Differentiation of heat semigroups and applications

K. D. Elworthy and X.-M. Li *

Mathematics Institute

University of Warwick

Coventry CV4 7AL,U.K.

1 Introduction

Consider the Stratonovich stochastic differential equation

$$dx_t = X(x_t) \circ dB_t + A(x_t)dt. \tag{1}$$

on $\mathbb{R}^n$ with coefficients $A : \mathbb{R}^n \to \mathbb{R}^n$ a smooth vector field and $X : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ a smooth map into the space of linear maps of $\mathbb{R}^m$ into $\mathbb{R}^n$, driven by the white noise determined by a Brownian motion $\{B_t : t \geq 0\}$ on $\mathbb{R}^m$. It can also be written

$$dx_t = \sum_1^m X^i(x_t) \circ dB_t^i + A(x_t)dt \tag{2}$$

*Research supported by SERC grant GR/H67263.

where $B_t = (B_t^1, \ldots, B_t^m)$ and $X^1, \ldots, X^m$ are the vector fields given by $X^i(x) = X(x)(e_i)$ for $e_1, \ldots, e_m$ the standard basis for $\mathbb{R}^m$. Let $\{F_t(x_0) : t \geq 0\}$ be the solution to (1) starting from $x_0 \in \mathbb{R}^n$: we will assume non-explosion, so a solution exists for all time. Then there is the associated Markov semigroup $\{P_t : t \geq 0\}$ on bounded measurable functions $f : \mathbb{R}^n \to \mathbb{R}$ and differential generator $\mathcal{A}$ given by

$$\mathcal{A}(f) = \sum_1^m Ł_{X^i} Ł_{X^i}(f) + Ł_A(f)$$

for C^2 maps f, where $Ł_A$ denotes Lie differentiation with respect to the vector field A, i.e. $Ł_A(f) : \mathbb{R}^n \to \mathbb{R}$ is the derivative of f in the direction $A(x)$ at each x in $\mathbb{R}^m$:

$$Ł_A(f)(x) = Df(x)\,(A(x))\,.$$

When (1) is non-degenerate, or equivalently when $\mathcal{A}$ is elliptic, it is well known that P_t is smoothing: it maps bounded measurable maps to C^∞ maps for each $t > 0$. This is a classical result from p.d.e. theory, or could be proved by Malliavin calculus. Indeed in [Bis84], though for Brownian motion on a compact manifold, Bismut obtained a formula for the derivative of $P_t f$ in terms of f, and not involving the derivative of f; thus clearly demonstrating the smoothing properties in this case. This then led to his celebrated formula for the logarithmic derivative of the heat kernel, $\nabla_x \log p_t(x, y)$, see also [Nor93].

It turns out that an approach used by Elliot & Kohlmann [EK89] to prove Malliavin's integration by parts formula can also be used to give such Bismut type formulae in a variety of situations and with elementary proofs (and statements). Here we will prove the basic formula on $\mathbb{R}^n$, taken from

[EL94] and various variations and applications which have been made of it: the details of these will appear elsewhere. In particular we look at

(a). state space $\mathbb{R}^n$, global Lipschitz coefficients [EL94],

(a)'. Variations: (i) higher derivatives of $P_t f$ (ii) the semigroup given by $\mathcal{A} + V$ where V is a potential, and the coefficients are time dependent [EL94]

(b). infinite dimensional state space, stochastic evolution equations [DPEZ95], [PZ93]

(c). non-linear K.P.P. equations [LZ]

(d). state space a compact manifold M [Elw92]

(e). more general coefficients on $\mathbb{R}^n$ or non-compact M [EL94]

(f). Sobolev estimates [EL94]

(g). heat equations for differential forms [Elw88] [Li92] [EL94].

There is also a version in preparation for the non-linear heat equations for maps between Riemannian manifolds, but we do not describe that here.

2 The basic formula on R^n

A. Assume our state space is $\mathbb{R}^n$ and consider now equation (1) written as the Itô stochastic differential equation with coefficients assumed to be globally

Lipschitz:

$$dx_t = X(x_t)dB_t + Z(x_t)dt. \tag{3}$$

Let $\{x_t \equiv F_t(x_0) : t \geq 0\}$ be the solution to (3) starting at $x_0 \in \mathbb{R}^n$. From [BF61] we can take a version of this to give a solution flow jointly continuous in space and time, and C^∞ in space for each time t.

Suppose there is a bounded C^∞ map $Y : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that $Y(x)$ is the right inverse of $X(x)$ for each $x \in \mathbb{R}^n$: $X(x)Y(x) = \mathbf{1}_{\mathbb{R}^m}$. Boundedness of Y is a uniform ellipticity condition on $\mathcal{A}$.

Let BC^k be the space of C^k functions on $\mathbb{R}^n$ with bounded derivatives of order 1 to k, and let v_t be the solution to

$$dv_t = DX(x_t)(v_t)dB_t + DZ(x_t)(v_t)dt \tag{4}$$

with initial value v_0. Then $v_t = DF_t(x_0)(v_0)$, the derivative of the flow at x_0 in the direction v_0.

Let δP_t be the semigroup on bounded measurable maps $\theta : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n; \mathbb{R})$ given by

$$\delta P_t(\theta)(x_0)(v_0) = \mathbb{E}\theta(x_t)(v_t).$$

The basic formula is the following [EL94]: *If f is a bounded BC^2 function on $\mathbb{R}^n$, then the derivative of $P_t f$ is given by*

$$D(P_t f)(x_0)(v_0) = \frac{1}{t}\mathbb{E}f(x_t)\int_0^t < Y(x_s)(v_s), dB_s >_{\mathbb{R}^m} . \tag{5}$$

Proof:

Let $T > 0$. Apply Itô's formula to $(t, x) \mapsto P_{T-t}f(x), 0 \leq t \leq T$, which is sufficiently smooth by parabolic regularity (or if X, A are BC^2 by direct differentiation of the probabilistic expression for $P_t f$):

$$P_{T-t}f(x_t) = P_T f(x_0) + \int_0^t D(P_{T-s}f)(x_s)(X(x_s)dB_s). \tag{6}$$

for $t \in [0,T)$. Letting $t \to T$,

$$f(x_T) = P_T f(x_0) + \int_0^T D(P_{T-s}f)(x_s)(X(x_s)dB_s).$$

Multiplying the above equation by $\int_0^T < Y(x_s)(v_s), dB_s >$ and taking expectations, we get:

$$\begin{aligned} & \mathbb{E} f(x_T) \int_0^T < Y(x_s)v_s, dB_s >= \mathbb{E} \int_0^T D(P_{T-s}f)(x_s)(v_s)ds \\ = \ & \mathbb{E} \int_0^T ((\delta P_{T-s})(Df))\,(x_s)(v_s)ds = \int_0^T ((\delta P_s)\,((\delta P_{T-s})(Df)))\,(x_0)(v_0)ds \\ = \ & \int_0^T (\delta P_T(Df))(x_0)(v_0)ds = T\,\delta P_T(Df)(x_0)(v_0) = T\,D(P_t f)(x_0)(v_0). \end{aligned}$$

Divide both sides by T to finish the proof. ■

3 Variations

A. Higher derivatives [EL94]: Similarly there is a formula for the second derivative of $P_t f$: *Suppose X and A are in BC^2. Let f be a bounded function in BC^2. For u_0, v_0 in $\mathbb{R}^n$, let $u_s = DF_s(x_0)(u_0)$ and $v_s = DF_s(x_0)(v_0)$. Then*

$$\begin{aligned} & D^2 P_t f(x_0)(u_0, v_0) \\ & = \tfrac{4}{t^2}\mathbb{E}\left\{ f(x_t) \int_{\frac{t}{2}}^t < Y(x_s)v_s, dB_s > \int_0^{\frac{t}{2}} < Y(x_s)u_s, dB_s > \right\} \\ & + \tfrac{2}{t}\mathbb{E}\left\{ f(x_t) \int_0^{\frac{t}{2}} < DY(x_s)(u_s)(v_s), dB_s > \right\} \\ & + \tfrac{2}{t}\mathbb{E}\left\{ f(x_t) \int_0^{\frac{t}{2}} < Y(x_s)D^2F_s(x_0)(u_0,v_0), dB_s > \right\} \end{aligned} \tag{7}$$

if DY is bounded.

B. Time dependent coefficients and with a potential term: Suppose the coefficients of (3) are time-dependent, written as X_t and Z_t. Suppose X_t and Z_t are globally Lipschitz for each t and assume their spatial derivatives are continuous in t. Denote by Y_t the right inverse of X_t, assumed bounded uniformly on each $[0,T]$. Let $V_\cdot(\cdot)\colon [0,\infty)\times\mathbb{R}^n \to \mathbb{R}$ be continuous, C^1 in x and bounded above with derivative bounded on each $[0,T]\times\mathbb{R}^n$, and $\mathcal{A}_t$ the related elliptic operator: for $f\in C^2$,

$$\begin{aligned}\mathcal{A}_t(f)(x) &= \frac{1}{2}\text{trace } D^2f(x)\left(X_t(x)(-),X_t(x)(-)\right)\\ &+ Df(x)\left(Z_t(x)\right)+V_t(x)f(x).\end{aligned}$$

Consider

$$\frac{\partial f_t(x)}{\partial t} = \mathcal{A}_t f_t(x), \quad t>0 \tag{8}$$

Let f_0 be a bounded measurable function, then any bounded $C^{2,1}$ solution to (8), i.e. a solution which is C^2 in x and C^1 in t, is given by:

$$f_t(x_0) = \mathbb{E} f_0\left(x_t^t\right) \mathrm{e}^{\int_0^t V_{t-s}\left(x_s^t\right)ds} \tag{9}$$

for $\{x_t^T\}$ the solution of

$$dx_t^T = X_{T-t}(x_t^T)dB_t + Z_{T-t}(x_t^T)dt$$

starting from x_0. See e.g. [Fre85]. Let

$$\alpha_t^T(x_0) = \mathrm{e}^{\int_0^t V_{T-s}\left(x_s^T\right)ds},$$

and let $\{v_t^T \equiv TF_t^T(v_0) : t\geq 0\}$ be the solution to

$$dv_t^T = DX_{T-t}(x_t^T)(v_t^T)\,dB_t + DZ_{T-t}(x_t^T)(v_t^T)dt$$

with $v_0^T = v_0$. Then *[EL94] if f_0 is a bounded function in BC^2,*

$$\begin{aligned} Df_t(x_0)(v_0) = \;& \tfrac{1}{t}\mathbb{E} f_0(x_t^t) \mathrm{e}^{\int_0^t V_{t-s}(x_s^t)ds} \int_0^t \langle Y_{t-s}(x_s^t)(v_s^t), dB_s\rangle \\ & + \tfrac{1}{t}\mathbb{E} f_0(x_t^t) \mathrm{e}^{\int_0^t V_{t-s}(x_s^t)ds} \int_0^t (t-r) DV_{t-r}(x_r^t)(v_r^t) dr. \end{aligned} \tag{10}$$

C. Heat kernels The formulae are equivalent to formulae for derivatives of the relevant heat kernels (transition probabilities). For this consider the (most complicated) case of the time dependent equation (8) with time dependent potential $\{V_t : t \geq 0\}$. There is then a heat kernel $k(s, x; t, y)$ such that the solution $\{f_t : t \geq 0\}$ to (8) is given by

$$f_t(x_0) = \int_{\mathbb{R}^n} k(0, x_0; t, y) f_0(y) dy$$

for $x_0 \in \mathbb{R}^n$. Then, differentiating in the direction v_0,

$$Df_t(x_0)(v_0) = \int_{\mathbb{R}^n} Dk(0, -; t, y)(x_0)(v_0) f_0(y) dy.$$

On the other hand let $p_s^t(x_0, -)$ be the density of x_s^t. Thus, by (10) *for all y in $\mathbb{R}^n$*

$$\begin{aligned} & Dk(0, -; t, y)(x_0)(v_0) \\ = \;& p_t^t(x_0, y)\mathbb{E}\left\{ \mathrm{e}^{\int_0^t V_{t-s}(x_s^t)ds} \frac{1}{t} \int_0^t Y_{t-s}(x_s^t)(v_s^t), dB_s > \mid x_t^t = y \right\} \\ & + \mathbb{E}\left\{ \mathrm{e}^{\int_0^t V_{t-s}(x_s^t)ds} \frac{1}{t} \int_0^t (t-r) DV_{t-r}(x_r^t)(v_r^t) dr \mid x_t^t = y \right\} \end{aligned}$$

In fact the equation holds for all y by continuity of both sides in y [Bis84] [Nor93] [Wat84]. For the time independent case with $V \equiv 0$ we have $k(0, x; t, y) = p_t^t(x, y)$ giving *the logarithmic gradient formula:*

$$\nabla \log k(0, -; t, y)(x_0) = \frac{1}{t}\mathbb{E}\left\{ \int_0^t (TF_s)^* X(x_s) dB_s \mid x_t = y \right\}. \tag{11}$$

4 Infinite dimensional state space: stochastic evolution equations

The results described here are taken from work by Da Prato, Elworthy, & Zabczyk [DPEZ95] and best illustrated by an example. For this consider the stochastic partial differential equation

$$du_t = \left(\frac{1}{2}\Delta u_t - u_t^{2k+1}\right) dt + dW_t \tag{12}$$

some natural number k, for $u_t : [0,\pi] \to \mathbb{R}$ with Dirichlet boundary conditions. Here d refers to the Itô differential (in time). We let the driving noise $\{W_t : t \geq 0\}$ be a cylindrical Wiener process on $L^2([0,\pi])$ and consider (12) in the mild sense [DPZ92b]. Take $H = L^2([0,\pi])$ as state space. Then a Markov semigroup P_t is induced on the space of bounded measurable maps $f : H \to \mathbb{R}$. With the use of Yosida approximations it is shown in [DPEZ95] that *if $f : H \to \mathbb{R}$ is bounded and measurable then $P_t f$ is Lipschitz on H and when restricted to the space of continuous functions $C([0,\pi])$ it is Frechét differentiable with derivative at u_0 in the direction v_0 given by*

$$DP_t f(u_0)(v_0) = \frac{1}{t}\mathbb{E} f(u_t) \int_0^t < v_s, dW_s >_H$$

for $u_0, v_0 \in C([0,\pi])$ where $\{v_t : t \geq 0\}$ satisfies the heat equation with random potential

$$\frac{\partial v_t}{\partial t} = \frac{1}{2}\Delta v_t - (2k+1)u_t^{2k} v_t \quad t > 0$$

on $(0,\pi)$ with Dirichlet boundary conditions.

In particular this shows that P_t maps bounded measurable functions to continuous functions, if $t > 0$; the 'strong Feller' property. This property is particular important since [Kha60], given irreducibility, it implies that the transition probabilities are mutually absolutely continuous and that there is at most one invariant probability measure, see [DPZ92a].

For the precise class of stochastic differential equations for which such a result holds we refer to [DPEZ95]. The basic point is that the non-linearity can be very irregular (e.g. $u \mapsto -u^{2k+1}$ is not defined on all of H) provided there is dissipativity. The case of Lipschitz coefficients, including a non-degenerate diffusion coefficient based on the method of §1 is considered in [PZ93].

Just as in finite dimensions, given an invariant measure, the existence of the formula will often show smoothing in the sense of mapping L^2 functions into some Sobolev space $L^{2,1}$, see section 6 below. There are now results on compactness of the inclusion of such $L^{2,1}$ into L^2 at least for Gaussian measures [Pes93], [DPMN92], and so it seems that there are many cases for which the infinite dimensional Markov semigroup consists of compact operators on L^2, [DPZ93].

5 Stochastic differential equations on manifolds

Any elliptic generator $\mathcal{A}$ coming from an s.d.e. (1) on $\mathbb{R}^n$, or more generally on a smooth manifold, can be written $\mathcal{A} = \frac{1}{2}\Delta + Z$ where Z is a vector field and Δ is the Laplace-Beltrami operator for some Riemannian metric on the state space ($\mathbb{R}^n$ or M). It is therefore natural to consider (1) on some smooth

Riemannian manifold (so $X(x)$ is now a linear map from $\mathbb{R}^m$ onto the tangent space T_xM at x to M for each x in M) and to assume that $\mathcal{A} = \frac{1}{2}\Delta + Z$. This occurs if $X(x)$ is surjective and induces the given Riemannian inner product $<,>_x$ on T_xM and

$$Z(x) = A^X(x) \equiv \frac{1}{2}\sum_1^m \nabla X^i(X^i(x)) + A(x)$$

using the Levi-Civita connection.

5.1 Compact manifolds

A. For compact M there is a version of the solution flow which is continuous in time onto the space of C^∞ diffeomorphisms of M [Elw78], [CE83], [Kun90]. The derivatives now consist of linear maps $T_{x_0}F_t : T_{x_0}M \to T_{x_t}M$ and if $v_t = T_{x_0}F_t(v_0)$ for $v_0 \in T_{x_0}M$ it satisfies the covariant equation

$$Dv_t = \nabla X(v_t) \circ dB_t + \nabla A(v_t)dt \tag{13}$$

along the solution $\{x_t : t \geq 0\}$ [Elw88]. In this situation the previous proofs go through in the same way, using covariant differentiation to replace the usual differentiation. Since we can take the adjoint $X(x)^* : T_xM \to \mathbb{R}^m$ as right inverse for $X(x)$ we obtain [EL94] our basic formula in the form

$$d(P_tf)(v_0) = \frac{1}{t}\mathbb{E}\left\{f(x_t)\int_0^t < v_s, X(x_s)dB_s >\right\} \tag{14}$$

where the left hand side is the differential of P_tf in the direction $v_0 \in T_{x_0}M$, c.f. [Elw92]. Here $t > 0$ and f need only to be bounded and measurable. There are analogous formulae to (7) and (10) for higher derivatives and for the case of a potential $V : M \to \mathbb{R}$, see [EL94].

B. There is an alternative approach which leads to Bismut's formula and some generalizations. Instead of using the derivative flow TF_t we use the 'Hessian flow' $\left\{W_t^Z : t \geq 0\right\}$ on TM which is defined to be the solution flow to the covariant equation along $\{x_t : t \geq 0\}$

$$\frac{D}{\partial t} W_t^Z(v_0) = -\frac{1}{2} \text{Ric}^{\#} \left(W_t^Z(v_0), -\right) + \nabla Z \left(W_t^Z(v_0)\right) \tag{15}$$

where $\text{Ric}^{\#}(v, -) \in T_x M$ for $v \in T_x M$ is defined in terms of the Ricci curvature by

$$\left\langle \text{Ric}^{\#}(v, -), w \right\rangle_x = \text{Ric}(v, w) \quad w \in T_x M.$$

(We are assuming $\mathcal{A} = \frac{1}{2}\Delta + Z$.) Then [Elw88], for $v_0 \in T_{x_0} M$,

$$d(P_t f)(v_0) = \mathbb{E} df \left(W_t^Z(v_0)\right). \tag{16}$$

Hence, arguing as before,

$$dP_t f(v_0) = \frac{1}{t} \mathbb{E} \left\{ f(x_t) \int_0^t < W_s^Z(v_0), X(x_s) dB_s > \right\}. \tag{17}$$

or, in a completely intrinsic form, if $\{\tilde{B}_t : 0 \leq t < \infty\}$ is the martingale part of the stochastic anti-development of $\{x_s : 0 \leq s \leq t\}$ using an orthonormal frame u_0 at x_0 and if $\{u_t : 0 \leq t < \infty\}$ is parallel translation of u_0 along $\{x_t : 0 \leq t < \infty\}$ then

$$d(P_t f)(v_0) = \frac{1}{t} \mathbb{E} \left\{ f(x_t) \int_0^t < W_s^Z(v_0), u_s d\tilde{B}_s > \right\}. \tag{18}$$

This reduces to Bismut's formula [Bis84] when $Z = 0$.

C. Although (18) is intrinsic (i.e. it is written entirely in terms of the underlying diffusion $\{x_t : t \geq 0\}$) whereas (14) uses the derivative flow which involves the stochastic differential equation used to obtain the diffusion, there can be advantages in using (14). One obvious one is the simplicity of it, especially in the basic form (5). Also given special situations it is appropriate to use specially chosen stochastic differential equations e.g. when M is a Lie group with bi-invariant Riemannian metric its solutions can be represented as solutions of a left invariant s.d.e. (1) (i.e. m=dim$M$ and the A and X^i are left invariant vector fields) e.g. see [Elw82]. The solution flow is given by $F_t(x_0) = g_t x_0, t \geq 0, x_0 \in M$ where $\{g_t : t \geq 0\}$ is the solution starting from the identity element $\mathbf{1}$. Equation(14) can be written [EL94] as

$$d(P_t f)(v_0) = \frac{1}{t}\mathbb{E}\left\{ f(g_t) \int_0^t < \mathrm{ad}(g_s)^{-1}(v_0), u_s d\check{B}_s >_{\mathbf{1}} \right\}.$$

for $v_0 \in T_{\mathbf{1}}M$ where $\{\check{B}_t : 0 \leq t < \infty\}$ is given by $\check{B}_t = \sum_1^m X^i(\mathbf{1})B_t^i$.

Recall that (1) is called a *gradient system with drift* if $X(x)$ is given by $X(x)(e) = \nabla < j(x), e >$ for $e \in \mathbb{R}^m$, where $j : M \to \mathbb{R}^m$ is an isometric immersion.

For such a system with drift, $\{W_t^Z(v_0) : t \geq 0\}$ is just the predictable projection of $\{TF_t(v_0) : t \geq 0\}$ using the filtration of $\{x_t : 0 \leq t < \infty\}$. Essentially if you filter out the redundant noise in $\{TF_t(v_0) : t \geq 0\}$ then $\{W_t^Z(v_0) : t \geq 0\}$ is obtained [EY93]:

$$\mathbb{E}\left\{TF_t(v_0) \mid \sigma\{x_s : 0 \leq s \leq t\}\right\} = W_t^Z(v_0)$$

where the conditional expectation is defined by parallel translating back to $T_{x_0}M$ then taking the classical conditional expectation of the resulting $T_{x_0}M$ valued process, then parallel translating back to $T_{x_t}M$.

D. Just as in §3C formulae (14), (17), (18) lead to formulae for the logarithmic derivatives of the heat kernel, and in particular Bismut's formula comes from (18).

5.2 More general coefficients on R^n and non-compact manifolds

A. For non-compact manifolds, including $\mathbb{R}^n$, when (1) does not have globally Lipschitz coefficients, there may be no version of the flow $\{F_t : t \geq 0\}$ continuous in space, even though we are assuming non-explosion [Elw78]. However we can still define $\{v_t : t \geq 0\}$ by (13) or (4) (this will be the derivative in probability in the direction v_0 so we will continue to write it as $T_{x_0}F_t(v_0)$) and hope that the formulae still hold with some integrability conditions on $\{v_t : t \geq 0\}$.

One crucial ingredient in the proofs is the possibility of differentiating under the expectation. From $P_tf(x_0) = \mathbb{E}f(F_t(x_0))$ we want to obtain

$$d(P_tf)_{x_0}(v_0) = \mathbb{E}\,(df)_{x_t}\,(v_t). \tag{19}$$

To see when this holds, the concept of strong 1-completeness is useful [Li92]. Our stochastic differential equation is *strongly 1-complete* if for each smooth curve $\sigma : [a, b] \to M$ there is a version of $\{F_t : t \geq 0\}$ which is jointly continuous when restricted to the image of σ. As a consequence its composition $F_t \circ \sigma$ is smooth on $[a, b]$ for each $t \geq 0$. The equation $dx_t = dB_t$ on $\mathbb{R}^n - \{0\}$ is strongly 1-complete if $n \geq 3$, even though it does not admit a global smooth flow. Strong 1-completeness follows from certain integrability conditions on v_t, see [Li92].

Using Fubini's theorem and the existence of smooth partial flows we can obtain [EL94]: *Suppose (1) is strongly 1-complete and $\mathbb{E}|T_xF_t|$ exists and is locally integrable in x. Then, for f bounded with continuous and bounded first derivative, P_tf has a weak derivative and (19) holds in the weak sense.* No ellipticity hypothesis was used here. These considerations yield the following [EL94]:

For s.d.e. (3) on $\mathbb{R}^n$, suppose there is a smooth bounded $Y : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ which is the right inverse to $X(x)$ at each $x \in \mathbb{R}^n$. Then (5) holds a.s. with respect to the Lebesgue measure for bounded measurable f, provided that

$$2 < DZ(x)(v), v > + \sum_1^m \left|DX^i(x)(v)\right|^2 + \sum_1^m \left\langle DX^i(x)v, v\right\rangle^2$$

is bounded on $\{(x,v) \in \mathbb{R}^n \times \mathbb{R}^n : |v| \le 1\}$.

For the manifold case and higher derivatives, see [EL94].

6 Sobolev Estimates

Let h be a smooth function on M. Assume the generator of the stochastic differential equation is $\Delta^h = \frac{1}{2}\Delta + 2\text{Ł}_{\nabla h}$, which is called the Bismut-Witten Laplacian. It is essentially self-adjoint on $L^2(M, \mathrm{e}^{2h}dx)$, the space of L^2 functions on M, for dx the Riemannian volume measure. In fact if δ^h is the adjoint of d in $L^2(M, \mathrm{e}^{2h}dx)$, then $\Delta^h = -(d + \delta^h)^2$. We shall denote its closure also by Δ^h.

Let $1 \le p \le \infty$. There is the Sobolev space $W^{p,1} = \{f : f, \nabla f \in L^p(M, \mathrm{e}^{2h}dx)\}$ with norm $|f|_{L^{p,1}} = |f|_{L^p} + |\nabla f|_{L^p}$. Suppose

$$k^2 =: \sup_{x \in M} E \int_0^t |T_x F_s|^2 ds < \infty.$$

Then [EL94] we have the following estimate from formula (14) for $d(P_t f)$: *Suppose (19) for all bounded functions in BC^1. Then (14) holds almost* *everywhere for any* $f \in L^p$, $1 < p \leq \infty$, *and, for* $t > 0$, P_t *gives a continuous map*

$$P_t\colon L^p(M, \mathrm{e}^{2h}dx) \to W^{p,1}(M, \mathrm{e}^{2h}dx), \quad 1 < p \leq \infty$$

with

$$|(P_t f)|_{L^{p,1}} \leq (1 + \frac{k_p}{t})|f|_{L^p}, \tag{20}$$

where $k_p = k$ *for* $2 \leq p \leq \infty$, *and* $k_p = c_p k^p$ *for* $1 < p < 2$ *and* c_p *a universal constant.*

Alternatively using (15), *assuming Ricci-2Hess(h) is bounded from below by a number c then the above estimate holds with* $k^2 = \frac{1}{c}\mathrm{e}^{ct}$. Here Hess(h) is the Hessian of the function h.

7 Flatness of the trough of approximate travelling waves of KPP equations

Consider the K.P.P. equation on $\mathbb{R}$ with parameter μ:

$$\begin{cases} \frac{\partial f_t^\mu}{\partial t} &= \frac{\mu^2}{2}\Delta f_t^\mu + \frac{1}{\mu^2}\hat{c}(1 - f_t^\mu) f_t^\mu \\ f_0^\mu(x) &= \mathrm{e}^{-\frac{\alpha}{\mu^2}x^2} \end{cases} \tag{21}$$

for α and $\hat{c}$ constants. Then following Freidlin e.g. [Fre85], it was shown in [ZE92] that

$$\lim_{\mu \to 0} |f_t^\mu(x)| = 0, \quad \text{if } \alpha x^2 > \hat{c}t(1 + 2\alpha t).$$

We shall show that this convergence is C^1 in x as an example of more general results from [LZ]. First by (10), with $B.$ now one-dimensional

$$
\begin{aligned}
&Df_t^\mu(x_0)(v_0) = \tfrac{1}{t}\mathbb{E}\mathrm{e}^{-\frac{\alpha}{\mu^2}(x_0+\mu B_t)^2}\mathrm{e}^{\frac{\hat{c}}{\mu^2}\int_0^t\left(1-f_{t-s}^\mu(x_0+\mu B_s)\right)ds}\int_0^t v_0\mu dB_s\\
&+\tfrac{1}{t}\mathbb{E}\mathrm{e}^{-\frac{\alpha}{\mu^2}(x_0+\mu B_t)^2}\mathrm{e}^{\frac{\hat{c}}{\mu^2}\int_0^t\left(1-f_{t-s}^\mu(x_0+\mu B_s)\right)ds}\int_0^t\left(-\tfrac{1}{\mu^2}\hat{c}\right)(t-r)Df_{t-r}^\mu(x_0+\mu B_r)(v_0)dr.
\end{aligned}
\tag{22}
$$

Let $V_t(x) = \hat{c}t - \frac{\alpha x^2}{2\alpha t+1}$. Let $\{y_t\}$ be the solution to

$$
dy_t = \mu dB_t - \frac{2\alpha}{2\alpha t+1}y_t dt
$$

starting from x_0. Then by the Girsanov-Cameron-Martin formula, (22) becomes

$$
\begin{aligned}
&\mathrm{e}^{-\frac{1}{\mu^2}\left(\hat{c}t-\frac{\alpha x^2}{2\alpha t+1}\right)}Df_t^\mu(x_0)(v_0)\\
=\;&\tfrac{\mu}{t}\tfrac{1}{\sqrt{2\alpha t+1}}\mathbb{E}\mathrm{e}^{-\frac{\hat{c}}{\mu^2}\int_0^t f_{t-r}^\mu(y_r)dr}v_0(y_t-y_0)\\
&-\tfrac{\hat{c}}{\mu^2 t}\tfrac{1}{\sqrt{2\alpha t+1}}\mathbb{E}\mathrm{e}^{-\frac{\hat{c}}{\mu^2}\int_0^t f_{t-r}^\mu(y_r)dr}\int_0^t(t-r)Df_{t-r}^\mu(y_r)(v_0)dr.
\end{aligned}
\tag{23}
$$

Note that $|Df_t^\mu|$ is uniformly bounded on $[0,T]$ by applying Gronwall's inequality to (10). Then letting $\mu \to 0$, we get the convergence to zero of $|Df_t^\mu(x)|$, when $\hat{c}t(1+2\alpha t) < \alpha x^2$.

8 Heat equations for forms

Let $\mathrm{e}^{\frac{1}{2}\Delta^h t}$ be the semigroup corresponding to $\frac{1}{2}\Delta^h$ (for Δ^h as in §6), defined by the spectral theorem. We shall use $P_t^{h,q}$ for its restriction to differential q-forms.

For a differential q-form θ, define a (q-1) form $\int_0^t \theta \circ dx_s$ by

$$
\begin{aligned}
\int_0^t \theta\circ dx_s(\alpha_0) =: \;& \tfrac{1}{q}\int_0^t \theta\left(X(x_s)dB_s, TF_s(\alpha_0^1),\dots,TF_s(\alpha_0^{q-1})\right)\\
&-\tfrac{1}{2}\int_0^t \delta^h\theta\left(TF_s(\alpha_0^1),\dots,TF_s(\alpha_0^{q-1})\right)\,ds
\end{aligned}
\tag{24}
$$

for $\alpha_0 = (\alpha_0^1, \dots, \alpha_0^{q-1})$ a (q-1) vector. Let M be a compact manifold. Then *[Li92],[EL94] if θ is a differential q-form,*

$$(P_t^{h,q}\theta)_{x_0} = \frac{1}{t}\mathbb{E}\int_0^t < X(x_s)dB_s, T_{x_0}F_s(\cdot) > \wedge \int_0^t \theta \circ dx_s \tag{25}$$

for a gradient system with drift ∇h induced by some isometric immersion $j : M \to \mathbb{R}^m$.

If $\theta = d\phi$ is an exact form, then $d(P_t^{h,q-1}\phi) = P_t^{h,q}\theta$. From (25) or by a direct proof as for (5), *for any (q-1)-form ϕ, $q = 1, 2, \dots,$*

$$dP_t^{h,q-1}\phi = \frac{1}{t}\mathbb{E}\{d\psi_t \wedge (F_t)^*\phi\}$$

for $\psi_t(x) = \int_0^t < j\left(F_s(x)\right), dB_s >_{\mathbb{R}^m}$.

References

[BF61] Y. Blagovescenskii and M. Freidlin. Some properties of diffusion processes depending on a parameter. *Soviet Math.*, 2(DAN 138):633–636, 1961.

[Bis84] J. M. Bismut. *Large deviations and the Malliavin calculus. Progress in Math. 45.* Birkhaűser, 1984.

[CE83] A. Carverhill and D. Elworthy. Flows of stochastic dynamical systems: the functional analytic approach. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, 65:245–267, 1983.

[DPEZ95] G. Da Prato, D. Elworthy, and Zabczyk. Strong Feller property for stochastic semilinear equations. *Stochastic Analysis and its Applications*, 13(1):35–45, 1995.

[DPMN92] G. Da Prato, P. Malliavin, and D. Nualart. Compact families of Wiener functionals. To appear, 1992.

[DPZ92a] G. Da Prato and J. Zabczyk. Regular densities of invariant measures for non-linear stochastic equations. Preprints di Matematica n. 149. Scuola Normale Superior, Pisa, 1992.

[DPZ92b] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Encyclopedia of Mathematics and its Applications 44, Cambridge University Press, Cambridge., 1992.

[DPZ93] G. Da Prato and J. Zabczyk. Strong Feller property and irreducibility for diffusions on Hilbert spaces. Preprint P.A.N. Warsaw, 1993.

[EK89] R. Elliott and M. Kohlmann. Integration by parts, homogeneous chaos expansions and smooth densities. *Annals of Probability*, 17(1):194–207, 1989.

[EL94] K.D. Elworthy and Xue-Mei Li. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.*, 125(1):252–286, 1994.

[Elw78] K.D. Elworthy. Stochastic dynamical systems and their flows. In A. Friedman and M. Pinsky, editors, *Stochastic Analysis*, pages 79–95. London, New York: Academic press, 1978.

[Elw82] K.D. Elworthy. *Stochastic Differential Equations on Manifolds.* Lecture Notes Series 70, Cambridge University Press, 1982.

[Elw88] K. D. Elworthy. Geometric aspects of diffusions on manifolds. In P. L. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-*

Flour XV-XVII, 1985-1987. Lecture Notes in Mathematics, volume 1362, pages 276–425. Springer-Verlag, 1988.

[Elw92] K. D. Elworthy. Stochastic flows on Riemannian manifolds. In M. A. Pinsky and V. Wihstutz, editors, *Diffusion processes and related problems in analysis, volume II. Birkhauser Progress in Probability*, pages 37–72. Birkhauser, Boston, 1992.

[EY93] K. D. Elworthy and M. Yor. Conditional expectations for derivatives of certain stochastic flows. In J. Azéma, P.A. Meyer, and M. Yor, editors, *Sem. de Prob. XXVII. Lecture Notes in Maths. 1557*, pages 159–172. Springer-Verlag, 1993.

[Fre85] M. Freidlin. *Functional Integration and Partial Differential Equations.* Princeton Unviersity Press, 1985.

[Kha60] R. Z. Khas'minskii. Strong Feller property from semilinear stochastic evolution equation and stabilization of the solution to the Cauchy problem for parabolic equations. *Theory of probability and its applications*, 5(2):179–196, 1960.

[Kun90] H. Kunita. *Stochastic Flows and Stochastic Differential Equations.* Cambridge University Press, 1990.

[Li92] Xue-Mei Li. Stochastic Flows on Noncompact Manifolds. Ph.D. thesis, University of Warwick., 1992.

[LZ] Xue-Mue Li and H.Z. Zhao. Gradient estimates and the smooth convergence of approximate travelling waves for reaction-diffusion equations.

[Nor93] J. Norris. Path integral formulae for heat kernels and their derivatives. *Probability Theory and Related Fields*, 94:525–541, 1993.

[Pes93] S. Peszat. Sobolev spaces of functions on an infinite-dimensional domain. Warwick Preprint 49/1993, 1993.

[PZ93] S. Peszat and J. Zabczyk. Strong Feller property and irreducibility for diffusions on Hilbert spaces. Preprint of Polish Academy of Sciences, (1993).

[Wat84] S. Watanabe. *Lectures on stochastic differential equations and Malliavin Calculus.* Bombay: Tata Institute of Fundamental Research, Springer-Verlag, 1984.

[ZE92] H. Zhao and K. D. Elworthy. The travelling wave solutions of scalar generalized KPP equations via classical mechanics and stochastic approaches. In A. Truman and I.M. Davies, editors, *Stochastics and Quantum Mechanics*, pages 298–316. World Scientific, 1992.