D. W. STROOCK

O. ZEITOUNI

Variations on a theme by Bismut

Astérisque, tome 236 (1996), p. 291-301

<http://www.numdam.org/item?id=AST_1996__236__291_0>

Variations on a Theme by Bismut

D. W. Stroock, O. Zeitouni

Abstract. — Let M be a compact, connected, Riemannian manifold of dimension d, let $\{P_t : t > 0\}$ denote the Markov semigroups on $C(M)$ determined by $\frac{1}{2}\Delta$, and let $p_t(x,y)$ denote the kernel (with respect to the Riemannian volume measure) for the operator P_t. (The existence of this kernel as a positive, smooth function is well-known, see e.g. [D].) Bismut's celebrated formula, presented in [B], equates $\nabla \log\big(p_t(\,\cdot\,,y)\big)$ with certain stochastic integrals (see (20) below.) Various derivations of this formula and its extensions can be found in [AM], [EL] and [N]. In this note, we give a quick derivation of Bismut's and related formulae by lifting considerations to the bundle of orthonormal frames, using Bochner's identity, and applying a little elementary stochastic analysis. Some consequences of these identities are then explored. In particular, after deriving a standard logarithmic Sobolev inequality, we present (see (26)) a sharp pointwise estimate on the logarithmic derivative of the heat kernel in terms of known estimates on the heat kernel itself.

§1 Bismut's Formula and Variations

Let $\mathcal{O}(M)$ denote the bundle of orthonormal frames associated to M, equipped with the Lévi–Civita connection. (Throughout, we will take our basic reference for differential geometry to be the book [BC]. In particular, see Chapter 7 for an explanation of $\mathcal{O}(M)$.) The advantage gained by moving considerations to $\mathcal{O}(M)$ is that many differential geometric quantities resemble their classical analogs. For example, if $(\mathbf{e}_1, \ldots, \mathbf{e}_d)$ denotes the standard orthonormal basis in $\mathbb{R}^d$ and $\mathfrak{E}_1, \ldots, \mathfrak{E}_d$ are the corresponding basic vector fields on $\mathcal{O}(M)$ (i.e., $\mathfrak{E}_k$ is the horizontal vector field on $\mathcal{O}(M)$ for which $d\pi\mathfrak{E}_k(\mathfrak{f}) = \mathfrak{f}\mathbf{e}_k$ at each $\mathfrak{f} \in \mathcal{O}(M)$), then we can define the gradient $\mathfrak{f} \in \mathcal{O}(M) \longmapsto \nabla_{\mathfrak{f}}\varphi \in \mathbb{R}^d$ for $\varphi \in C^1(M)$ so that

$$\nabla\varphi(\mathfrak{f}) = \nabla_{\mathfrak{f}}\varphi = \sum_1^d \mathfrak{E}_k(\mathfrak{f})(\varphi \circ \pi)\mathbf{e}_k, \tag{1}$$

where $\pi : \mathcal{O}(M) \longrightarrow M$ denotes the fiber map. Similarly, if, for $F \in C^2\big(\mathcal{O}(M)\big)$,

$$\Delta F = \sum_{k=1}^d \mathfrak{E}_k^2 F, \tag{2}$$

then

$$\Delta\varphi \equiv \Delta(\varphi \circ \pi), \quad \varphi \in C^2(M), \tag{3}$$

is well-defined as a function on M and, in fact, gives the action of the standard Laplacian (Laplace–Beltrami operator) on φ.

Next, let $\phi : T\big(\mathcal{O}(M)\big) \longrightarrow o(d)$ (the Lie algebra of $d \times d$ skew symmetric matrices) denote the connection 1-form determined by the Lévi–Civita connection (cf. §5.2 in [BC]). That is, for any $\mathfrak{f} \in \mathcal{O}(M)$ and $X_\mathfrak{f} \in T_\mathfrak{f}\big(\mathcal{O}(M)\big)$, $\phi(X_\mathfrak{f})$ is determined so that

$$\mathbf{H}X_\mathfrak{f} \equiv X_\mathfrak{f} - \lambda\big(\phi(X_\mathfrak{f})\big)$$

is the horizontal component of $X_\mathfrak{f}$, where, for $A \in o(d)$, $\lambda(A) \in T\big(\mathcal{O}(M)\big)$ is the vertical vector field such that

$$\lambda(A)F(\mathfrak{f}) = \frac{d}{ds}F\big(R_{e^{sA}}\mathfrak{f}\big)\Big|_{s=0}, \quad \mathfrak{f} \in \mathcal{O}(M),$$

and $R_\mathcal{O} : \mathcal{O}(M) \longrightarrow \mathcal{O}(M)$ is the natural right action given by $R_\mathcal{O}\mathfrak{f}\mathbf{v} = \mathfrak{f}\mathcal{O}\mathbf{v}$ for $\mathcal{O} \in \mathrm{O}(d)$, $\mathfrak{f} \in \mathcal{O}(M)$, and $\mathbf{v} \in \mathbb{R}^d$. Then, the curvature 2-form $\Phi : T\big(\mathcal{O}(M)\big) \times T\big(\mathcal{O}(M)\big) \longrightarrow o(d)$ is defined to be the horizontal part of the exterior derivative $d\phi$ of ϕ:

$$\Phi(X_\mathfrak{f}, Y_\mathfrak{f}) = d\phi\big(\mathbf{H}X_\mathfrak{f}, \mathbf{H}Y_\mathfrak{f}\big), \quad X_\mathfrak{f}, Y_\mathfrak{f} \in T_\mathfrak{f}\big(\mathcal{O}(M)\big). \tag{4}$$

As a consequence of the fact that the Lévi–Civita connection is torsion free and the second structural equation (cf. Theorem 4 in §6.2 of [BC]), one finds (cf. §5.3 of [BC]) that the commutator of $\mathfrak{E}_k$ with $\mathfrak{E}_\ell$ is vertical and is given by

$$\big[\mathfrak{E}_k, \mathfrak{E}_\ell\big](\mathfrak{f}) = -\lambda\big(\Phi_{k,\ell}(\mathfrak{f})\big), \quad \text{where } \Phi_{k,\ell} \equiv \Phi\big(\mathfrak{E}_k, \mathfrak{E}_\ell\big). \tag{5}$$

In particular, for $\varphi \in C^2(M)$,

$$\mathfrak{E}_k\mathfrak{E}_\ell(\varphi \circ \pi) = \mathfrak{E}_\ell\mathfrak{E}_k(\varphi \circ \pi),$$

and, for $\varphi \in C^3(M)$,

$$\begin{aligned}\mathfrak{E}_k^2\mathfrak{E}_\ell(\varphi \circ \pi) &= \mathfrak{E}_\ell\mathfrak{E}_k^2(\varphi \circ \pi) - \lambda\big(\Phi_{k,\ell}\big)\mathfrak{E}_k(\varphi \circ \pi) \\ &= \mathfrak{E}_\ell\mathfrak{E}_k^2(\varphi \circ \pi) - \sum_{j=1}^{d}\big(\Phi_{k,\ell}\mathbf{e}_k, \mathbf{e}_j\big)_{\mathbf{R}^d}\mathfrak{E}_j(\varphi \circ \pi).\end{aligned}$$

Hence, after summing with respect of k, we arrive at the Bochner identity

$$\Delta\nabla\varphi = \nabla\Delta\varphi + \mathrm{Ric}\nabla\varphi, \tag{6}$$

where Ric : $\mathcal{O}(M) \longrightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ is the Ricci curvature (symmetric) matrix

$$\text{(7)} \qquad \mathrm{Ric}_{i,j} = -\sum_{k=1}^{d} \left(\Phi_{k,i}\mathbf{e}_k, \mathbf{e}_j\right)_{\mathbf{R}^d}.$$

Bochner's identity is the starting point for a great deal of analysis on M. To wit, let $\{P_t : t > 0\}$ denote the Markov semigroups on $C(M)$ determined by $\frac{1}{2}\Delta$. Then, as an application of (6), we find that

$$\text{(8)} \qquad \frac{d}{dt}\nabla P_t\varphi = \frac{1}{2}\Delta\nabla P_t\varphi - \frac{1}{2}\mathrm{Ric}\nabla P_t\varphi,$$

where the action of Δ on an $\mathbb{R}^d$-valued function is component by component; and, from (8), one has

$$\frac{d}{dt}\left|\nabla P_t\varphi\right|^2 = \left(\nabla P_t\varphi, \Delta\nabla P_t\varphi\right)_{\mathbf{R}^d} - \left(\nabla P_t\varphi, \mathrm{Ric}\nabla P_t\varphi\right)_{\mathbf{R}^d}.$$

At the same time, an easy computation leads to

$$\Delta\left|\nabla P_t\varphi\right|^2 = 2\left(\nabla P_t\varphi, \Delta\nabla P_t\varphi\right)_{\mathbf{R}^d} + 2\left\|\mathrm{Hess}(P_t\varphi)\right\|^2_{\mathrm{H.S.}},$$

where $\mathrm{Hess}f \equiv \big((\mathfrak{E}_k\mathfrak{E}_\ell f)\big)$ is the Hessian matrix of $f \in C^2(M)$ and $\|\cdot\|_{\mathrm{H.S.}}$ is the standard Hilbert–Schmidt norm for $d \times d$ matrices. Hence, we find that

$$\begin{aligned} \frac{d}{dt}\left|\nabla P_t\varphi\right|^2 &= \frac{1}{2}\Delta\left|\nabla P_t\varphi\right|^2 - \left\|\mathrm{Hess}(P_t\varphi)\right\|^2_{\mathrm{H.S.}} - \left(\nabla P_t\varphi, \mathrm{Ric}\nabla P_t\varphi\right)_{\mathbf{R}^d} \\ &\le \frac{1}{2}\Delta\left|P_t\varphi\right|^2 - \alpha\left|\nabla P_t\varphi\right|^2, \end{aligned}$$

where

$$\text{(9)} \qquad \alpha \equiv \inf\left\{\left(\mathbf{e}, \mathrm{Ric}(\mathfrak{f})\mathbf{e}\right)_{\mathbf{R}^d} : \mathfrak{f} \in \mathcal{O}(M) \text{ and } |\mathbf{e}| = 1\right\}.$$

In particular, for $T \in (0, \infty)$,

$$\frac{d}{dt}P_{T-t}\left(\left|\nabla P_t\varphi\right|^2\right) \le -\alpha P_{T-t}\left(\left|\nabla P_t\varphi\right|^2\right), \quad t \in (0, T),$$

and so

$$\text{(10)} \qquad \left|\nabla P_T\varphi\right|^2 \le e^{-\alpha T}P_T\left(|\nabla\varphi|^2\right), \quad T \in (0, \infty).$$

The estimate in (10) is very useful as it stands. For example, when $\alpha > 0$, it leads immediately to the well known fact that the spectral gap for Δ as an operator on $L^2(M)$ is at least α. However, as Bismut [B] noticed, (8) can be effectively combined with elementary probability theory to replace estimates like (10) with intriguing equalities. To see this, let $(\mathfrak{W}, \mathcal{B}_{\mathfrak{W}}, \mu)$ be the standard Wiener space of $\mathbb{R}^d$-valued paths and, for each $\mathfrak{f} \in \mathcal{O}(M)$, use $\mathfrak{F}_{\mathfrak{f}} : [0,\infty) \times \mathfrak{W} \longrightarrow \mathcal{O}(M)$ to denote the progressively measurable solution to the Stratonovich stochastic differential equation

$$d\mathfrak{F}_{\mathfrak{f}}(t,\mathbf{w}) = \sum_1^d \mathfrak{E}_k\big(\mathfrak{F}_{\mathfrak{f}}(t,\mathbf{w})\big) \circ d\mathbf{w}(t)_k \quad \text{with } \mathfrak{F}_{\mathfrak{f}}(0,\mathbf{w}) = \mathfrak{f}. \tag{11}$$

Next, define $A_{\mathfrak{f}} : [0,\infty) \times \mathfrak{W} \longrightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ by the integral equation

$$A_{\mathfrak{f}}(t,\mathbf{w}) = \mathbf{I} - \frac{1}{2}\int_0^t A_{\mathfrak{f}}(\tau,\mathbf{w})\mathrm{Ric}\big(\mathfrak{F}_{\mathfrak{f}}(\tau,\mathbf{w})\big)\,d\tau, \quad t \in [0,\infty). \tag{12}$$

Then, from (8) and Itô's formula, one finds that, for each $T \in (0,\infty)$,

$$M(t,\mathbf{w}) = A_{\mathfrak{f}}(t\wedge T,\mathbf{w})\big[\nabla P_{T-t\wedge T}\varphi\big]\big(\mathfrak{F}_{\mathfrak{f}}(t\wedge T,\mathbf{w})\big) \tag{13}$$

is an $\mathbb{R}^d$-valued martingale. In particular, this means that

$$\big[\nabla P_T\varphi\big](\mathfrak{f}) = \mathbb{E}\Big[A_{\mathfrak{f}}(T)\nabla\varphi\big(\mathfrak{F}_{\mathfrak{f}}(T)\big)\Big]. \tag{14}$$

Since it is obvious that (cf. (9))

$$\big\|A_{\mathfrak{f}}(T,\mathbf{w})\big\|_{\mathrm{op}}^2 \le e^{-\alpha T}, \quad (T,\mathbf{w}) \in [0,\infty)\times\mathfrak{W}, \tag{15}$$

(14) represents a considerable sharpening of (10). For example, from (14) and (15), we know that

$$\big|\nabla P_T\varphi\big| \le e^{-\frac{\alpha T}{2}} P_T\big(|\nabla\varphi|\big). \tag{16}$$

(Notice that although $\nabla\psi$ is defined only on $\mathcal{O}(M)$, $|\nabla\psi|$ is well-defined on M itself.) To see why (16) represents an improvement on (10), we follow the reasoning of D. Bakry and M. Emery [BM] to derive from it the logarithmic Sobolev inequality

$$P_T(\varphi\log\varphi) - P_T\varphi\log P_T\varphi \le \left(\frac{1}{2}\int_0^T e^{-\alpha t}\,dt\right) P_T\left(\frac{|\nabla\varphi|^2}{\varphi}\right), \quad \varphi \in C^1\big(M;(0,\infty)\big). \tag{17}$$

Indeed, note that,

$$\begin{aligned} 2\frac{d}{dt}P_t\big(P_{T-t}\varphi\log P_{T-t}\varphi\big) &= P_t\left(\frac{\big|\nabla P_{T-t}\varphi\big|^2}{P_{T-t}\varphi}\right) \\ &\le e^{\alpha(t-T)}P_t\left(\frac{\big(P_{T-t}|\nabla\varphi|\big)^2}{P_{T-t}\varphi}\right) \le e^{\alpha(t-T)}P_T\left(\frac{|\nabla\varphi|^2}{\varphi}\right), \end{aligned}$$