x, y)\right)^{\frac{1}{2}} & \text{when } n = 1, \\ C_n T^{-\frac{n}{2}} \left(1 + E_L(T, x, y)\right)^{\frac{3n}{2} - 2} & \text{when } n \geq 2. \end{cases}$$

When L is the Laplacian for some metric on $\mathbb{R}^d$, one knows that

$$E_L(T, x, y) = \frac{\text{dist}(x, y)^2}{T} \, .$$

Hence, when $n \in \{1, 2\}$, our results coincide with Sheu's. However, when $n \geq 3$, our estimate represents an improvement on his. For those who are wondering whether this improvement is a consequence of the difference between his and our settings, the answer is "no": our methods can be modified to cover Sheu's setting and show that his estimate can be improved when $n \geq 3$. Furthermore, it seems likely that Sheu's own methods can be sharpened and made to reproduce our estimates[1].

2. Perturbation of Brownian Paths on a Manifold.

This section gives a brief summary of results from §2 in [9]. For additional background, the reader may want to consult [1] and [4].

We begin by introducing some notation. Let

$$\mathfrak{W} = \left\{ \mathbf{w} \in C\Big([0, \infty)\,;\, \mathbb{R}^d\Big) : \mathbf{w}(0) = \mathbf{0} \right\},$$

thought of as a Frechét space with the topology of uniform convergence on finite intervals. Let $\mathcal{B}$ be the Borel field over $\mathfrak{W}$, and use μ to denote the standard Wiener measure on $(\mathfrak{W}, \mathcal{B})$. Also, for each $t \in [0, \infty)$, let $\mathcal{B}_t$ be the σ-algebra over $\mathfrak{W}$ generated by the maps $\mathbf{w} \rightsquigarrow \mathbf{w}(\tau)$ as τ runs over $[0, t]$.

Next, recall that, for each $\mathfrak{e} \in O(M)$, the tangent space $T_\mathfrak{e} O(M)$ decomposes into the direct sum of the vertical subspace $V_\mathfrak{e} O(M)$ and the horizontal subspace $H_\mathfrak{e} O(M)$. Furthermore, if, for $O \in O(d)$, $R_O : O(M) \longrightarrow O(M)$ is defined so that $R_O \mathfrak{e} \mathbf{v} = \mathfrak{e} O \mathbf{v}$ for all $(\mathfrak{e}, \mathbf{v}) \in O(M) \times \mathbb{R}^d$, and, for $\mathcal{A} \in o(d)$, $\lambda(\mathcal{A})$ is the vector field on $O(M)$ determined by

$$\lambda(\mathcal{A})_\mathfrak{e} = \frac{d}{ds} R_{e^{s\mathcal{A}}} \mathfrak{e} \bigg|_{s=0},$$

---

[1] After seeing our results, E.P. Hsu re-examined Sheu's method and claims to have made the required improvements.

then $V_{\mathfrak{e}}O(M) = \{\lambda(\mathcal{A})_{\mathfrak{e}} : \mathcal{A} \in o(d)\}$. The *connection* 1-form ϕ is the $o(d)$-valued 1-form on $O(M)$ such that $\lambda\left(\phi(\mathfrak{T}_{\mathfrak{e}})\right)_{\mathfrak{e}}$ is the vertical part of $\mathfrak{T}_{\mathfrak{e}} \in T_{\mathfrak{e}}O(M)$. Hence, if the *solder* form ω is the $\mathbb{R}^d$-valued 1-form defined by $d\pi\mathfrak{T}_{\mathfrak{e}} = \mathfrak{e}\omega(\mathfrak{T}_{\mathfrak{e}})$, then, for any $\mathfrak{e} \in O(M)$ and $\mathfrak{T}_{\mathfrak{e}} \in T_{\mathfrak{e}}O(M)$:

$$\mathfrak{T}_{\mathfrak{e}} = \mathfrak{E}\left(\omega(\mathfrak{T}_{\mathfrak{e}})\right)_{\mathfrak{e}} + \lambda\left(\phi(\mathfrak{T}_{\mathfrak{e}})\right)_{\mathfrak{e}}$$

gives the decomposition of $\mathfrak{T}_{\mathfrak{e}}$ into its horizontal and vertical components. Finally, given $\mathfrak{e} \in O(M)$, we define the bilinear map

$$(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \Phi(\boldsymbol{\xi}, \boldsymbol{\eta})_{\mathfrak{e}} \in o(d),$$

called *curvature* form, so that

$$\mathfrak{e}\,\Phi(\boldsymbol{\xi}, \boldsymbol{\eta})_{\mathfrak{e}}\mathbf{v} = \mathrm{Riem}\,(\mathfrak{e}\boldsymbol{\xi}, \mathfrak{e}\boldsymbol{\eta})\,\mathfrak{e}\mathbf{v}, \qquad \mathbf{v} \in \mathbb{R}^d,$$

where

$$\mathrm{Riem}\,(X_x, Y_x)\,Z_x \equiv \nabla_{[X,Y]_x}Z - [\nabla_X, \nabla_Y]_x\,Z$$

is the usual Riemann curvature tensor on M. In this setting, the Ricci curvature matrix $\mathrm{Ric}_{\mathfrak{e}}$ is the symmetric element of $\mathrm{Hom}(\mathbb{R}^d; \mathbb{R}^d)$ determined by

$$\mathrm{Ric}_{\mathfrak{e}}\,\mathbf{v} = \sum_{i=1}^{d} \Phi\left(\boldsymbol{\xi}^i, \mathbf{v}\right)_{\mathfrak{e}} \boldsymbol{\xi}^i, \qquad \mathbf{v} \in \mathbb{R}^d,$$

for any choice of orthonormal basis $\{\boldsymbol{\xi}^i\}_1^d$ in $\mathbb{R}^d$.

For every frame $\mathfrak{e} \in O(M)$, we define $\mathfrak{F}_{\mathfrak{e}} : [0, \infty) \times \mathfrak{W} \longrightarrow O(M)$ to be the μ-almost surely unique $\{\mathcal{B}_t : t \in (0, \infty)\}$-progressively measurable solution to the Stratonovich stochastic differential equation

$$d\mathfrak{F}_{\mathfrak{e}}(t, \mathbf{w}) = \mathfrak{E}\left(\circ\, d\mathbf{w}(t)\right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \quad \text{with} \quad \mathfrak{F}_{\mathfrak{e}}(0, \mathbf{w}) = \mathfrak{e}.$$

Because $\Delta = \sum_1^d d\pi\mathfrak{E}(\boldsymbol{\xi}^i)^2$ for any orthonormal basis $\{\boldsymbol{\xi}^i\}_1^d$, Itô's formula tells us that

$$(2.1) \quad \mathbb{E}^\mu\left[f \circ \pi(\mathfrak{F}_{\mathfrak{e}}(t, \mathbf{w}))\right] = [P_t f]\,(\pi(\mathfrak{e}))$$

$$\equiv \int_M f(y)p_t(\pi(\mathfrak{e}), y)\lambda_M(dy), \quad (t, \mathfrak{e}) \in (0, \infty) \times O(M).$$

In order to use (2.1) to analyze $p_t(x, y)$, we need to introduce two perturbations of the paths $\mathfrak{F}_{\mathfrak{e}}(\,\cdot\,, \mathbf{w})$. To this end, we first define the Ricci flow

$$A \colon [0, \infty) \times C\left([0, \infty); O(M)\right) \longrightarrow \mathrm{Hom}\left(\mathbb{R}^d; \mathbb{R}^d\right)$$

so that

$$(2.2) \qquad \dot{A}(t, \mathfrak{F}) + \tfrac{1}{2} \operatorname{Ric}_{\mathfrak{F}(t)} A(t, \mathfrak{F}) = \mathbf{0} \quad \text{with} \quad A(0, \mathfrak{F}) = I,$$

and set $A_\mathfrak{e}(t, \mathbf{w}) = A(t, \mathfrak{F}_\mathfrak{e}(\,\cdot\,, \mathbf{w}))$.

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

$$(t, \mathbf{w}) \in [0, \infty) \times \mathfrak{W} \longmapsto$$
$$\left(\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(t, \mathbf{w}), \overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(t, \mathbf{w}) \right) \in C^\infty\!\left(\mathbb{R} \,;\, \mathbb{R}^d \right) \times C^\infty(\mathbb{R} \,;\, O(M))$$

is a right-continuous, $\{\mathcal{B}_t : t \in [0, \infty)\}$-progressively measurable map such that

(a) $\left(\left[\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](0), \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](0) \right) = (\mathbf{w}, \mathfrak{F}_\mathfrak{e}(\,\cdot\,, \mathbf{w})) \,;$

(b) $\mathfrak{e}_\mathfrak{e}(s) \equiv \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(0, \mathbf{w}) \right](s)$ satisfies

$$\mathfrak{e}_\mathfrak{e}'(s) \equiv \frac{d}{ds} \mathfrak{e}_\mathfrak{e}(s) = \mathfrak{E}(\mathbf{v})_{\mathfrak{e}_\mathfrak{e}(s)}$$

with $\mathfrak{e}_\mathfrak{e}(0) = \mathfrak{e}$;

(c) for each $s \in \mathbb{R}$, the μ-distribution of

$$\mathbf{w} \rightsquigarrow \left(\left[\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](s), \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](s) \right)$$

is the same as that of $\mathbf{w} \rightsquigarrow \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}_\mathfrak{e}(s)}(\,\cdot\,, \mathbf{w}) \right)$;

(d) for each $t \in [0, \infty)$,

$$\left[\overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(t, \mathbf{w}) \right]'(0) \equiv \frac{d}{ds} \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(t, \mathbf{w}) \right](s) \Big|_{s=0}$$
$$= \mathfrak{E} \left(A_\mathfrak{e}(t, \mathbf{w}) \mathbf{v} \right)_{\mathfrak{F}_\mathfrak{e}(t, \mathbf{w})}$$
$$+ \lambda \left(\int_0^t \Phi \left(\circ \, d\mathbf{w}(\tau), A_\mathfrak{e}(\tau, \mathbf{w}) \mathbf{v} \right)_{\mathfrak{F}_\mathfrak{e}(\tau, \mathbf{w})} \right)_{\mathfrak{F}_\mathfrak{e}(t, \mathbf{w})} ;$$

(e) for each $t \in [0, \infty)$,

$$\left[\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(t, \mathbf{w}) \right]'(0) \equiv \frac{d}{ds} \left[\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(t, \mathbf{w}) \right](s) \Big|_{s=0}$$
$$= A_\mathfrak{e}(t, \mathbf{w}) \mathbf{v}$$
$$+ \int_0^t \left(\int_0^\tau \Phi \left(\circ \, d\mathbf{w}(\tau'), A_\mathfrak{e}(\tau', \mathbf{w}) \mathbf{v} \right)_{\mathfrak{F}_\mathfrak{e}(\tau', \mathbf{w})} \right) \circ d\mathbf{w}(\tau).$$

Our interest in the backward perturbation derives from the fact that, because of (a) and (b),

$$\mathfrak{E}(\mathbf{v})_{\mathfrak{e}} U = \frac{d}{ds} \mathbb{E}^{\mu} \left[F\left(\left[\overleftarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](s), \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](s) \right) \right] \Bigg|_{s=0}$$

when $U(\mathfrak{e}) = \mathbb{E}^{\mu}\left[F\left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w})\right) \right].$

In particular, by (a)—(d), when $U(\mathfrak{e}) \equiv \mathbb{E}^{\mu}[G(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w}))f\circ\pi(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w}))]$, and one can make sense out of all the derivatives involved,

$$(2.3) \quad \mathfrak{E}(\mathbf{v})_{\mathfrak{e}} U = \mathbb{E}^{\mu}\left[G\left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w})\right) \left(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})A_{\mathfrak{e}}(T,\mathbf{w})\mathbf{v}f\right) \right]$$

$$+ \mathbb{E}^{\mu}\left[\frac{d}{ds} G\left(\left[\overleftarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](s), \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](s) \right) \Bigg|_{s=0} f\circ\pi(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})) \right].$$

As (2.3) makes explicit, the backward perturbation allows us to *differentiate under the expectation value*. However, it leaves us with a derivative on f, which we now want to remove with a second perturbation. Namely, the *forward perturbation*

$$(t,\mathbf{w}) \in [0,\infty) \times \mathfrak{W} \longmapsto$$
$$\left(\overrightarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w}), \overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w}) \right) \in C^{\infty}\left(\mathbb{R}\,;\mathbb{R}^{d} \right) \times C^{\infty}(\mathbb{R}\,;O(M))$$

is a right-continuous, $\{\mathcal{B}_{t} : t \in [0,\infty)\}$-progressively measurable map such that

(a') $\left(\left[\overrightarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](0), \left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](0) \right) = \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w}) \right);$

(b') $\left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(0,\mathbf{w}) \right](s) \equiv \mathfrak{e};$

(c') for each $s \in \mathbb{R}$ the $\mu_{s,\mathfrak{e},\mathbf{v}}$-distribution of

$$\mathbf{w} \rightsquigarrow \left(\left[\overrightarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](s), \left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w}) \right](s) \right)$$

is the same as the μ-distribution of $\mathbf{w} \rightsquigarrow (\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w}))$, where $\mu_{s,\mathfrak{e},\mathbf{v}}$ is the probability measure on $\mathfrak{W}$ such that, for each $T \in (0,\infty)$ and $B \in \mathcal{B}_{T}$,

$$\mu_{s,\mathfrak{e},\mathbf{v}}(B) = \mathbb{E}^{\mu}\left[\left[R_{\mathfrak{e},\mathbf{v}}(T,\mathbf{w}) \right](s), B \right]$$

with

$$[R_{\mathfrak{e},\mathbf{v}}(T,\mathbf{w})](s) \equiv \exp\left[-\int_0^T \left(\int_0^s A\left(t, \left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w})\right](\sigma)\right)\mathbf{v}\,d\sigma, d\mathbf{w}(t)\right)_{\mathbb{R}^d}\right.$$
$$\left.-\frac{1}{2}\int_0^T \left|\int_0^s A\left(t, \left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w})\right](\sigma)\right)\mathbf{v}\,d\sigma\right|_{\mathbb{R}^d}^2 dt\right];$$

(d') for each $t \in [0,\infty)$,

$$\left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w})\right]'(0) \equiv \frac{d}{ds}\left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w})\right](s)\bigg|_{s=0}$$
$$= t\mathfrak{E}\left(A_{\mathfrak{e}}(t,\mathbf{w})\mathbf{v}\right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})}$$
$$+ \lambda\left(\int_0^t \tau\Phi\left(\circ\,d\mathbf{w}(\tau), A_{\mathfrak{e}}(\tau,\mathbf{w})\mathbf{v}\right)_{\mathfrak{F}_{\mathfrak{e}}(\tau,\mathbf{w})}\right);$$

(e') for each $t \in [0,\infty)$,

$$\left[\overrightarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w})\right]'(0) \equiv \frac{d}{ds}\left[\overrightarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(t,\mathbf{w})\right](s)\bigg|_{s=0}$$
$$= tA_{\mathfrak{e}}(t,\mathbf{w})\mathbf{v}$$
$$+ \int_0^t \left(\int_0^\tau \tau'\Phi\left(\circ\,d\mathbf{w}(\tau'), A_{\mathfrak{e}}(\tau',\mathbf{w})\mathbf{v}\right)_{\mathfrak{F}_{\mathfrak{e}}(\tau',\mathbf{w})}\right)\circ d\mathbf{w}(\tau).$$

The virtue of the forward perturbation is that, by (a')—(d'),

$$T\mathbb{E}^\mu\left[G\left(\mathbf{w},\mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w})\right)\left(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})A_{\mathfrak{e}}(T,\mathbf{w})\mathbf{v}f\right)\right]$$
$$= \mathbb{E}^\mu\left[\int_0^T \left(A_{\mathfrak{e}}(t,\mathbf{w})\mathbf{v}, d\mathbf{w}(t)\right)_{\mathbb{R}^d} G\left(\mathbf{w},\mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w})\right) f\circ\pi\left(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})\right)\right]$$
$$- \mathbb{E}^\mu\left[\frac{d}{ds}G\left(\left[\overrightarrow{\Theta_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w})\right](s), \left[\overrightarrow{\mathfrak{F}_{\mathfrak{e},\mathbf{v}}}(\,\cdot\,,\mathbf{w})\right](s)\right)\bigg|_{s=0} f\circ\pi\left(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})\right)\right]$$

when G depends only on $(\mathbf{w},\mathfrak{F}_{\mathfrak{e}}) \upharpoonright [0,T]$. Hence, after combining the preceding with (2.3), we see that

$$T\mathfrak{E}(\mathbf{v})_{\mathfrak{e}}U = \mathbb{E}^\mu\left[\int_0^T \left(A_{\mathfrak{e}}(t,\mathbf{w})\mathbf{v}, d\mathbf{w}(t)\right)_{\mathbb{R}^d} G\left(\mathbf{w},\mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w})\right)\right.$$

(2.4)
$$\left.f\circ\pi\left(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})\right)\right]$$

$$+ \mathbb{E}^\mu\left[\left[\mathcal{D}_{T,\mathbf{v}}G\right]\left(\mathbf{w},\mathfrak{F}_{\mathfrak{e}}(\,\cdot\,,\mathbf{w})\right) f\circ\pi\left(\mathfrak{F}_{\mathfrak{e}}(T,\mathbf{w})\right)\right]$$

when G depends only on $(\mathbf{w}, \mathfrak{F}) \upharpoonright [0, T]$ and

$$
\begin{aligned}
[\mathcal{D}_{T, \mathbf{v}} G] \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,, \mathbf{w}) \right) &= T \left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} G \right] \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,, \mathbf{w}) \right) \\
&\quad - \left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} G \right] \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,, \mathbf{w}) \right),
\end{aligned}
\tag{2.5}
$$

with

$$
\left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} G \right] \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,, \mathbf{w}) \right) \equiv \frac{d}{ds} G \left(\left[\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](s), \left[\overleftarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](s) \right) \Big|_{s=0}
$$

and

$$
\left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} G \right] \left(\mathbf{w}, \mathfrak{F}_{\mathfrak{e}}(\,\cdot\,, \mathbf{w}) \right) \equiv \frac{d}{ds} G \left(\left[\overrightarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](s), \left[\overrightarrow{\mathfrak{F}_{\mathfrak{e}, \mathbf{v}}}(\,\cdot\,, \mathbf{w}) \right](s) \right) \Big|_{s=0}.
$$

For future reference, we record a simple fact about the action of $\overleftarrow{\mathcal{D}_{\mathbf{v}}}$ and $\overrightarrow{\mathcal{D}_{\mathbf{v}}}$ on integrals. Namely, assume that $I : [0, \infty) \times \mathfrak{W} \longrightarrow \mathbb{R}$ is a right-continuous, μ-almost surely continuous, $\{\mathcal{B}_t : t \in [0, \infty)\}$-progressively measurable function for which there are right-continuous, μ-almost surely continuous, $\{\mathcal{B}_t : t \in [0, \infty)\}$-progressively measurable versions of $(t, \mathbf{w}) \rightsquigarrow \left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} I(t, \cdot) \right](\mathbf{w})$ and $(t, \mathbf{w}) \rightsquigarrow \left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} I(t, \cdot) \right](\mathbf{w})$. Then,

(2.6)

$$
J(t, \mathbf{w}) = \int_0^t I(\tau, \mathbf{w}) \, d\tau \implies
\begin{cases}
\left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} J(t, \cdot) \right](\mathbf{w}) = \displaystyle\int_0^t \left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} I(\tau, \cdot) \right](\mathbf{w}) \, d\tau, \\[3mm]
\left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} J(t, \cdot) \right](\mathbf{w}) = \displaystyle\int_0^t \left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} I(\tau, \cdot) \right](\mathbf{w}) \, d\tau,
\end{cases}
$$

and, for any $\boldsymbol{\xi} \in \mathbb{R}^d$ (cf. (e) and (e') above):

(2.7)

$$
J(t, \mathbf{w}) = \int_0^t I(\tau, \mathbf{w}) \, d \left(\boldsymbol{\xi}, \mathbf{w}(\tau) \right)_{\mathbb{R}}^d \implies
$$

$$
\begin{cases}
\left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} J(t, \cdot) \right](\mathbf{w}) = \displaystyle\int_0^t \left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} I(\tau, \cdot) \right](\mathbf{w}) \, d \left(\boldsymbol{\xi}, \mathbf{w}(\tau) \right)_{\mathbb{R}}^d \\[3mm]
\qquad\qquad + \displaystyle\int_0^t I(\tau, \mathbf{w}) \, d \left(\boldsymbol{\xi}, \left[\overleftarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(\tau, \mathbf{w}) \right]'(0) \right)_{\mathbb{R}}^d, \\[5mm]
\left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} J(t, \cdot) \right](\mathbf{w}) = \displaystyle\int_0^t \left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} I(\tau, \cdot) \right](\mathbf{w}) \, d \left(\boldsymbol{\xi}, \mathbf{w}(\tau) \right)_{\mathbb{R}}^d \\[3mm]
\qquad\qquad + \displaystyle\int_0^t I(\tau, \mathbf{w}) \, d \left(\boldsymbol{\xi}, \left[\overrightarrow{\Theta_{\mathfrak{e}, \mathbf{v}}}(\tau, \mathbf{w}) \right]'(0) \right)_{\mathbb{R}}^d.
\end{cases}
$$

Additionally, we want to record the action of $\overleftarrow{\mathcal{D}_\mathbf{v}}$ and $\overrightarrow{\mathcal{D}_\mathbf{v}}$ on a few special functions. First, by using (2.2) and the method of variation of parameters, one can see (cf. (d) above) that

$$\left[\overleftarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)\right](\mathbf{w})$$

$$= -\frac{1}{2} A_\mathfrak{e}(t,\mathbf{w}) \int_0^t A_\mathfrak{e}(\tau,\mathbf{w})^{-1} \left(\mathfrak{C}\left(A_\mathfrak{e}(\tau,\mathbf{w})\mathbf{v}\right) \mathrm{Ric} \right.$$

$$\left. +\lambda\left(\int_0^\tau \Phi\left(\circ\, d\mathbf{w}(\tau'),\, A_\mathfrak{e}(\tau',\mathbf{w})\mathbf{v}\right)\right)\mathrm{Ric} \right)_{\mathfrak{F}_\mathfrak{e}(\tau,\mathbf{w})} A_\mathfrak{e}(\tau,\mathbf{w})\,d\tau.$$

At the same time, a simple computation shows that, for any $\mathcal{A} \in o(d)$, $\lambda(\mathcal{A})\,\mathrm{Ric} = [\mathrm{Ric}, \mathcal{A}]$. Hence,

(2.8)

$$\left[\overleftarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)\right](\mathbf{w})$$

$$= -\frac{1}{2} A_\mathfrak{e}(t,\mathbf{w}) \int_0^t A_\mathfrak{e}(\tau,\mathbf{w})^{-1} \left(\mathfrak{C}\left(A_\mathfrak{e}(\tau,\mathbf{w})\mathbf{v}\right) \mathrm{Ric} \right.$$

$$\left. + \left[\mathrm{Ric}, \int_0^\tau \Phi\left(\circ\, d\mathbf{w}(\tau'),\, A_\mathfrak{e}(\tau',\mathbf{w})\mathbf{v}\right)_{\mathfrak{F}_\mathfrak{e}(\tau',\mathbf{w})}\right] \right)_{\mathfrak{F}_\mathfrak{e}(\tau,\mathbf{w})} A_\mathfrak{e}(\tau,\mathbf{w})\,d\tau.$$

Similarly (cf. (d'))

(2.9)

$$\left[\overrightarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)\right](\mathbf{w})$$

$$= -\frac{1}{2} A_\mathfrak{e}(t,\mathbf{w}) \int_0^t A_\mathfrak{e}(\tau,\mathbf{w})^{-1} \left(\tau\mathfrak{C}\left(A_\mathfrak{e}(\tau,\mathbf{w})\mathbf{v}\right) \mathrm{Ric} \right.$$

$$\left. + \left[\mathrm{Ric}, \int_0^\tau \tau'\Phi\left(\circ\, d\mathbf{w}(\tau'),\, A_\mathfrak{e}(\tau',\mathbf{w})\mathbf{v}\right)_{\mathfrak{F}_\mathfrak{e}(\tau',\mathbf{w})}\right] \right)_{\mathfrak{F}_\mathfrak{e}(\tau,\mathbf{w})} A_\mathfrak{e}(\tau,\mathbf{w})\,d\tau.$$

Furthermore,

(2.10)

$$\left[\overleftarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)^{-1}\right](\mathbf{w}) = -A_\mathfrak{e}(t,\mathbf{w})^{-1} \left[\overleftarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)\right](\mathbf{w}) A_\mathfrak{e}(t,\mathbf{w})^{-1}$$

$$\left[\overrightarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)^{-1}\right](\mathbf{w}) = -A_\mathfrak{e}(t,\mathbf{w})^{-1} \left[\overrightarrow{\mathcal{D}_\mathbf{v}} A_\mathfrak{e}(t,\,\cdot\,)\right](\mathbf{w}) A_\mathfrak{e}(t,\mathbf{w})^{-1}.$$

Finally, we will need to know how $\overleftarrow{\mathcal{D}_{\mathbf{v}}}$ and $\overrightarrow{\mathcal{D}_{\mathbf{v}}}$ act on horizontal derivatives of Φ. Thus, let $\mathbf{V} = \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n$ and $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R}^d \times \mathbb{R}^d$ be given. Then,

$$\left[\overleftarrow{\mathcal{D}_{\mathbf{V}}} \mathfrak{E}(\mathbf{v})_{\mathfrak{F}_{\mathfrak{e}}(t,\cdot)} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta}) \right](\mathbf{w})$$

$$= \left(\mathfrak{E}\left(\mathbf{V} \otimes (A_{\mathfrak{e}}(t, \mathbf{w})\mathbf{V}) \right) + \lambda\left(\overleftarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w}) \right) \circ \mathfrak{E}(\mathbf{V}) \right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta})$$

where (cf. (d))

$$\overleftarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w}) \equiv \int_0^t \Phi\left(\circ \, d\mathbf{w}(\tau), A_{\mathfrak{e}}(\tau, \mathbf{w})\mathbf{v} \right)_{\mathfrak{F}_{\mathfrak{e}}(\tau, \mathbf{w})},$$

But, for any $\mathcal{A} \in o(d)$,

$$[\lambda(\mathcal{A}), \mathfrak{E}(\mathbf{V})] = \mathfrak{E}(\mathcal{A}\mathbf{V}), \quad \text{where} \quad \mathcal{A}\mathbf{V} \equiv \sum_{m=1}^n \mathbf{v}^1 \otimes \cdots \otimes \mathcal{A}\mathbf{v}^m \otimes \cdots \otimes \mathbf{v}^n,$$

and

$$\lambda(\mathcal{A}) \Phi(\boldsymbol{\xi}, \boldsymbol{\eta}) = [\Phi(\boldsymbol{\xi}, \boldsymbol{\eta}), \mathcal{A}] + \Phi(\mathcal{A}\boldsymbol{\xi}, \boldsymbol{\eta}) + \Phi(\boldsymbol{\xi}, \mathcal{A}\boldsymbol{\eta}).$$

Hence,

(2.11)

$$\left[\overleftarrow{\mathcal{D}_{\mathbf{v}}} \mathfrak{E}(\mathbf{V})_{\mathfrak{F}_{\mathfrak{e}}(t,\cdot)} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta}) \right](\mathbf{w})$$

$$= \mathfrak{E}\left(\mathbf{V} \otimes (A_{\mathfrak{e}}(t, \mathbf{w})\mathbf{v}) \right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta})$$

$$+ \mathfrak{E}\left(\overleftarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w})\mathbf{V} \right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta})$$

$$+ \left[\overleftarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w}), \mathfrak{E}(\mathbf{V})_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta}) \right]$$

$$+ \mathfrak{E}(\mathbf{V})_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \left(\Phi\left(\overleftarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w})\boldsymbol{\xi}, \boldsymbol{\eta} \right) + \Phi\left(\boldsymbol{\xi}, \overleftarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w})\boldsymbol{\eta} \right) \right).$$

Similarly,

(2.12)

$$\left[\overrightarrow{\mathcal{D}_{\mathbf{v}}} \mathfrak{E}(\mathbf{V})_{\mathfrak{F}_{\mathfrak{e}}(t,\cdot)} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta}) \right](\mathbf{w})$$

$$= t\mathfrak{E}\left(\mathbf{V} \otimes (A_{\mathfrak{e}}(t, \mathbf{w})\mathbf{v}) \right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta})$$

$$+ \mathfrak{E}\left(\overrightarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w})\mathbf{V} \right)_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta})$$

$$+ \left[\overrightarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w}), \mathfrak{E}(\mathbf{V})_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \Phi(\boldsymbol{\xi}, \boldsymbol{\eta}) \right]$$

$$+ \mathfrak{E}(\mathbf{V})_{\mathfrak{F}_{\mathfrak{e}}(t,\mathbf{w})} \left(\Phi\left(\overrightarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w})\boldsymbol{\xi}, \boldsymbol{\eta} \right) + \Phi\left(\boldsymbol{\xi}, \overrightarrow{\mathcal{A}_{\mathfrak{e}}}(t, \mathbf{w})\boldsymbol{\eta} \right) \right),$$

where (cf. (d'))

$$\overrightarrow{\mathcal{A}_\mathfrak{e}}(t, \mathbf{w}) \equiv \int_0^t \tau \Phi \left(\circ \, d\mathbf{w}(\tau), A_\mathfrak{e}(\tau, \mathbf{w}) \mathbf{v} \right)_{\mathfrak{F}_\mathfrak{e}(\tau, \mathbf{w})},$$

3. Estimates on Derivatives of the Heat Flow.

Let $f : M \longrightarrow (0, \infty)$ be a smooth function. In this section we will combine (2.5)— (2.12) with elementary estimates on Itô stochastic integrals to estimate (cf. (2.1))

$$\frac{\mathfrak{E}^{(n)} \left((P_T f) \circ \pi \right)}{(P_T f) \circ \pi}.$$

To be more precise, this section is devoted to the proof of the following statement.

Theorem 3.1. *Set*

$$Y_f(T, \mathbf{w}) \equiv \frac{f \circ \pi(\mathfrak{F}_\mathfrak{e}(T, \mathbf{w}))}{[P_T f](\pi(\mathfrak{e}))}.$$

Then, for any $n \geq 1$, there is a constant $C_n < \infty$, depending on n and the derivatives of the Riemann curvature, such that

$$(3.2) \quad \left\| \frac{\mathfrak{E}^{(n)}(P_T f) \circ \pi}{(P_T f) \circ \pi} \right\|_{(\mathbb{R}^d)^{\otimes n}}$$

$$\leq C_n T^{-\frac{n}{2}} \left(1 + \log \|Y_f(T, \mathbf{w})\|_{L^\infty(\mu)} \right)^{\frac{n}{2}}, \quad T \in (0, 1].$$

Given the calculations made in §2, the proof of Theorem 3.1 is basically an exercise in bookkeeping, most of which is taken care of by the next lemma.

Lemma 3.3. *Let $\{\mathbf{v}^m\}_1^\infty \subseteq \mathbb{R}^d$ be given, and, using induction, define $\{J_n\}_1^\infty$ so that (cf. (2.6))*

$$J_n(t, \mathbf{w}) = \begin{cases} \displaystyle\int_0^{t \wedge 1} \left(A_\mathfrak{e}(\tau, \mathbf{w}) \mathbf{v}^1, d\mathbf{w}(\tau) \right)_{\mathbb{R}}^d, & \text{if } n = 1; \\[4mm] J_n(t, \mathbf{w}) = \left[\mathcal{D}_{t, \mathbf{v}^n} J_{n-1}(t, \cdot) \right](\mathbf{w}), & \text{if } n \geq 1. \end{cases}$$

Then, for each $n \geq 1$, J_n is a finite linear combination of elements from (cf. the Appendix)

$$(3.4) \quad \mathcal{P}_n \equiv \bigcup_{k \leq n \leq k + 2\ell} \mathcal{P}(k, \ell).$$

Proof. Let $\boldsymbol{\Psi}$ denote the class of functions $\Psi : [0, \infty) \times \mathfrak{W} \longmapsto \mathbb{R}$ which can be expressed as a finite product in which each factor is either constant or has the form

$$\big(\boldsymbol{\eta}, A_\mathfrak{e}(t \wedge 1, \mathbf{w})\boldsymbol{\xi}\big)_\mathbb{R}^d, \quad \big(\boldsymbol{\eta}, A_\mathfrak{e}(t \wedge 1, \mathbf{w})^{-1}\boldsymbol{\xi}\big)_\mathbb{R}^d, \quad \text{or} \quad \Big(\mathfrak{e}^{(m)}\Phi(\boldsymbol{\xi}, \boldsymbol{\eta})_{\mathfrak{F}_\mathfrak{e}(t,\mathbf{w})}\boldsymbol{\xi}', \boldsymbol{\eta}'\Big)_\mathbb{R}^d,$$

for some $\boldsymbol{\xi}, \boldsymbol{\xi}', \boldsymbol{\eta}, \boldsymbol{\eta}' \in \mathbb{R}^d$, and $m = 0, 1, 2, \ldots$. We will say that J is an *integral of order* (k, ℓ) if

$$J(t, \mathbf{w}) = \int_0^t \Psi_{k+\ell}(\tau_{k+\ell}, \mathbf{w}) \left(\int_0^{\tau_{k+\ell}} \Psi_{k+\ell-1}(\tau_{k+\ell-1}, \mathbf{w}) \cdots \right.$$

$$\left. \left(\int_0^{\tau_2} \Psi_1(\tau_1, \mathbf{w}) \, d\alpha_1(\tau_1, \mathbf{w}) \right) \cdots d\alpha_{k+\ell-1}(\tau_{k+\ell-1}, \mathbf{w}) \right) d\alpha_{k+\ell}(\tau_{k+\ell}, \mathbf{w}),$$

where, for each $m = 1, \ldots, k + \ell$, $\Psi_m \in \boldsymbol{\Psi}$ and

$$\alpha_m(t, \mathbf{w}) = \begin{cases} t, & \text{for } \ell \text{ subscripts } m; \\ (\boldsymbol{\xi}_m, \mathbf{w}(t))_\mathbb{R}^d, & \text{otherwise, and } \boldsymbol{\xi}_m \in \mathbb{R}^d. \end{cases}$$

Since, by Lemma A.1, every integral of order (k, ℓ) is an element of $\mathcal{P}_n$, it suffices for us to check that, for each $n \geq 1$, J_n is a finite linear combination of integrals of order (k, ℓ) with $k \leq n \leq k + 2\ell$. But, trivially, this is the case when $n = 1$, and the general case follows by induction when one takes into account the computations in (2.5)— (2.12). $\square$

Proof of Theorem 3.1. Given Lemma 3.3 and (2.4), it is clear that, for any $n \geq 1$ and $\mathbf{V} = \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n \in \big(\mathbb{R}^d\big)^{\otimes n}$, there is an $I_\mathbf{V} \in \text{span}\,(\mathcal{P}_n)$ such that

$$T^n \left[\frac{\mathfrak{E}(\mathbf{V})P_T f \circ \pi}{P_T f \circ \pi} \right](\mathfrak{e}) = \mathbb{E}^\mu \left[I_\mathbf{V}(T, \mathbf{w}) \frac{f \circ \pi(\mathfrak{F}_\mathfrak{e}(T, \mathbf{w}))}{P_T f(\pi(\mathfrak{e}))} \right], \quad T \in (0, 1].$$

Thus, (3.2) follows from (A.4). $\square$

4. Derivation of the Main Estimate.

At this point, the derivation of (1.1) is quite easy. Indeed, the first step is to observe that, for any $n \geq 1$ and $\mathbf{V} = \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^n$, $\mathfrak{E}(\mathbf{V}) \log P_T f$ can be

estimated by quantities of the sort already estimated. To be more precise, we denote by $\Sigma_{m,n}$ the set of all partitions of the set $\mathcal{S}_n = \{1, \ldots, n\}$ into non-empty *disjoined* increasingly ordered blocks $\beta^1, \ldots, \beta^m$. If $\boldsymbol{\beta} = (\beta_1, \ldots \beta_k)$, with $1 \leq \beta_1 < \cdots < \beta_k \leq n$, is a block in a partition of $\mathcal{S}_n$, we set $|\boldsymbol{\beta}| = k$ and $\mathbf{V}^{\boldsymbol{\beta}} = \mathbf{v}^{\beta_1} \otimes \cdots \otimes \mathbf{v}^{\beta_k}$. Then, using induction, one can check that:

$$(4.1) \quad \mathfrak{E}(\mathbf{V}) \log P_T f$$
$$= \sum_{m=1}^{n} \frac{(-1)^{m-1}(m-1)!}{(P_T f)^m} \sum_{\{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^m\} \in \Sigma_{m,n}} \prod_{\ell=1}^{m} \mathfrak{E}\left(\mathbf{V}^{\boldsymbol{\beta}^\ell}\right) P_T f.$$

But, by (3.2), for any $T \in (0, 1]$:

$$\left| \frac{\mathfrak{E}\left(\mathbf{V}^{\boldsymbol{\beta}^\ell}\right) P_T f}{P_T f} \left(\pi(\mathfrak{e})\right) \right|$$
$$\leq C_{|\boldsymbol{\beta}^l|} \left\| \mathbf{V}^{\boldsymbol{\beta}^\ell} \right\|_{(\mathbb{R}^d)^{\otimes n}} T^{-\frac{|\boldsymbol{\beta}^l|}{2}} \left(1 + \log \|Y_f(T, \mathbf{w})\|_{L^\infty(\mu)}\right)^{\frac{|\boldsymbol{\beta}^\ell|}{2}},$$

and so we now know that

$$\left\| \left[\mathfrak{E}^{(n)} \log \left((P_T f) \circ \pi\right)(\mathfrak{e}) \right] \right\|_{(\mathbb{R}^d)^{\otimes n}}$$
$$\leq C'_n T^{-\frac{n}{2}} \left(1 + \log \|Y_f(T, \mathbf{w})\|_{L^\infty(\mu)}\right)^{\frac{n}{2}}, \quad T \in (0, 1],$$

for a $C'_n < \infty$ which depends only on n and bounds on the Riemann curvature and its derivatives. In particular, since $p_{2T}(x, y) = [P_T(p_T(\cdot, y))](x)$, we find that

$$\left\| \left[\mathfrak{E}^{(n)} \log p_{2T}(\pi(\cdot), y) \right](\mathfrak{e}) \right\|_{(\mathbb{R}^d)^{\otimes n}}$$
$$\leq C'_n T^{-\frac{n}{2}} \left(1 + \log \left\| \frac{p_T(\cdot, y)}{p_{2T}(\pi(\mathfrak{e}), y)} \right\|_{C(M;\mathbb{R})} \right), \quad T \in (0, 1].$$

Finally, since (cf. [5] or [3]) there is a $\theta \in (0, 1)$, so that

$$\frac{\theta}{T^{\frac{d}{2}}} \exp\left[-\frac{\text{dist}(x, y)^2}{\theta T}\right] \leq p_T(x, y) \leq \frac{1}{\theta T^{\frac{d}{2}}}, \quad (T, x, y) \in (0, 1] \times M \times M,$$

the estimate in (1.1) follows after an elementary manipulation.

5. Appendix.

We collect here some elementary facts and estimates. To begin with, let $\mathcal{P}$ denote the vector space of right-continuous, $\{\mathcal{B}_t : t \in [0, \infty)\}$-progressively measurable functions $I : [0, \infty) \times \mathfrak{W} \longmapsto \mathbb{R}$, and define

$$I^*(T, \mathbf{w}) = \sup\{|I(t, \mathbf{w})| : t \in [0, T]\}, \qquad (T, \mathbf{w}) \in [0, \infty) \times \mathfrak{W}.$$

Next, for $\ell \in \mathbb{N}$, let $\mathcal{P}(0, \ell)$ denote the class of $I \in \mathcal{P}$ for which $T^{-\ell} I^*(T, \mathbf{w})$ is in $L^\infty(\mu)$; and, for $(k, \ell) \in (\mathbb{N} \setminus \{0\}) \times \mathbb{N}$, let $\mathcal{P}(k, \ell)$ be the set of $I \in \mathcal{P}$ with the property that

$$\varlimsup_{R \to \infty} R^{-2} \log \left(\sup_{T \in (0, \infty)} \mu \left(T^{-\frac{k}{2} - \ell} I^*(T, \mathbf{w}) \geq R^k \right) \right) < 0.$$

Lemma A.1. *For each (k, ℓ), $\mathcal{P}(k, \ell)$ is a vector space, and if $I \in \mathcal{P}(k, \ell)$ and $J \in \mathcal{P}(m, n)$, then $IJ \in \mathcal{P}(k+m, \ell+n)$. In addition, if $I \in \mathcal{P}(k, \ell)$ and*

$$J(t, \mathbf{w}) = \int_0^t I(\tau, \mathbf{w}) \, d\tau, \qquad (t, \mathbf{w}) \in [0, \infty) \times \mathfrak{W},$$

then $J \in \mathcal{P}(k, \ell+1)$. Finally, if $\{I_i\}_1^d \subseteq \mathcal{P}(k, \ell)$ and

$$J(t, \mathbf{w}) = \sum_{i=1}^d \int_0^t I_i(\tau, \mathbf{w}) \, dw_i(\tau), \qquad (t, \mathbf{w}) \in [0, \infty) \times \mathfrak{W},$$

then $J \in \mathcal{P}(k+1, \ell)$.

Proof. Everything except the final statement is completely elementary. Moreover, the final statement follows immediately once one recalls the fact that (cf. [7]), for any progressively measurable $\mathbf{I} \colon [0, \infty) \times \mathfrak{W} \longrightarrow \mathbb{R}^d$,

$$\mu \left(\sup_{t \in [0,T]} \left| \int_0^t (\mathbf{I}(\tau, \mathbf{w}), d\mathbf{w}(\tau))_{\mathbb{R}^d} \right| \geq \rho R \text{ and } \int_0^T |\mathbf{I}(t, \mathbf{w})|_{\mathbb{R}^d}^2 \, dt \leq R^2 \right)$$
$$\leq \exp \left(-\frac{\rho^2}{2} \right).$$

$\square$

Lemma A.2. *Let X and Y be random variables satisfying the conditions*

$$\mathbb{E}\left[e^{X^2}\right] \leq M < \infty, \quad \mathbb{E}[Y] = 1, \quad Y \geq 0, \quad and \; \|Y\|_\infty < \infty.$$

Then, for each $p \in (0, \infty)$ there is a universal $C_p \in (0, \infty)$ such that

$$(A.3) \qquad \mathbb{E}\left[|X|^p \, Y\right] \leq C_p \left(\log \|Y\|_\infty + \log M\right)^{\frac{p}{2}}.$$

Proof. By Jensen's inequality, it suffices to handle p's which are even integers. Further, without loss in generality, we will assume that $X > 0$ and $Y > 0$. In particular, by Jensen's inequality, this means that

$$\mathbb{E}\left[X^2 Y\right] - \mathbb{E}\left[Y \log Y\right] = \mathbb{E}\left[\log \left(\frac{e^{X^2}}{Y}\right) Y\right] \leq \log \mathbb{E}\left[e^{X^2}\right];$$

and therefore

$$A_2 \equiv \mathbb{E}\left[X^2 Y\right] \leq \mathbb{E}\left[Y \log Y\right] + \log M \leq \log \|Y\|_\infty + \log M.$$

More generally, if $A_{2m} \equiv \mathbb{E}\left[X^{2m} Y\right]$, then

$$\begin{aligned} A_{2(m+1)} &= A_{2m}\mathbb{E}\left[X^2 \frac{X^{2m}Y}{A_{2m}}\right] \\ &\leq A_{2m}\left(\mathbb{E}\left[\frac{X^{2m}Y}{A_{2m}} \log \frac{X^{2m}Y}{A_{2m}}\right] + \log M\right) \\ &\leq m\mathbb{E}\left[X^{2m}Y \log X^2\right] + \left(\log \|Y\|_\infty + \log M\right) A_{2m} - A_{2m} \log A_{2m} \\ &\leq \frac{A_{2(m+1)}}{2} + \left(m \log 2m + \log \|Y\|_\infty + \log M\right) A_{2m}, \end{aligned}$$

where we have used the easily verified facts that

$$x \log x \geq -e^{-1} \quad and \quad \log x \leq \frac{x}{2m} + \log 2m - 1 \quad for \; all \; x \in (0, \infty).$$

Clearly, the preceding shows that

$$A_{2(m+1)} \leq 2 \left(m \log 2m + \log \|Y\|_\infty + \log M\right) A_{2m},$$

and the desired estimate follows immediately from this.

Finally, for $I \in \mathcal{P}(k, \ell)$, with $k \geq 1$, define

$$M(I) = \inf \left\{ M \in [1, \infty) : \sup_{T \in (0, \infty)} \mathbb{E} \left[\exp \left(\sqrt[\frac{2}{k}]{\frac{I^*(T, \mathbf{w})}{MT^{\frac{k}{2}+\ell}}} \right) \right] \leq M \right\},$$

and conclude from (A.3) that

$(A.4)$ $\mathbb{E}^\mu \left[I^*(T)Y \right] \leq C_k M(I) T^{\frac{k}{2}+\ell} \left(\log M(I) + \log \|Y\|_\infty \right)^{\frac{k}{2}}$

         for all $T > 0$ and non-negative Y's with $\mathbb{E}^\mu[Y] = 1$.

□

References.

[1] R. L. Bishop and R. Crittenden, *Geometry of Manifolds*, Pure and Applied Math **15**, Academic Press (1964).

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

[3] E. B. Davies, *Heat Kernels and Spectral Theory*, Cambridge Tracts in Mathematics **92**, Cambridge University Press (1989).

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

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

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

[7] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 2nd edition Grundlehren der mathematischen Wissenschaften **293**, Springer-Verlag (1994).

[8] S.-J. Sheu *Some estimates of the transition density of a nondegenerate diffusion markov process*, Ann. Prob. **19** (1991).

[9] D. W. Stroock, *An estimate on the Hessian of the heat kernel*, Festschrift in Honor of K. Itô, Springer-Verlag (1995) to appear.

[10] D. W. Stroock and O. Zeitouni, *Variations on a theme by Bismut*, Astérisque Volume in Honor of P.A. Meyer and J. Neveu, to appear.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
2-272, 77 MASS. AVE.
CAMBRIDGE, MA 02139, USA.

AND

HARVARD UNIVERSITY
1 OXFORD ST.
CAMBRIDGE, MA 02138, USA.
E-mail addresses: turetsky@math.harvard.edu
                    dws@math.mit.edu