Proceedings of the International Conference on Stochastic Analysis and Partial Differential Equations, Northwestern University, 2005.

HEAT EQUATIONS ON RIEMANNIAN MANIFOLDS AND BISMUT'S FORMULA

ELTON P. HSU

CONTENTS

1. INTRODUCTION

Let M be a Riemannian manifold and $p_M(t,x,y)$ the (minimal) heat kernel on M. Bismut's formula ([2]) is a probabilistic representation of the covariant derivative of $p_M(t,x,y)$ of the following form:

$$T\nabla \log p_M(T,x,y) = \mathbb{E}_x \left[\int_0^T M_s \, dW_s \middle| X_T = y \right]. \tag{1.1}$$

Here the right side is the expected value of a functional of Brownian motion $\{X_t\}$ on M starting from x and conditioned to return to y at time T, the process $\{W_t\}$ is the (stochastic) anti-development of X, which by definition is a euclidean Brownian motion, and $\{M_s\}$ is the multiplicative Feynmann-Kac functional determined by the Ricci curvature of the manifold. The usefulness of such a formula can be

The research was supported in part by the NSF grant DMS 0407819.

seen more clearly from another equivalent form of the formula. Let $\{P_t\}$ be the heat semigroup defined by

$$P_t f(x) = \int_M p_M(t, x, y) f(y)\, dy.$$

Since the heat kernel $p_M(t, x, y)$ is also the transition density function of Brownian motion on M, we have also

$$P_t f(x) = \mathbb{E}_x f(X_t).$$

It is easy to verify that Bismut's formula (1.1) is equivalent to the following statement: for all reasonable smooth functions f on M,

$$T \nabla_x P_T f(x) = \mathbb{E}_x \left[f(X_T) \int_0^T M_s\, dW_s \right]. \tag{1.2}$$

The significance of this representation is that the right side does not contain the gradient of the function f. Suppose that f is harmonic on M, i.e., it satisfies the equation $\Delta_M f = 0$, where Δ_M is the Laplace-Beltrami operator on M. Suppose also that $P_T f = e^{\Delta_M T/2} f = f$ and that we are allowed to use (1.2) for this function f. Then the left side is simply $T\nabla f$; thus the gradient of a harmonic function is expressed explicitly in terms of the function itself, which potentially opens ways of using this formula to study gradient estimates of harmonic functions (see Thalmaier and Wang [9]). We can also see a possible use of Bismut's formula in financial mathematics, for in many cases, the price of a financial derivative takes the form $\mathbb{E}_x f(X_T)$, where f represents the payoff at the expiry (with the expiration time T) and $X_0 = x$ represents the initial prices of the underlying assets. Thus $\nabla_x \mathbb{E}_x f(X_T)$ are the rate of change (sensitivity) of the price of the derivative with respect to the initial prices, which are called the derivative Greeks in financial mathematics. Thus Bismut's formula gives a useful way of computing the Greeks in terms of the payoff function itself (see Nualart [7]).

Bismut's original approach to his formula (see Bismut [2]) is to calculate the variation of Brownian paths with respect to the starting point $X_0 = x$. If we use $X^x = \{X_t^x\}$ to denote a Brownian motion on M starting from x, then, as a first step to calculate the covariant derivative $\nabla_x \mathbb{E}_x f(X_T) = \nabla_x \mathbb{E} f(X_T^x)$, we need to calculate the variation of the terminal position X_T^x of the Brownian motion with respect to the initial point x. This is to be followed by an integration by parts in the path space to remove the differentiation from the function f. Nowadays this approach is largely abandoned in favor of a more streamlined approach based on heat equations and stochastic calculus on manifolds. The end form of this approach is the fruit of research by many mathematicians (including Elworthy, Stroock, Driver and their collaborators, and myself). The first part of this paper is to give an exposition of this approach to Bismut's formula. In the second part, we consider an extension of Bismut's formula to vector bundles. Norris [6] discussed such an extension by following the original variation method of Bismut. Bismut's formula for vector bundles can also be treated by heat equations and stochastic calculus on manifolds. We will illustrate this assertion by the simple case of a trivial bundle. The proof of Bismut's formula in this case, simple as it is, has already contained some good indications of what one might expect to do in a more general setting. A full version of this result will be a part of the forthcoming work [1].

2. LAPLACE-BELETRAMI OPERATOR AND BOCHNER'S HORIZONTAL LAPLACIAN

We start with the Laplace-Beltrami operator Δ_M, which is the infinitesimal generator for Brownian motion on M as a diffusion process. We assume that M is a Riemannian manifold equipped with the Levi-Civita connection and use ∇ to denote covariant differentiation on tensor fields on M. The Laplace-Beltrami operator, which generalizes the usual Laplace operator on euclidean space, is defined by

$$\Delta_M f = \mathrm{div}(\mathrm{grad} f),$$

where the gradient and the divergence are defined with respect to the Riemannian metric on M. The gradient $\mathrm{grad} f$ is the dual of the differential df; thus it is the unique vector field defined by the relation

$$\langle \mathrm{grad} f, X\rangle = df(X) = Xf, \qquad \forall X \in \Gamma(TM).$$

In local coordinates $x = \left\{x^i\right\}$, let $X_i = \partial/\partial x^i$ be the partial differentiations along the coordinate variables. The Riemannian metric can be written as

$$ds^2 = g_{ij}dx^i dx^j, \quad g_{ij} = \langle X_i, X_j\rangle.$$

In terms of local coordinates, the Riemannian metric is given by

$$\nabla f = g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j}.$$

The divergence $\mathrm{div} X$ of a vector field X is defined to be the contraction of the (1,1)-tensor ∇X. If $X = a^i \dfrac{\partial}{\partial x^i}$ in local coordinates, then it is easy to verify that

$$\operatorname{div} X = \frac{1}{\sqrt{G}} \frac{\partial\left(\sqrt{G}\, a^i\right)}{\partial x^i}.$$

Combining the local expressions the gradient and divergence, we obtain the familiar local formula for the Laplace-Beltrami operator:

$$\Delta_M f = \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i} \left(\sqrt{G} g^{ij} \frac{\partial f}{\partial x^j} \right).$$

Thus Δ_M is a nondegenerate second order elliptic operator on M. For our purpose, the following description is more important. For any orthonormal basis $\{X_i\}$ of $T_x M$, we have

$$\Delta_M f = \operatorname{trace} \nabla^2 f = \sum_{i=1}^{d} \nabla^2 f(X_i, X_i).$$

Let $\mathscr{O}(M)$ be the orthonormal frame bundle of M and $\pi : \mathscr{O}(M) \to M$ the canonical projection. Recall that the fundamental horizontal vector fields H_i (with respect to the Levi-Civita connection) are the unique horizontal vector fields on $\mathscr{O}(M)$ such that $\pi_* H_i(u) = ue_i$, where $\{e_i\}$ is the canonical basis for $\mathbb{R}^d$ (see Kobayashi and Nomizu [5]). Bochner's horizontal Laplacian is the second order elliptic operator on $\mathscr{O}(M)$ defined by

$$\Delta_{\mathscr{O}(M)} = \sum_{i=1}^{d} H_i^2.$$

Bochner's horizontal Laplacian $\Delta_{\mathscr{O}(M)}$ is the lift of the Laplace-Beltrami operator Δ_M to the orthonormal frame bundle $\mathscr{O}(M)$. More precisely, let $f \in C^\infty(M)$, and $\widetilde{f} = f \circ \pi$ its lift to $\mathscr{O}(M)$. Then for any $u \in \mathscr{O}(M)$,

$$\Delta_M f(x) = \Delta_{\mathscr{O}(M)} \widetilde{f}(u), \tag{2.1}$$

where $x = \pi u$. The obvious advantage of $\Delta_{\mathscr{O}(M)}$ over Δ_M is that it is (intrinsically) in the form of the sum of $n = \dim M$ canonically defined vector fields. As a consequence, the diffusion process generated by $\Delta_{\mathscr{O}(M)}/2$ is the solution of a standard Stratonovich stochastic differential equation on $\mathscr{O}(M)$. This allows us to take the full advantage of stochastic calculus (stochastic differential equations) on manifolds. The price to be paid is that we need to work in the orthonormal frame bundle $\mathscr{O}(M)$, a much larger space than the manifold M itself.

3. Brownian motion on a Riemannian manifold

Let M be a Riemannian manifold. The heat kernel $p_M(t,x,y)$ is the (minimal) fundamental solution of the parabolic operator

$$L = \frac{\partial}{\partial t} - \frac{1}{2}\Delta_M,$$

where Δ_M is the Laplace-Beltrami operator on M. The heat kernel can be constructed geometrically by the method of parametrix starting from an approximate heat kernel in local coordinates (see Chavel [3]). The heat kernel is also the transition density function of Brownian motion on M, i.e.,

$$\mathbb{P}_x\{X_t \in A\} = \int_A p_M(t,x,y)\,dy.$$

Once $p_M(t,x,y)$ is given, the process X can be constructed by the standard method in the theory of Markov processes (see Chung [4]). This construction of Brownian motion on a Riemannian manifold, although simple and direct, will not show explicitly how the geometry of the manifold affects the behavior of Brownian motion.

We now briefly discuss the well-known Eells-Elworthy-Malliavin construction of Brownian motion on a Riemannian manifold. Instead of starting from the heat kernel $p_M(t,x,y)$, which shows the averaged behavior of the process to be constructed, this approach goes directly to the path level. The approach is geometric and the heat kernel (or the transition density function) is obtained as a product rather than a building block of the construction. The extra flexibility we gain from the manifold to the path space over the manifold will be crucial to making effective use Itô's formula for Brownian motion on a Riemannian manifold.

Let $\{W_t\}$ be a standard Brownian motion on $\mathbb{R}^d$ (starting from zero) and consider the following Stratonovich stochastic differential equation on $\mathscr{O}(M)$:

$$dU_t = H_i(U_t) \circ dW_t.$$

By general theory of stochastic differential equations, it has a unique solution, which is a diffusion process generated by Bochner's horizontal Laplacian $\Delta_{\mathscr{O}(M)}/2$. This process is called a horizontal Brownian motion. Using the relation of $\Delta_{\mathscr{O}(M)}$ and Δ_M in (2.1) it is easy to verify that the projection $X_t = \pi U_t$ is a diffusion process generated by $\Delta_M/2$, namely, a Brownian motion on M. One immediately consequence of this statement is that for any $f \in C^\infty(M)$, the function

$$f(x,t) = \mathbb{E}_x f(X_t)$$

is the solution of the initial value problem

$$\frac{\partial f}{\partial t} = \frac{1}{2}\Delta_M f, \qquad f(x,0) = f(x). \tag{3.1}$$

Equivalently, $f(x,t) = P_t f(x)$, where $P_t = e^{\Delta_M t/2}$ is the heat kernel semigroup, and

$$\mathbb{E}_x f(X_t) = P_t f(x) = \int_M p_M(t,x,y) f(y)\, dy.$$

The above identity can be regarded as a probabilistic identification of the heat kernel.

4. Hodge-de Rham Laplacian and Weitzenböck formula

Recall that the Laplace-Beltrami operator Δ_M is the trace of the Hessian $\nabla^2 f$:

$$\Delta_M f(x) = \sum_{i=1}^d \nabla^2 f(X_i, X_i),$$

where $\{X_i\}$ is any orthonormal basis of $T_x M$. The Laplace-Beltrami operator Δ_M on functions can be extended to tensor fields on M by the same relation:

$$\Delta_M \theta = \sum_{i=1}^d \nabla^2 \theta(X_i, X_i).$$

This is the so-called rough Laplacian on tensor fields. We will concentrate on the case where θ is a 1-form, or a vector field (by duality).

A 1-form θ on M can be lifted to its scalarization $\tilde\theta$ on the orthonormal frame bundle $\mathscr{O}(M)$ defined by

$$\tilde\theta(u) = u^{-1}\theta(\pi u).$$

Thus $\tilde\theta$ is an $\mathbb{R}^d$-valued function on $\mathscr{O}(M)$ and is $O(d)$-invariant in the sense that $\tilde\theta(ug) = g\tilde\theta(u)$ for $g \in O(d)$. The horizontal derivatives $H_i\tilde\theta$ and Bochner's horizontal Laplacian $\Delta_{\mathscr{O}(M)}\tilde\theta$ are well defined, and we have a relation for 1-forms similar to (2.1):

$$\Delta_{\mathscr{O}(M)}\tilde\theta(u) = u^{-1}\Delta_M\theta(x), \quad \pi u = x. \tag{4.1}$$

We now turn to the Hodge-de Rham Laplacian on 1-forms. Although the covariant (rough) Laplacian Δ_M is naturally associated with Brownian motion on a manifold, it does not commute with the exterior differentiation, a natural operation on differential forms. Geometrically more significant is the Hodge-de Rham Laplacian

$$\Box_M = -(d\delta + \delta d).$$

Here δ is the (formal) adjoint of the exterior differentiation d with respect to the canonical inner product on the space of differential forms on M:

$$(d\alpha, \beta) = (\alpha, \delta\beta).$$

Note that with the sign convention we have chosen $\square_M$ coincides with Δ_M on functions. Using the fact that $d^2 = 0$ (and hence also $\delta^2 = 0$) we verify easily that $d\,\square_M = \square_M d$. Thus we have two natural Laplacians on the space of differential forms: $\square_M$ is associated closely with geometry because it commutes with d (and hence also ∇), and Δ_M is the direct descendant from Bochner's Laplacian $\Delta_{\mathscr{O}(M)}$, which is the generator of horizontal Brownian motion. The difference between $\square_M$ and Δ_M on differential forms is given by the Weitzenböck formula (see de Rham [8]). It takes particular simple form in the case of 1-forms:

$$\square_M\theta = \Delta_M - \mathrm{Ric}_M\theta,$$

where $\mathrm{Ric}_M : T_xM \to T_xM$ is the Ricci curvature transform. Note that the last term is a matrix multiplication (zeroth order operator), which opens a way of using Feynman-Kac functionals to pass from Δ_M to $\square_M$.

For our purpose it is more convenient to lift the Weitzenböck formula from the manifold M to the orthonormal frame bundle $\mathscr{O}(M)$. Let

$$\mathrm{Ric}_u := u^{-1}\mathrm{Ric}_{\pi u}u : \mathbb{R}^d \to \mathbb{R}^d$$

be the scalarized Ricci transform at a frame $u \in \mathscr{O}(M)$. Let

$$\square_{\mathscr{O}(M)} = \Delta_{\mathscr{O}(M)} - \mathrm{Ric}.$$

Then $\square_{\mathscr{O}(M)}$ is a lift of the Hodge-de Rham Laplacian in the sense that

$$\square_{\mathscr{O}(M)}\tilde{\theta}(u) = u^{-1}\square_M\theta(x), \quad \pi u = x.$$

5. Heat equations for functions and 1-forms

We now consider the following initial-value problem for a 1-form $\theta = \theta(t,x)$:

$$\begin{cases} \dfrac{\partial\theta}{\partial t} = \dfrac{1}{2}\square_M\theta, & (t,x) \in (0,\infty) \times M; \\[2ex] \theta(0,x) = \theta(x), & x \in M. \end{cases} \tag{5.1}$$

We can rewrite the above equation on $\mathscr{O}(M)$. Let $\tilde{\theta}$ be the scalarization of θ. Then the above equation is equivalent to

$$\begin{cases} \dfrac{\partial \tilde{\theta}}{\partial t} = \dfrac{1}{2}\Box_{\mathscr{O}(M)}\tilde{\theta}, & (t,u) \in (0,\infty) \times \mathscr{O}(M); \\ \\ \tilde{\theta}(0,u) = \tilde{\theta}_0(u), & u \in \mathscr{O}(M). \end{cases} \tag{5.2}$$

Recall that the horizontal Brownian motion is generated by Bochner's Laplacian

$$\Delta_{\mathscr{O}(M)} = \sum_{i=1}^n H_i^2.$$

Let $\widetilde{\theta}$ be the scalarization of θ and

$$\widetilde{\theta}(t,u) = \mathbb{E}_u \widetilde{\theta}(U_t).$$

Then it satisfies the heat equation

$$\frac{\partial \widetilde{\theta}}{\partial t} = \frac{1}{2}\Delta_{\mathscr{O}(M)}\widetilde{\theta}.$$

Furthermore, $\widetilde{\theta}(t,u)$ is the scalarization of a 1-form $\theta(t,x)$ on M, which satisfies the heat equation

$$\frac{\partial \theta}{\partial t} = \frac{1}{2}\Delta_M \theta.$$

However, what we want is the solution of the heat equation with the covariant (rough) Laplacian Δ_M replaced by the Hodge-de Rham Laplacian $\Box_M$. By the Weitzenböck formula

$$\Box_{\mathscr{O}(M)} = \Delta_{\mathscr{O}(M)} - \mathrm{Ric}_M.$$

The solution of the heat equation (5.1) can be obtained by using a matrix version of the well-known Feynman-Kac formula. Let M_t be the matrix-valued multiplicative functional determined by

$$\frac{dM_t}{dt} + \frac{1}{2}M_t \widetilde{\mathrm{Ric}}_{U_t} = 0, \qquad M_0 = I_n. \tag{5.3}$$

Then an easy application of Itô's formula gives

$$d\left\{M_s\tilde{\theta}(t-s,U_s)\right\} = M_s \sum_{i=1}^n H_i \widetilde{\theta}(t-s,U_s)\, dW_s^i,$$

which shows that $\left\{M_s\tilde{\theta}(t-s,U_s), 0 \le s \le t\right\}$ is a martingale. Equating the expected values at $s=0$ and $s=t$, we obtain a probabilistic representation of the solution of the heat equation in the form

$$\tilde{\theta}(t,u) = \mathbb{E}_u\left\{M_t\tilde{\theta}(U_t)\right\}.$$

Correspondingly, the solution of (5.1) is given by

$$\theta(t,x) = \mathbb{E}_x \left\{ M_t U_t^{-1} \theta(X_t) \right\},$$

where U is the horizontal lift of a Brownian motion X.

6. Bismut's formula for the heat kernel on functions

With the preparations of the previous two sections, the proof of Bismut's formula for the functional heat kernel becomes plain sailing. Let

$$f(x,t) = \mathbb{E}_x f(X_t) = P_t f(x).$$

It is the solution of the initial value problem

$$\frac{\partial f}{\partial t} = \frac{1}{2}\Delta_M f, \qquad f(x,0) = f(x).$$

We now use the same notation $P_t = e^{\Box_M t/2}$ to denote the heat semigroup (acting on functions and 1-forms) generated by the Hodge-de Rham Laplacian $\Box_M$. Let

$$\theta(x,t) = \nabla f(x,t) = \nabla \mathbb{E}_x f(X_t).$$

Since the exterior differentiation d (hence also the gradient ∇ by the canonical identification) commutes with $\Box_M$, it also commutes with the semigroup P_t and we have

$$\theta(x,t) = \nabla P_t f(x) = P_t(\nabla f)(x).$$

This equality simply means that $\theta(x,t)$ solves the equation

$$\frac{\partial \theta}{\partial t} = \frac{1}{2}\Box_M \theta.$$

On the other hand, the discussion in the last section shows that the process $M_t \widetilde{\theta}(U_t, T-t)$ is a martingale, hence for any $0 \le t \le T$,

$$\nabla \mathbb{E}_x f(X_T) = \mathbb{E}_u \left[M_t \widetilde{\theta}(U_t, T-t) \right].$$

Integrating with respect to t from 0 to T, we obtain

$$\begin{aligned} T\nabla \mathbb{E}_x f(X_T) &= \mathbb{E}_u \left[\int_0^T M_t \widetilde{\theta}(U_t, T-t)\, dt \right] \\ &= \mathbb{E}_u \left[\int_0^T M_t\, dW_t \int_0^T \langle \widetilde{\theta}(U_t, T-t), dW_t \rangle \right]. \end{aligned}$$

Finally, applying Itô's formula to $f(X_t, T-t) = \widetilde{f}(U_t, T-t)$ and using the fact that $f(x,t)$ solves the heat equation we obtain immediately that

$$f(X_T) - \mathbb{E}_x f(X_T) = \int_0^T \langle \widetilde{\theta}(U_t, T-t), dW_t \rangle.$$

It follows that

$$\nabla \mathbb{E}_x f(X_T) = \mathbb{E}_x \left[f(X_T) \int_0^T M_s \, dW_s \right].$$

This is exactly Bismut's formula we wanted to prove.

7. Extension of Bismut's formula to vector bundles

The same line of thought can be applied to prove extensions of Bismut's formula for heat kernels for vector bundles. In this section we discuss the simplest case of a trivial line bundle $F = \mathbb{R}^2$ over $M = \mathbb{R}$. It illustrates some extra steps one must take beyond what have already appeared in the functional case.

The trivial bundle $F = \mathbb{R}^2$ is covered by a coordinate system of two variables $z = (x, y)$, where x is the base variable and y the fibre variable. We will take the base motion to be a standard one dimensional Brownian motion $X_t = W_t$. The fibre motion typically should be driven by both W and t, but we will take the simple case where the fibre motion is determined by an ordinary differential equation

$$dY_t = V(X_t, Y_t) \, dt.$$

The total process is given by $Z_t = (X_t, Y_t)$.

Now consider a smooth function $f : F = \mathbb{R}^2 \to \mathbb{R}$ and $f(z,t) = \mathbb{E}_z f(Z_t)$. What we need is a probabilistic representation of $\nabla_x f(z, T)$.

The generator for the diffusion process $\{Z_t\}$ is easily identified:

$$L = \frac{1}{2} \partial_x^2 + V(x,y) \partial_y.$$

Hence $f(z,t)$ is the solution of

$$\frac{\partial f}{\partial t} = Lf. \tag{7.1}$$

This means that the process $f(Z_t, T-t)$ is a martingale and

$$f(Z_T) = \mathbb{E}_z f(Z_T) + \int_0^T f_x(Z_t, T-t) \, dW_t. \tag{7.2}$$

This identity will be needed later.

Next, differentiating (7.1) with respect to x we see that derivative $f_x = \partial f(x,y,t)/\partial x$ solves satisfies

$$\frac{\partial f_x}{\partial t} = Lf_x + [\partial_x, L]f.$$

The commutator $[\partial_x, L]$ can be computed easily:

$$[\partial_x, L]f = V_x(x,y)f_y.$$

Therefore the equation for f_x is

$$\frac{\partial f_x}{\partial t} = Lf_x + V_x f_y.$$

Here we see that the fibre derivative $f_y = \partial f(x,y,t)/\partial y$ appears on the right side. This means that we cannot directly use the Feynman-Kac technique to deal with this equation. However, a probabilistic representation of the solution is still possible. In fact, we can verify by Itô's formula that

$$f_x(Z_t, T-t) + \int_0^t V_x(Z_s)f_y(Z_s, T-s)\,ds, \quad 0 \le t \le T,$$

is a martingale, hence for $0 \le t \le T$,

$$f_x(z,T) = \mathbb{E}_z f_x(Z_t, T-t) + \int_0^t V_x(Z_s)f_y(Z_s, T-s)\,ds.$$

Integrating with respect to t, we have

$$\begin{aligned} T\nabla_x \mathbb{E}_z f(Z_T) &= \mathbb{E}_z\left[\int_0^T f_x(Z_t, T-t)\,dt\right] \\ &\quad + \mathbb{E}_z \int_0^T V_x(Z_t)f_y(Z_t, T-t)(T-t)\,dt. \end{aligned} \tag{7.3}$$

The first term can be treated in the same way as we have done in the last section, namely, we write it as

$$\begin{aligned} \mathbb{E}_z\left[\int_0^T f_x(Z_t, T-t)\,dt\right] &= \mathbb{E}_z\left[\int_0^T f_x(Z_t, T-t)\,dW_t \int_0^T I\,dW_t\right] \\ &= \mathbb{E}_z\left[f(Z_T)\int_0^T I\,dW_t\right]. \end{aligned}$$

Here we have used (7.2) in the second step. The question now is how to deal with the second term in (7.3) involving $f_y(Z_t, T-t)$. It turns out that this term can be treated by the Feynman-Kac technique. To see this, we first derive the heat equation for $f_y(z,t)$. Differentiating the equation (7.1) for $f(z,t)$ with respect to y we have

$$\frac{\partial f_y}{\partial t} = Lf_y + [\partial_y, L]f.$$

Recall that $L = (1/2)\partial_x^2 + V\partial_y$, hence $[\partial_y, L]f = V_y(z)f_y$ and

$$\frac{\partial f_y}{\partial t} = Lf_y + V_y(z)f_y.$$

This is an equation satisfied by f_y itself, not involving f_x. Introduce the Feynman-Kac functional

$$e_t = \exp\left[\int_0^t V_y(Z_s)\,ds\right].$$

Then a straightforward application of Itô's formula shows that the process

$$N_t = f_y(Z_t, T-t)\,e_t$$

is a martingale. Let

$$g_t = \int_0^t e_s^{-1} V_x(Z_s)(T-s)\,ds.$$

Then the second term on the right side of (7.3) involving f_y becomes

$$\begin{aligned}
\int_0^T V_x(Z_t)f_y(z_t, T-t)(T-t)\,dt \\
&= \int_0^T N_t dg_t \\
&= N_T g_T - N_0 g_0 - \int_0^T g_t\,dN_t.
\end{aligned}$$

The expected value of the last term vanishes because it is a martingale. We also have $N_T = f_y(Z_T)e_T$ and $g_0 = 0$, we have

$$\mathbb{E}_z \int_0^T V_x(Z_t)f_y(Z_t, T-t)(T-t)\,dt = \mathbb{E}_z\left[f_y(Z_T)Y_T\right],$$

where

$$Y_t = e_t \int_0^t e_s^{-1} V_x(Z_s)(T-s)\,ds.$$

Putting things together, we obtain a Bismut's formula (for the trivial bundle) in the following form

$$T\nabla_x \mathbb{E}_z f(z_T) = \mathbb{E}_x\left[f(Z_T)\int_0^T I\,dW_t + f_y(Z_T)Y_T\right].$$

Note that Y_t is Norris' derived process (see Norris [6]) satisfying the equation

$$dY_t = V_y(Z_t)Y_t\,dt + V_x(Z_t)(T-t)\,dt.$$

In a more general setting, the derived equation is a stochastic differential equation which in general cannot be solved explicitly.

References

[1] Arnaudon, M., Hsu, E.P., Qian, Z.M., and Thalmaier, A., Bismut's formula for vector bundles and applications (2006).

[2] Bismut, J-M., *Large Deviation and the Malliavin Calculus*, Birkhäuser (1984).

[3] Chavel, I., *Eigenvalues in Riemannian Geometry*, Academic Press.

[4] Chung, K. L. and Walsh, J. B., *Markov Processes, Brownian Motion, and Time Symmetry*, Springer, Berlin (2004).

[5] Kobayashi, S. and Nomizu, K., *Foundations of Differential Geometry*, volume one, Interscience Publishers (1963).

[6] Norris, J., Path integral formulae for heat kernels and their derivatives, *Probability Theory and Related Fields*, **94**, 524–541 (1993).

[7] Nualart, D., *The Malliavin Calculus and Related Topics*, 2nd edition, Springer, Berlin (2006).

[8] de Rham, G., *Riemannian Manifolds*, Springer, New York (1984).

[9] Thalmaier, Anton and Wang, F.-Y., Gradient estimates for harmonic functions on regular domains in Riemannian manifolds, *Journal of Functional Analysis*, **155**, 109-124 (1998).

Department of Mathematics, Northwestern University, Evanston, IL 60208

E-mail address: `elton@math.northwestern.edu`