Higher order derivatives of heat semigroups on spheres and Riemannian symmetric spaces

K. D. Elworthy
Mathematics Institute, University of Warwick,
Coventry CV4 7AL, England

October 26, 2020

Abstract

* As a very special case of a more general procedure a formula is derived for the Hessian of the solutions $P_t f$ of the heat equation for functions on the sphere S^n. The formula demonstrates that for higher order derivatives there can be a spectrum of decay/growth rates, unlike the generic situation for first derivatives which is fundamental for Bakry-Emery theory. The method used is then applied for higher derivatives for spheres, and could be used for compact Riemannian symmetric spaces.
Key words *stochastic analysis, stochastic flows, symmetric spaces, heat semigroup, Bakry-Emery, diffusion of symmetric tensors, semi-group domination.*
Mathematics subject classification 58J65 (58J70 60H30 60J60 43A85)

1 Introduction

A well known and fundamental result concerning the heat-semigroup $\{P_t\}_{t\geq 0}$ of a complete Riemannian manifold M is that of Bakry-Emery theory, [1]

$$|\nabla P_t(f)| \leq e^{-ct} P_t(|\nabla f|) \qquad \text{iff } \; c|v|^2 \leq \mathrm{Ric}(v,v) \text{ all } v \in TM \tag{1}$$

where $\mathrm{Ric} : TM \bigoplus TM \to \mathbf{R}$ is the Ricci curvature of M. Bakry-Emery theory, [2], [3], shows how to extend it to much more general classes of heat semi-groups, and it can then be used to define the notion of generalised Ricci curvature bounded below in much more general situations than Riemannian geometry. An obvious question is whether similar expressions hold for higher derivatives of $P_t f$ with exponential rates given in terms of the geometry of the Riemannian manifold. With this in mind we obtain expressions for the second and third derivatives of $P_t f$ when M is a sphere with its standard Riemannian structure, Theorems 4.1 and 5.5 respectively, and also give expressions for all symmetrised derivatives in Theorem 5.1. These suggest that the situation is more complicated, and possibly more interesting, than expected. The approach we give, based on earlier work with Yves LeJan & Xue-Mei Li, [13], can be extended to arbitrary compact Riemannian symmetric spaces, and should give similar formulae. However we have not done this.

The second derivative, or Hessian, is symmetric. This is not true in general for higher derivatives; see Subsection 5.1 below. It is simpler to compute the symmetrised versions. For the symmetrised versions the exponential rate is controlled by a Weitzenböck term, in the sense of [13], [8], which for spheres turns out to be essentially the Weitzenböck term for the Lichnerowicz Laplacian, eg see [5]. For general M, the latter has been shown by Bettiol & Mendes, [6], to characterise sectional curvature bounds. However our exponential rate can be expected to involve derivatives of curvature for general M; for spheres these vanish.

A relevant result by James Thompson, [27], is that if M is compact then for each $p \in N$ and $\epsilon > 0$ there is a constant $C_p(\epsilon) > 0$ such that for all C^1 functions $f : M \to \mathbf{R}$

$$|\nabla^p P_t f|_\infty \leq C_p(\epsilon) e^{-\lambda t} |\nabla f|_\infty \text{ for all } t > \epsilon$$

where $\lambda > 0$ is the spectral gap of M. This involves the smoothing behaviour of the semigroup when $p > 1$, which is why t needs be kept away from 0. It demonstrates that there is uniform rate of decay for all derivatives as $t \to \infty$, but our formulae suggests that a more detailed analysis involving the directions of the derivatives could be rewarding. Indeed for spheres our estimate (68) shows that for derivatives taken in orthonormal directions the rate of decay increases with $1 \leq p \leq n$. For $p = 2$ it is bounded above by e^{-nt}, (51). For S^n with $\frac{1}{2}\triangle$ the spectral gap, is $\frac{n}{2}$.
An excellent survey of work on higher order derivative formulae can be found in the introduction to Xue-Mei Li's article, [19]. Much of this concerns the technically harder problem of considering derivatives of the heat

kernels. Usually just the first and second derivatives are discussed, though a notable early example giving path integral formulae is Norris's work, [25]. See Section 55 for the result of applying [19] to our situation on S^n.

Our treatment here of $S^n$is as a very special illustrative example of the more general situation described in [14]. We give the necessary geometric background, and give the proof of a simple case concerning the expectation of representations of diffusing Lie group elements. As pointed out in section 4.1, below, there are alternative methods for S^n, and a purely algebraic one could be the most economical.

1.1 Acknowledgement

This was written for the 80th Birthday of Sergio Albeverio, and I am very happy to be able to record my appreciation of Sergio as a mathematical colleague, and my enjoyment at having known him personally for close to half of those 80 years. Thanks also to the organisers of the joyous and stimulating workshop in the beautiful city of Verona in honour of that birthday.

2 Brownian motion on spheres as symmetric spaces

2.1 The sphere as a symmetric space

Consider the sphere S^n as the set of unit vectors of $\mathbf{R}^n$ with its induced topology, differential structure, and Riemannian metric. It is acted on transitively and smoothly by the special orthogonal group $SO(n+1)$. Let x_0 be a given point in S^n; we can take it to be the North Pole, $(0,0,0...,1)$. This identifies the subgroup $SO(n+1)_{x_0}$, of those $\theta \in SO(n+1)$ which fix x_0, with $SO(n)$. We have the projection

$$p : SO(n+1) \to S^n \qquad\qquad p(k) = k(x_0) \quad k \in SO(n+1) \tag{2}$$

which identifies S^n with the quotient space $SO(n+1)/SO(n)$. It is a principal bundle with group $SO(n)$. For us the main import of that will be that there is the right action of $SO(n)$

$$SO(n+1) \times SO(n) \to SO(n+1) \qquad\qquad (k,g) \mapsto k.g$$

with $p(k.g) = p(k)$.
We want p to be a Riemannian submersion. This means that we have an

inner product $\langle -, - \rangle_k$ on each tangent space $T_k SO(n+1)$ such that $T_k p : T_k SO(n+1) \to T_{p(k)} S^n$, the derivative of p at k is an orthogonal projection. We also want this Riemannian structure to be bi-invariant and so it suffices to take

$$\langle A, B \rangle_{Id} = -\frac{1}{2} \operatorname{trace} AB^* \qquad A, B \in \mathfrak{so}(\mathfrak{n}+\mathfrak{1}) \cong T_e SO(n+1).$$

With this choice, if $\{k_t\}_{t \geq 0}$ is a Brownian motion on $SO(n+1)$ starting at the identity Id, then $\{x_t\}_{t\geq 0}$ with $x_t = p(k_t) = k_t.x_0$ is a Brownian motion on S^n from x_0. Moreover if we define $\xi_t : S^n \to S^n$ by $\xi_t(y) = k_t.y$ we have a stochastic flow of Brownian motions on the sphere. For example see [7] or [13]. In particular if P_t denotes the heat semi-group acting on continuous functions on S^n then

$$P_t f(y) = \mathbf{E} f(\xi_t(y)) \qquad f : S^n \to \mathbf{R} \quad y \in S^n \tag{3}$$

Recall that $f_t = P_t f : S^n \to \mathbf{R},\ t \geq 0$ is the classical solution to the heat equation $\frac{df_t}{dt} = \frac{1}{2} \triangle f_t,\ \ f_0 = f$ on R^n. Here $\triangle$ is the Laplace Beltrami operator, $\triangle = \operatorname{div} \operatorname{grad}$, on S^n.

2.2 Derivatives of the heat semigroup.

Assume now that f is C^∞, then we can differentiate equation (3) in the direction of some $v \in T_{x_0} S^n$ to give

$$d(P_t f)(v) = \mathbf{E}\{ df_{x_t}(T_{x_t} \xi_t(v))\}. \tag{4}$$

Recall that the derivative of f gives a differential one-form $df_y : T_y S^n \to \mathbf{R}$, and the derivative of the flow gives, random, linear isomorphisms,
$T_y \xi_t : T_y S^n \to T_{\xi_t(y)} S^n$, for $y \in S^n$.

2.2.1 Aside on calculus on spheres

In order to differentiate again we need a connection on S^n. This gives a covariant derivative operator ∇ with which tensor fields such as df can be differentiated in tangent directions. Equivalently it gives a differentiation operator $\frac{D}{dt}$ of tensor fields along C^1 curves σ, and a parallel translation operator $/\!/_t : T_{\sigma(0)} S^n \to T_{\sigma(t)} S^n$ of tangent vectors, or of other tensors. These are related, for example by

$$\frac{D}{dt} V_t = /\!/_t \frac{d}{dt} /\!/_t^{-1} V_t \qquad V_t \in T_{\sigma(t)} S^n, \tag{5}$$

and if $v = \dot{\sigma}(0)$

$$\nabla_v(df) = \frac{D}{dt}(df_{\sigma(t)})|_{t=0} = \frac{d}{dt}(df_{\sigma(t)} /\!\!/_t)|_{t=0} \tag{6}$$

Stratonovich calculus allows these operations to be extended, almost surely, to the situation where σ is a continuous semi-martingale, such as our Brownian motion $\{x_t\}_t$. Also we can differentiate our stochastic flow successively, for example to get $\nabla_{u_0} T\xi_t : T_{x_0}S^n \to T_{x_t}S^n$ for $u_0 \in T_{x_0}S^n$, given by

$$\nabla_{u_0} T\xi_t(v_0) = \frac{D}{ds}(T_{\sigma(s)}\xi_t(/\!\!/_s v_0)|_{s=0} \qquad u_0, v_0 \in T_{x_0}S^n \quad \dot{\sigma}(0) = u_0. \tag{7}$$

All this holds for any Riemannian manifold, and there is a unique connection, the Levi-Civita connection, for which parallel translations consist of orthogonal transformations and also

$$\frac{D}{\partial s}\frac{\partial}{\partial t} f(\sigma(s,t)) = \frac{D}{\partial t}\frac{\partial}{\partial s} f(\sigma(s,t)) \tag{8}$$

for a two parameter $\sigma(s,t)$ and $f : M \to \mathbf{R}$, both smooth.

We will use this. For S^n it has the natural definition that $\frac{D}{dt}V_t$ is obtained by considering the vector field V_t along σ as having values in $\mathbf{R}^{n+1}$, differentiating this in t as usual and projecting the result back to $T_{\sigma(t)}S^n$.

Note: for $\{e_j\}_{j=1}^n$ an orthonormal base for T_yS^n, $y \in S^n$

$$\triangle f(y) = \mathrm{trace}(\nabla(df))_y = \Sigma_{j=1}^n \nabla_{e_j}(df)e_j. \tag{9}$$

The *Hessian*, $\mathrm{Hess}(f)$, of f is just the second derivative considered as a bilinear form

$$\mathrm{Hess}(f)_y = \nabla_{-}(df)(-) : T_yS^n \times T_yS^n \to \mathbf{R}. \tag{10}$$

By equation (8), the Hessian is symmetric and so determines a linear map on the symmetric tensor product $T_yS^n \bigodot T_yS^n$ by

$$\mathrm{Hess}(f)(u \odot v) = \nabla_u(df)(v) \qquad u, v \in T_yS^n. \tag{11}$$

2.3 Higher derivatives of P_tf

Using the Levi-Civita connection we can differentiate equation (4) again to obtain, for $u_0, v_0 \in T_{x_0}S^n$:

$$\mathrm{Hess}(P_t f)(u_0 \odot v_0) = \mathbf{E}\{\mathrm{Hess}(f)(T_{x_0}\xi_t u_0 \odot T_{x_0}\xi_t v_0) + df_{x_t} \nabla_{u_0}(T\xi_t)(v_0)\}. \tag{12}$$

An important simplification arises since our flow is a flow of isometries. In this situation covariant second order derivatives of the flow vanish, see [5]. Thus for $u_0, v_0 \in T_{x_0}S^n$:

$$\mathrm{Hess}(P_t f)_{x_0}(u_0 \odot v_0) = \mathbf{E}\{\mathrm{Hess}(f)_{x_t}(T_{x_0}\xi_t u_0 \odot T_{x_0}\xi_t v_0)\}. \tag{13}$$

and repeating the differentiaton, for $k = 1, 2, ...$ and $u_0^1, ..., u_0^k, v_0 \in T_{x_0}S^n$:

$$\nabla^{(k)} d(P_t f)(u_0^k, ..., u_0^1, v_0) = \mathbf{E}\{\nabla^k(df)(T_{x_0}\xi_t u_0^k, ..., T_{x_0}\xi_t u_0^1, T_{x_0}\xi_t v_0)\} \tag{14}$$

But the derivatives are not symmetric when $k \geq 2$ and $n \geq 2$; the curvature intervenes. See Subsection 5.1 below.

We can get a more precise formula from formula (12) by computing the conditional expectation of

$$T_{x_0}\xi_t \odot T_{x_0}\xi_t : T_{x_0}S^n \odot T_{x_0}S^n \to T_{x_t}S^n \odot T_{x_t}S^n$$

with respect to the σ-algebra $\mathcal{F}_t$ generated by the Brownian motion $\{x_s : 0 \leq s \leq t\}$. This technique, of filtering out the redundant noise, has been a basic tool for looking at first derivatives since [12]. It is described in detail in [10]. In essence the conditional expectation is obtained by parallel translation back to the initial point:
Write $u_t = T_{x_0}\xi_t(u_o)$ and $v_t = T_{x_0}\xi_t(v_o)$ and set $\overline{u_t \odot v_t} = \mathbf{E}\{u_t \odot v_t | \mathcal{F}_t\}$; then, essentially by definition,

$$\overline{u_t \odot v_t} = (/\!/_t \odot /\!/_t)\mathbf{E}\{/\!/_t^{-1} u_t \odot /\!/_t^{-1} v_t | \mathcal{F}_t\}. \tag{15}$$

Since $/\!/_t^{-1} u_t \odot /\!/_t^{-1} v_t$ lies in a fixed vector space its conditional expectation makes classical sense. There is no problem about integrability, and any choice of parallel translation in $TS^n \odot TS^n$ will do, [10].
We proceed to calculate this conditional expectation using techniques from [13], see also [8].

3 Decomposition and conditioning of $T\xi_t \odot T\xi_t$

3.1 Decomposition of the flow

Remember ξ_t is just the action of the Brownian motion $\{k_t\}_t$, on $SO(n+1)$, on our sphere. Also the Brownian motion $\{x_t\}_t$, from x_0 on the sphere, is given by $x_t = p(k_t) = k_t.x_0$. From [9] we have a skew product decomposition:

$$k_t = \tilde{x}_t.g_t \tag{16}$$

where $\{g_t\}_t$ is a Brownian motion on $SO(n)$ from the identity, independent of $\{\mathcal{F}_t\}_{t\geq 0}$, and $\{\tilde{x}_t\}_t$ is a diffusion process adapted to $\{\mathcal{F}_t\}_{t\geq 0}$ with $p(\tilde{x}_t) = x_t$ for $t \geq 0$. In fact $\{\tilde{x}_t\}_t$ is the "horizontal lift" of Brownian motion on S^n from the identity, and is the conditioned process of $\{x_t\}_t$ given $\{\mathcal{F}_t\}_{t\geq 0}$. Moreover if we write $\tilde{\xi}_t : S^n \to S^n$ for $y \mapsto \tilde{x}_t.y$ then parallel translation $\{/\!/_t\}_{t\geq 0}$ along $\{x_t\}_t$ is given by

$$/\!/_t = T_{x_0}\tilde{\xi}_t : T_{x_0}S^n \to T_{x_t}S^n. \tag{17}$$

See [11] or [13] for more.

Identifying $g \in SO(n)$ with its action on S^n let $\rho^{\odot^2}$ denote the representation of $SO(n)$ on $T_{x_0}S^n \odot T_{x_0}S^n$ given by

$$\rho^{\odot^2}(g)(u_0 \odot v_0) = T_{x_0}L_g u_0 \odot T_{x_0}L_g v_0. \tag{18}$$

From above, using the independence of g_t from $\mathcal{F}_t$, we have:

Lemma 3.1 For a C^2 function $f : S^n \to \mathbf{R}$ and $u_0, v_0 \in T_{x_0}S^n$

$$\mathrm{Hess}(P_t f)(u_0, v_0) = \mathbf{E}\left\{\mathrm{Hess}(f)_{x_t}\left((/\!/_t \odot /\!/_t)\mathbf{E}\left\{\rho^{\odot^2}(g_t)(u_0 \odot v_0)\right\}\right)\right\}. \tag{19}$$

We go on to compute the second expectation appearing above.

3.2 Expectations of representations of random matrices: an elementary lemma

The following is a very special case of a similar result for finite dimensional representations of certain, possibly time inhomogenous, diffusions on possibly infinite dimensional groups. It is essentially Theorem 3.4.1 of [13],

but see [8], or below, for a corrected sign in equation (3.19) of [13]. For completeness the simple proof is given here for Brownian motions on finite dimensional Lie groups. The more general cases will be discussed in [14]. The integrability of $\rho(g_t)v$ was proved by Baxendale, [4], for Wiener processes on Polish groups acting on Banach spaces.
Let G be a finite dimensional Lie group with right invariant metric, determined by an inner product $\langle -, - \rangle_e$ on its Lie algebra $\mathfrak{g}$ identified with the tangent space T_eG at the identity $e \in G$. We will use the Maurer-Cartan form determined by *right translations* R_g, rather than the more usual left translations L_g. It is the $\mathfrak{g}$-valued one-form ϖ given by:

$$\varpi_g := T_g(R_g)^{-1} : T_g\mathcal{G} \to \mathfrak{g} := T_e\mathcal{G} \qquad g \in G.$$

The co-differential, the adjoint d^* of d, maps one-forms to functions. It acts on ϖ component wise: let $\{\alpha^j\}_j$ be an orthonormal base for $\mathfrak{g}$ and define the scalar one forms ϖ^j by $\varpi^j(v) = \langle \varpi(v), \alpha^j \rangle_{\mathfrak{g}}$. Then $d^*\varpi(g) := \sum_j d^*\varpi^j(g)\alpha^j \in \mathfrak{g}$ for $g \in G$.

Note that

$$\varpi^j(v) = \langle A^{\alpha^j}(g), v \rangle_g \qquad g \in G, v \in T_gG$$

for $A^{\alpha^j}(g) = TR_g(\alpha^j)$, the right invariant vector field corresponding to α^j. Therefore

$$d^*\varpi(g) = -\sum_j \operatorname{div} A^{\alpha^j}(g)\alpha^j \in \mathfrak{g} \qquad g \in G. \tag{20}$$

The divergence of a vector field measures the infinitesimal rate of change of Riemannian volume μ, say, under its flow. For us the Riemannian volume is a right Haar measure. However the flow of a right invariant vector field is left translation by its 1-parameter subgroup ie $L_{e^{t\alpha^j}}$ for A^{α^j}. It follows that if the right Haar measure is also left invariant, in other words if G is unimodular then $d^*\varpi = 0$. This holds in particular for G a compact Lie group; the situation of our main present interest. In general $(L_g)_*\mu$ is again right invariant and so a multiple $m(g)$ say of μ. This version $m : G \to \mathbf{R}(>0)$ of the modular function of G is a group homomorphism. Since μ corresponds to a right invariant top dimensional form it is given by

$$m(g) = |\det \mathrm{Ad}_g|$$

for the adjoint action $\mathrm{Ad}_g = (TR_g)^{-1}TL_g : \mathfrak{g} \to \mathfrak{g}$.
Thus,

$$
\begin{aligned}
d^*\varpi^j(g) &= -\operatorname{div} A^{\alpha^j}(g) = -\frac{d}{dt}\frac{d\left((L_{e^{-t\alpha^j}})_*(\mu)\right)}{d\mu}|_{t=0} && (21)\\
&= -\frac{d}{dt}|\det \mathrm{Ad}_{e^{-t\alpha^j}}|_{t=0} && (22)\\
&= \operatorname{trace}\mathrm{ad}_{\alpha^j} = -\sum_k \langle ad^*_{\alpha_k}\alpha_k, \alpha_j\rangle && (23)
\end{aligned}
$$

for $\mathrm{ad} : \mathfrak{g} \to \mathbf{L}(\mathfrak{g}; \mathfrak{g})$ the adjoint representation, $\mathrm{ad}_\alpha(\beta) = [\alpha, \beta]$.

Lemma 3.2 Let $\rho : G \to GL(V)$ be a smooth representation of G on a real finite dimensional vector space V and denote by $\rho_* : \mathfrak{g} \to \mathbf{L}(V; V)$ the derivative of ρ at the identity element.
Let $\{g_t\}_t$ be Brownian motion on G from the identity.
Then $\rho(g_t)v$ is integrable for each $v \in V$ and $t \geq 0$ and its expectation is differentiable in t with

$$
\frac{d}{dt}\mathbf{E}\{\rho(g_t)v\} = \lambda^\rho(\mathbf{E}\{\rho(g_t)v\}) \tag{24}
$$

where $\lambda^\rho \in \mathbf{L}(V; V)$ is given by

$$
\lambda^\rho = \frac{1}{2}\operatorname{Comp}\sum_j(\rho_*(\alpha^j)\otimes\rho_*(\alpha^j)) + \frac{1}{2}\sum_k \mathrm{ad}^*_{\alpha^k}\alpha^k \tag{25}
$$

with

$$
\operatorname{Comp} : \mathbf{L}(V; V)\otimes\mathbf{L}(V; V) \to \mathbf{L}(V; V)
$$

the composition map $A \otimes B \mapsto AB$. For unimodular groups, and in particular for compact Lie groups, the term $\sum_k \mathrm{ad}^*_{\alpha^k}\alpha^k$ vanishes.

Proof. By Itô's formula, as in equation (4.1) of [13],

$$
\rho(g_t)(v) = v + M_t^{d\rho v} + \int_0^t \frac{1}{2}\triangle(\rho(-)v)(g_s)\,ds, \tag{26}
$$

where $\{M_t^{d\rho v}\}_t$ is the continuous local martingale in V

$$
M_t^{d\rho v} = \int_0^t d\rho_{g_s}(T_eR_{g_s} \circ dB_s) \tag{27}
$$

where $\{B_s\}_{s\geq 0}$ is the Brownian motion on $\mathfrak{g}$ given by $dB_s := \varpi_{g_s} \circ dg_s$.

Now since $\rho : G \to GL(V)$ is a group homomorphism we see,

$$(d\rho)_k = \rho_* \circ \varpi_k(-)\rho(k) : T_k G \to \mathbf{L}(V;V) \qquad \text{for any } k \in G. \tag{28}$$

Thus

$$M_t^{d\rho v} = \int_0^t (\rho_*(\,dB_s)\rho(g_s)v)\,. \tag{29}$$

Also, using the right invariance of the Laplacian,

$$\triangle(\rho)(g_s) = \triangle(\rho \circ R_{g_s})(e) = \triangle(\rho)(e)\rho(g_s) \in \mathbf{L}(V;V).$$

From (28) and (23) we see

$$\begin{aligned}
\triangle(\rho)(e) &= -d^*(d\rho)(e) = -d^*(\rho(-)\rho_* \circ \varpi_-)(e) \\
&= \sum_j \rho_*(\alpha^j)\rho_*(\alpha^j) - d^*(\rho_* \circ \varpi)(e) \\
&= \sum_j \rho_*(\alpha^j)\rho_*(\alpha^j) + \rho_* \sum_k (\mathrm{ad})^*_{\alpha^k}\alpha^k \\
&= 2\lambda^\rho.
\end{aligned}$$

Thus equation (26) reduces to the linear equation with constant coefficients

$$d\rho(g_t)(v) = \rho_*(\,dB_t)\rho(g_t)v + \lambda^\rho \rho(g_t)v\;dt \tag{30}$$

For compact Lie groups the result is immediate since the local martingale $\{M_t^{d\rho v}\}_t$ will be bounded and so a martingale. In general we can use a stopping time argument or the basic existence theorems for equations with Lipschitz coefficients to see that $\{\rho(g_t)(v)\}_{0\le t\le T}$ is bounded in L^2 for each $T \ge 0$, so the local martingale has integrable quadratic variation and so is a martingale [26]. $\square$

3.3 Calculation for S^n

We must calculate $\mathbf{E}\left\{\rho^{\odot^2}(g_t)(u_0 \odot v_0)\right\}$ to make use of our Hessian formula (19) for S^n. By Lemma 3.2 we have

$$\mathbf{E}\left\{\rho^{\odot^2}(g_t)(u_0 \odot v_0)\right\} = \mathbb{W}_t(u_0 \odot v_0) \tag{31}$$

where $\mathbb{W}_t = \mathbb{W}_t^{\rho^{\odot^2}} : T_{x_0}S^n \odot T_{x_0}S^n \to T_{x_0}S^n \odot T_{x_0}S^n$ satisfies $\mathbb{W}_0(u_0 \odot v_0) = u_0 \odot v_0$ and

$$\frac{d}{dt}\mathbb{W}_t(u_0 \odot v_0) = \lambda^{\rho^{\odot^2}}(\mathbb{W}_t(u_0 \odot v_0)) \qquad t \ge 0.$$

Here $2\lambda^{\rho^{\odot^2}} = \sum_j \rho_*^{\odot^2}(\alpha^j)\rho_*^{\odot^2}(\alpha^j)$.

We will use a more algebraic formulation of our representation $\lambda^{\rho^{\odot^2}}$ defined in (18):

3.3.1 Identification of $\mathfrak{m}$ with $T_{x_0}S^n$

As for any smooth left action of a Lie group we have a linear map $\alpha \mapsto X^\alpha$ from $\mathfrak{so}(n+1)$ to smooth vector fields on S^n. It is given by

$$X^\alpha(y) = \frac{d}{ds}((\exp s\alpha).y)|_{s=0}.$$

In particular the derivative $T_e p : \mathfrak{so}(n+1) \to T_{x_0}S^n$ at the identity of our projection $p : SO(n+1) \to S^n$ has $T_e p(\alpha) = X^\alpha(x_0)$. It is important, [5] page 182, or [18] page 469, to note the minus sign in the identity

$$[X^\alpha, X^\beta] = -X^{[\alpha,\beta]} \qquad \alpha, \beta \in \mathfrak{so}(n+1). \tag{32}$$

Let $\mathfrak{m}$ be the orthogonal complement of $\mathfrak{so}(n)$ in $\mathfrak{so}(n+1)$. A fundamental symmetric space property is that $\mathfrak{m}$ is invariant under the adjoint action, Ad, of $SO(n)$ on $\mathfrak{so}(n+1)$, and so under $\mathrm{ad} : \mathfrak{so}(n) \to GL(\mathfrak{so}(n+1))$, its derivative at the identity e. There is the following important standard lemma with versions for more general symmetric spaces:

Lemma 3.3 There are the commutative diagrams:

1. For $g \in SO(n)$

$$\begin{tikzcd} \mathfrak{m} \arrow[r, "T_e p"] & T_{x_0}S^n \\ \mathfrak{m} \arrow[u, "\mathrm{Ad}_g"] \arrow[r, "T_e p"'] & T_{x_0}S^n. \arrow[u, "TL_g"'] \end{tikzcd}$$

2. For $\alpha \in \mathfrak{so}(n)$

$$\begin{tikzcd} \mathfrak{m} \arrow[r, "T_e p"] & T_{x_0}S^n \\ \mathfrak{m} \arrow[u, "\mathrm{ad}_\alpha"] \arrow[r, "T_e p"'] & T_{x_0}S^n. \arrow[u, "\nabla_{(-)}X^\alpha"'] \end{tikzcd}$$

Proof. For 1. :

$$\begin{aligned} TL_gT_ep(\alpha) &= TL_gX^\alpha(x_0) = \frac{d}{ds}g\exp(s\alpha).x_0|_{s=0} \\ &= \frac{d}{ds}g\exp(s\alpha)g.x_0|_{s=0} = X^{\mathrm{Ad}_g\alpha}(x_0) \\ &= T_ep(\mathrm{Ad}_g\alpha). \end{aligned}$$

For 2. : if $v \in \mathfrak{m}$

$$\begin{aligned} T_ep(\mathrm{ad}_\alpha(v)) &= T_ep([\alpha, v]) = -[X^\alpha, X^v] \\ &= -\nabla_{X^\alpha(0)}X^v + \nabla_{X^v(0)}X^\alpha \\ &= \nabla_{X^v(0)}X^\alpha \\ &= \nabla_{(-)}X^\alpha \circ T_ep(v). \end{aligned}$$

□

We will identify $\mathfrak{m}$ with $T_{x_0}S^n$ by T_ep. By the lemma the representation $\rho^{\odot^2} : SO(n) \to GL(T_{x_0}S^n \odot T_{x_0}S^n)$ gets identified with $\mathrm{Ad} \otimes \mathrm{Ad} : SO(n) \to GL(\mathfrak{m} \odot \mathfrak{m})$ using the restriction of the adjoint action. Then we have $\rho_*^{\odot^2}(\alpha) = \mathrm{ad}_\alpha \otimes \mathrm{Id} + \mathrm{Id} \otimes \mathrm{ad}_\alpha$, and so

$$\lambda^{\rho^{\odot^2}} = \frac{1}{2}\sum_j \{\mathrm{ad}_{\alpha^j} \circ \mathrm{ad}_{\alpha^j} \otimes \mathrm{Id} + \mathrm{Id} \otimes \mathrm{ad}_{\alpha^j} \circ \mathrm{ad}_{\alpha^j} + 2\mathrm{ad}_{\alpha^j} \otimes \mathrm{ad}_{\alpha^j}\}. \tag{33}$$

3.3.2 Curvature identities

Let $R : TM \oplus TM \to \mathbf{L}(TM; TM)$ denote the curvature tensor, with Kobayashi & Nomizu's convention, so for tangent vectors u, v, w at a point z we have a tangent vector $R(u, v)w$ at z, and for S^n :

$$R(u, v)w = \langle v, w\rangle u - \langle u, w\rangle v. \tag{34}$$

The Ricci curvature, $\mathrm{Ric} : TM \oplus TM \to \mathbf{R}$ is given as the trace,
$\mathrm{Ric}(u, v) = \mathrm{trace}\, R(-, u)v$, with $\mathrm{Ric}^\sharp : TM \to TM$ given by

$$\mathrm{Ric}^\sharp(u) = \sum_j R(u, e_j)e_j$$

for a suitable o.n. base.
For S^n:

$$\mathrm{Ric}(u, v) = (n-1)\langle u, v\rangle. \tag{35}$$

In our situation, from [18] page 231, and [5] page 193 taking account of Besse's different sign convention for R:
for $u, v, w \in \mathfrak{m}$ with $\mathfrak{m}$ identified with $T_{x_0}S^n$

$$R(u,v)w = -[[u,v],w]. \tag{36}$$

Noting that $\mathrm{ad}_\alpha : \mathfrak{k} \to \mathfrak{k}$ is skew-symmetric for $\alpha \in \mathfrak{so}(n+1)$ we see from this that if also $a \in \mathfrak{m}$,

$$\begin{aligned} \langle R(u,v)w, a\rangle &= \langle \mathrm{ad}_w([u,v]), a\rangle & (37)\\ &= -\langle \mathrm{ad}_u v, \mathrm{ad}_w a\rangle. & (38) \end{aligned}$$

From this, for $u, v \in \mathfrak{m}$,

$$\mathrm{Ric}(u,v) = -\mathrm{trace}_{\mathfrak{m}} \mathrm{ad}_u \mathrm{ad}_v \tag{39}$$

Here we have written $\mathrm{trace}_{\mathfrak{m}}$ to emphasise that the trace is taken for $\mathrm{ad}_u\mathrm{ad}_v : \mathfrak{m} \to \mathfrak{m}$. Indeed there are the fundamental relations:

$$[\mathfrak{g},\mathfrak{g}] \subset \mathfrak{g}, \quad [\mathfrak{g},\mathfrak{m}] \subset \mathfrak{m}, \quad [\mathfrak{m},\mathfrak{m}] \subset \mathfrak{g} \tag{40}$$

where for us $\mathfrak{g} = \mathfrak{so}(n)$. See for example [5] page 193, or [18] page 226. Therefore ad_u interchanges $\mathfrak{g}$ and $\mathfrak{m}$ so

$$\mathrm{Ric}(u,v) = -\mathrm{trace}_{\mathfrak{g}} \mathrm{ad}_u \mathrm{ad}_v = -\frac{1}{2}\,\mathrm{trace}\,\mathrm{ad}_u\mathrm{ad}_v \tag{41}$$

as in [5] page 194.

3.3.3 Decomposition of $V \odot V$

To go further we shall decompose $\mathfrak{m} \odot \mathfrak{m}$ into irreducible components for $\rho^{\odot^2}$.

For a real, n-dimensional, inner product space $V, \langle,\rangle$, the inner product, being symmetric and bilinear, determines a linear map $\langle - \rangle : V \otimes V \to \mathbf{R}$ given by

$$\langle u \otimes v\rangle = \langle u, v\rangle.$$

It is invariant under the action

$$u \odot v \mapsto Uu \odot Uv : \; U \in O(V)$$

of the orthogonal group $O(V)$ of V. Its kernel in $V \odot V$, denoted by $\mathcal{H}$, is therefore also invariant. It has codimension one and its elements are sometimes called "traceless" or "harmonic"; the latter because of the representation of symmetric tensors as homogeneous polynomials, [16], [6]. The space

$V \otimes V$ has a distinguished element $\Xi := \sum_j e_j \odot e_j$ for $\{e_j\}_j$ an orthonormal basis of V . It corresponds to the identity when $V \otimes V$ is identified with $\mathbf{L}(V; V)$ using the inner product. Using the inner product of $V \odot V$ inherited from that of $V \otimes V$ we see

$$\langle \Xi, u \odot v \rangle = \sum_j \langle u, e_j \rangle \langle v, e_j \rangle = \langle u \odot v \rangle.$$

Thus Ξ is the Riesz representative of $\langle - \rangle$ and so orthogonal to the kernel $\mathcal{H}$ and invariant under our orthogonal action. We write

$$\begin{aligned} V \odot V &= \mathbf{R}\Xi \oplus \mathcal{H} \qquad \text{with} && (42)\\ u \odot v &= \frac{1}{n}\langle u, v \rangle \Xi \oplus \left(u \odot v - \frac{1}{n}\langle u, v \rangle \Xi \right). && (43) \end{aligned}$$

Let $\mathcal{P}_{\mathcal{H}} : V \odot V \to V \odot V$ be the orthogonal projection onto $\mathcal{H}$, so

$$\mathcal{P}_{\mathcal{H}}(u \odot v) = u \odot v - \frac{1}{n}\langle u, v \rangle \Xi. \tag{44}$$

3.3.4 Computations

From (36), starting to compute $\lambda^{\rho^{\odot^2}}$ from formula (33), with our orthonormal base $\{\alpha^j\}_j$ for $\mathfrak{so}(n)$, and $u, v, a, b \in \mathfrak{m}$,

$$\begin{aligned} \sum_j \langle \mathrm{ad}_{\alpha^j} u \otimes \mathrm{ad}_{\alpha^j} v, a \odot b \rangle &= \sum_j \langle \mathrm{ad}_u \alpha^j \otimes \mathrm{ad}_v \alpha^j, a \odot b \rangle \\ &= \frac{1}{2} \sum_j \{ \langle \mathrm{ad}_u a, \alpha^j \rangle \langle \mathrm{ad}_v b, \alpha^j \rangle + \\ &\qquad + \langle \mathrm{ad}_v a, \alpha^j \rangle \langle \mathrm{ad}_u b, \alpha^j \rangle \} \\ &= -\frac{1}{2} \{ \langle R(u, a)v, b \rangle + \langle R(v, a)u, b \rangle \} \\ &= -\frac{1}{2} \{ \langle R(u, a)v, b \rangle + \langle R(u, b)v, a \rangle \} \\ &= -\langle R^\sharp(u, -)v,_\rangle (a \odot b), \end{aligned} \tag{45}$$

where $R^\sharp(u, -)v \in \mathfrak{m} \odot \mathfrak{m}$ is the dual to $a \odot b \mapsto \frac{1}{2}\langle R(u, a)v + R(v, a)u, b \rangle$. For S^n using (34)

$$\begin{aligned} \langle R(u, a)v + R(v, a)u, b \rangle &= \langle a, v \rangle \langle u, b \rangle + \langle b, v \rangle \langle u, a \rangle - 2\langle u, v \rangle \langle a, b \rangle \\ &= 2\langle u \odot v, a \odot b \rangle - 2\langle u, v \rangle \langle a, b \rangle \end{aligned}$$

whence

$$\sum_j \mathrm{ad}_{\alpha^j} u \otimes \mathrm{ad}_{\alpha^j} v = -u \odot v + \langle u, v\rangle \Xi. \tag{46}$$

Furthermore, using (41), for any $w \in \mathfrak{m}$

$$\begin{aligned}
\sum_j \langle \mathrm{ad}_{\alpha^j} \circ \mathrm{ad}_{\alpha^j} u, w\rangle &= -\sum_j \langle \mathrm{ad}_u \alpha^j, \mathrm{ad}_w \alpha^j\rangle \\
&= \sum_j \langle \mathrm{ad}_w \circ \mathrm{ad}_u \alpha^j, \alpha^j\rangle \\
&= -\operatorname{Ric}(u, w). \qquad (47)
\end{aligned}$$

We now see from (33), (47), (45)

$$\lambda^{\rho^{\odot 2}}(u \odot v) = -\frac{1}{2}\{\operatorname{Ric}^\sharp u \odot v + u \odot \operatorname{Ric}^\sharp(v)\} - R^\sharp(u, -)v. \tag{48}$$

Using the explicit expressions, (35) and (46), for S^n this yields:

$$\begin{aligned}
\lambda^{\rho^{\odot 2}}(u \odot v) &= -(n-1)(u \odot v) - u \odot v + \langle u, v\rangle \Xi \\
&= -n\mathcal{P}_{\mathcal{H}}(u \odot v). \qquad (49)
\end{aligned}$$

4 Main result for S^n

Theorem 4.1 For $x_0 \in S^n$ and u_0, v_0 in the tangent space $T_{x_0}S^n$ and a C^2 map $f : S^n \to \mathbf{R}$ the second derivative $\operatorname{Hess} P_t f$ of the solution to the heat equation

$$\begin{aligned}
\frac{d}{dt} P_t f &= \frac{1}{2} \triangle P_t f \\
P_0 f &= f
\end{aligned}$$

is given by

$$\begin{aligned}
\operatorname{Hess} P_t f(u_0, v_0) &= \frac{1}{n}(1 - e^{-nt})\langle u_0, v_0\rangle P_t(\triangle f)(x_0) + e^{-nt}\mathbf{E}\{\operatorname{Hess}(f)_{x_t}(/\!/_t u_0, /\!/_t v_0)\} \\
&= \frac{1}{n}\langle u_0, v_0\rangle P_t(\triangle f)(x_0) + e^{-nt}\mathbf{E}\{\operatorname{Hess}(f)_{x_t}\mathcal{P}_{\mathcal{H}}(/\!/_t u_0, /\!/_t v_0). \qquad (50)
\end{aligned}$$

In particular if u_0 and v_0 are orthogonal,

$$||\operatorname{Hess} P_t f(u_0, v_0|| \le e^{-nt} P_t(||\operatorname{Hess} f||)(x_0)||u_0||\,||v_0|| \qquad t \ge 0. \tag{51}$$

Proof. From (19) and from (31),

$$\mathrm{Hess}(P_t f)(u_0, v_0) = \mathbf{E}\left\{\mathrm{Hess}(f)_{x_t}\left((/\!/_t \otimes /\!/_t)\mathbb{W}_t(u_0 \odot v_0)\right)\right\}$$

where, using (31),

$$\begin{aligned}\frac{d}{dt}\mathbb{W}_t(u_0 \odot v_0) &= \lambda^{\rho^{\odot^2}}(\mathbb{W}_t(u_0 \odot v_0)) \quad t \geq 0 && (52)\\ &= -n\mathcal{P}_{\mathcal{H}}(\mathbb{W}_t(u_0 \odot v_0)). && (53)\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{W}_t(u_0 \odot v_0) &= \frac{1}{n}\langle u_0, v_0\rangle\Xi + e^{-nt}\left(u_0 \odot v_0 - \frac{1}{n}\langle u_0, v_0\rangle\Xi\right)\\ &= \frac{1}{n}(1 - e^{-nt})\langle u_0, v_0\rangle\Xi + e^{-nt}(u_0 \odot v_0).\end{aligned}$$

Write $\Xi_t := (/\!/_t \otimes /\!/_t)\Xi$. We now see

$$\begin{aligned}\mathrm{Hess}(P_t f)(u_0, v_0) &= \frac{1}{n}\langle u_0, v_0\rangle\mathbf{E}\{\mathrm{Hess}(f)_{x_t}(\Xi_t)\}\\ &\qquad + e^{-nt}\mathbf{E}\{\mathrm{Hess}(f)\mathcal{P}_{\mathcal{H}}(/\!/_t u_0 \odot /\!/_t v_0))\}\end{aligned}$$

equivalently

$$\begin{aligned}\mathrm{Hess}(P_t f)(u_0, v_0) &= \frac{1}{n}(1 - e^{-nt})\langle u_0, v_0\rangle\mathbf{E}\{\mathrm{Hess}(f)_{x_t}\Xi_t\}\\ &\qquad + e^{-nt}\mathbf{E}\{\mathrm{Hess}(f)_{x_t}(/\!/_t u_0 \odot /\!/_t v_0)\}.\end{aligned}$$

Since $\mathrm{Hess}(f)_{x_t}\Xi_t = \triangle f(x_t)$ the results follow.

$\square$

4.1 Two alternative approaches

4.1.1 Algebraic

The form of formula (50) is not surprising given the symmetries of the sphere, and the decomposition of our representation of $\mathfrak{so}(n)$ into irreducible components. Indeed for $g \in \mathfrak{so}(n+1)$ and $y \in S^n$, we have $P_t f(gy) = P_t(f \circ g)(y)$. Using the fact that for $g \in \mathfrak{so}(n)$ left translation by g preserves the law of Brownian motion from x_0, we see that $\lambda^{\rho^{\odot^2}}$ must be invariant under the action of $\rho^{\odot^2}(\mathfrak{so}(n))$. It follows that it must be constant on the irreducible components $\mathcal{H}$ and $\mathbf{R}\Xi$ of $\mathfrak{m} \odot \mathfrak{m}$ for that action. Since we must have $\text{Hess}(P_t f)(\Xi) = \triangle(P_t f)(x_0) = P_t(\triangle f)(x_0)$ we see the second constant must be zero. To compute the first constant we could proceed as in [13] Corollary 3.4.4, page 50 and relate $\sum_j \rho_*^{\odot^2}(\alpha_j) \circ \rho_*^{\odot^2}(\alpha_j)$ with the Casimir element of our representation, [17] 6.2. That way there need be no mention of curvature. However we have preferred to introduce curvature since it gives a geometric interpretation of the constants, and also our approach applies in greater generality.

4.1.2 Doubly damped parallel translation

In [20] and [19] Xue-Mei Li obtains second derivative formulae on rather general Riemannian manifolds M by differentiating the standard first derivative formula $dP_t(f)(v_0) = \mathbf{E}\{df_{\xi_t(x_0)} W_t(v_0)\}$ with $\{W_t\}$ *damped*, or *Dohrn -Guerra*, parallel translation, and $\{\xi_t\}_t$, a gradient stochastic flow. This gives a term under the expectation of the form $df_{\xi_t(x_0)} \nabla_{u_0} W_t(v_0)$. If we filter out the redundant noise, i.e. condition, $\nabla_{u_0} W_t(v_0)$, this term becomes $df_{\xi_t(x_0)} W_t^{(2)}(u_0, v_0)$ for a certain process $W_t^{(2)}(u_0, v_0) \in T_{x_t}M$ which she calls the *doubly damped parallel translation*. For our sphere

$$W_t^{(2)}(u_0, v_0) = e^{-\frac{1}{2}(n-1)t} /\!\!/_t \int_0^t e^{-(n-1)s} \left(\langle u_0, v_0 \rangle dB_s - \langle u_0, dB_s \rangle v_0 \right) \tag{54}$$

for $\{B_t\}_t$ the stochastic anti-development of our Brownian motion on S^n. and her formula gives:

$$\begin{aligned} \text{Hess}(P_t f)(u_0, v_0) &= e^{-(n-1)t} \mathbf{E}\{\text{Hess}(f)(/\!\!/_t u_0, /\!\!/_t v_0)\} \\ &\quad + \mathbf{E}\{df(W_t^{(2)}(u_0, v_0))\}. \end{aligned} \tag{55}$$

5 Extensions

5.1 Higher order derivatives

To consider 3rd order, or higher derivatives $\nabla^{(k)} d(P_t f)(u_0^k, ..., u_0^1, v_0)$, we have formula (14) but have to recall that the higher derivatives are not symmetric. To deal with this we could look at the representation theory of $so(n)$ on the full tensor algebra $\bigotimes^k \mathfrak{m}$ but this will involve sub-representations such as on $\wedge^k \mathfrak{m}$ which are not relevant to us. It seems easier to keep to the symmetric tensor products and then adjust with curvature terms as done for third derivatives below. For S^n or other symmetric spaces this is much helped by the vanishing of the covariant derivatives of the curvature.

5.1.1 Symmetrised derivatives

For the symmetrised version, for each $p = 2, 3, ...$ we use the map

$$\mathcal{C} : \bigotimes^{p} \mathbf{R}^n \to \bigodot^{p-2} \mathbf{R}^n$$

given by

$$\mathcal{C}(u^1 \otimes ... \otimes u^p) = \mathcal{C}(u^1 \odot ... \odot u^p) = \sum_{i<j} \langle u^i, u^j \rangle \bigodot^{k \neq i,j} u^k.$$

Let $\bigodot^p(\mathbf{R}^n)_{\mathcal{H}}$ denote the kernel of $\mathcal{C}$ in $\bigodot^p \mathbf{R}^n$. These are the *traceless* or *harmonic* elements. It is invariant under the representation $\rho^{\odot}$ of $SO(n)$:

$$\rho^{\odot}(g)(u^1 \odot ... \odot u^p) = (\rho(g)u^1 \odot ... \odot \rho(g)u^p)$$

for any given orthogonal representation ρ of $SO(n)$ on $\mathbf{R}^n$.

If ρ is irreducible we have the decomposition of $\bigodot^p \mathbf{R}^n$ into irreducible factors under $\rho^{\odot}$;

$$\bigodot^{p} \mathbf{R}^n = \bigodot^{p}(\mathbf{R}^n)_{\mathcal{H}} + \bigodot^{p-2}(\mathbf{R}^n)_{\mathcal{H}} \bigodot \Xi + ... + \bigodot^{p-2k} (\mathbf{R}^n)_{\mathcal{H}} \bigodot (\bigodot^{k} \Xi) + ... \tag{56}$$

For example see [6] or [16].

For $p = 3$ the decomposition is

$$\begin{aligned} u \odot v \odot w \quad = \quad & \left(u \odot v \odot w - \frac{1}{n+2} \mathcal{C}(u \odot v \odot w) \odot \Xi \right) \\ & \oplus \frac{1}{n+2} \mathcal{C}(u \odot v \odot w) \odot \Xi \end{aligned} \tag{57}$$

We can give a precise formula for arbitrarily high symmetric derivatives:

Theorem 5.1 For $p = 1, 2, ...$ and smooth $f : S^n \to \mathbf{R}$ the symmetrised p-th covariant derivative of the solution $P_t f$ to the heat equation

$$\frac{\partial}{\partial t} P_t f = \frac{1}{2} \triangle P_t f \qquad\qquad P_0 f = f$$

is given by

$$\nabla^p(P_t f)(u_0^1 \odot ... \odot u_0^p) = \mathbf{E}\{\nabla^p(f)W_t^{[p]}(u_0^1 \odot ... \odot u_0^p)\} \tag{58}$$

where the damped parallel translation $W_t^{[p]} : \bigodot^p T_{x_0}M \to \bigodot^p T_{x_t}M$ is given in terms of the decomposition (56) of $\bigodot^p T_{x_0}M$ by

$$W_t^{[p]} = W_{\mathcal{H},t}^{[p]} + W_{\mathcal{H},t}^{[p-2]} \odot /\!\!/_t|_\Xi + W_{\mathcal{H},t}^{[p-4]} \odot /\!\!/_t|_\Xi \odot /\!\!/_t|_\Xi + ... \tag{59}$$

where $/\!\!/_t|_\Xi$ refers to parallel translation restricted to Ξ, and

$$W_{\mathcal{H},t}^{[q]} : \bigodot^q (T_{x_0}M)_\mathcal{H} \to \bigodot^q (T_{x_t}M)_\mathcal{H}$$

is the restriction of $W_t^{[q]}$ to the harmonic tensors and is given by

$$W_{\mathcal{H},t}^{[q]} U_0 = e^{-\frac{q}{2}(n+q-2)t} /\!\!/_t U_0 \qquad U_0 \in \bigodot^q (T_{x_0}S^n)_\mathcal{H}. \tag{60}$$

Proof. The same argument that gave formula (19) yields

$$\nabla^p(P_t f)(U_0) = \mathbf{E}\{\nabla^p(f)(/\!\!/_t \mathbf{E}\{\rho^\odot(g_t)U_0\})\} \qquad U_0 \in \bigodot^p (T_{x_0}S^n).$$

Now

$$\rho^\odot(V \odot \Xi) = \rho^\odot(V) \odot \Xi \qquad V \in \bigodot^q (T_{x_0}S^n)$$

so formulae (58) and (59) hold with $W_{\mathcal{H},t}^{[q]}$ the restriction of $\mathbf{E}\{\rho^\odot(g_t)\}$ to $\bigodot^q (T_{x_0}S^n)_\mathcal{H}$. To calculate this we have, by Lemma 3.2

$$\frac{d}{dt}\mathbf{E}\{\rho^\odot(g_t)U_0\} = \lambda^{\rho^\odot}\mathbf{E}\{\rho^\odot(g_t)U_0\}$$

for $\lambda^{\rho^\odot} = \frac{1}{2}\,\mathrm{Comp}\sum_r(\rho_*^\odot(\alpha^r) \otimes \rho_*^\odot(\alpha^r))$.

Since $\rho_*^\odot(\alpha^r)(u^1 \odot ... \odot u^q) = \sum_\ell \rho_*(\alpha^r)u^\ell \odot^{j\neq\ell} u^j$ we have

$$\rho_*^\odot(\alpha^r)\rho_*^\odot(\alpha^r)(u^1 \odot ... \odot u^q) = A^r + B^r$$

where

$$A^r = \sum_\ell \rho_*(\alpha^r)^2 u^\ell \odot^{j\neq\ell} u^j$$

and

$$B^r = 2\sum_{j<k} \rho_*(\alpha^r)u^j \odot \rho_*(\alpha^r)u^k \odot^{\ell\neq j,k} u^\ell.$$

From formula (47) and the fact that $\mathrm{Ric}^\sharp(u) = (n-1)u$ for $u \in TS^n$ we have

$$\sum_r A^r(U_0) = -q(n-1)(U_0) \qquad U_0 \in \bigodot^q (T_{x_0}S^n),$$

while by (46), for $U_0 \in \bigodot^q(T_{x_0}S^n)_{\mathcal{H}}$,

$$\sum_r B^r(U_0) = -q(q-1)U_0 + 2\mathcal{C}(U_0)\odot \Xi = -q(q-1)U_0,$$

giving (60), to complete the proof. □

In particular for $p = 3$, using the decomposition (57) we obtain:

Corollary 5.2 . For $U_0 = u_0^1 \odot u_0^2 \odot u_0^3 \in \bigodot^3 T_{x_0}M$, with parallel translate $U_t \in \bigodot^3 T_{x_t}M$ along the Brownian paths, we have

$$\begin{aligned} \nabla^3(P_tf)(U_0) &= e^{-\frac{3}{2}(n+1)t}E\{\nabla^3 f(U_t)\} \\ &+ \frac{e^{-\frac{1}{2}(n-1)t}}{n+2}(1-e^{-(n+2)t})\mathbf{E}\{\nabla^3 f(\mathcal{C}(U_t)\odot \Xi_t)\}. \end{aligned} \tag{61}$$

5.1.2 The full derivative, p=3

Recall that with our use of Kobayashi & Nomizu's sign conventions the curvature $R^{\nabla^E} : TM \times TM \to \mathcal{L}(E;E)$ of a connection ∇^E on a vector bundle E over a manifold M is given by definition by

$$R^{\nabla^E}(u,v)S(x) = \nabla^E_u \nabla^E_V S - \nabla^E_v \nabla^E_U S - \nabla^E_{[U,V](x)} S \tag{62}$$
$$= (\nabla^E)^2 S(u,v) - (\nabla^E)^2 S(v,u) \tag{63}$$

for $u = U(x)$, $v = V(x)$ some $x \in M$, with U, V vector fields and S a section of E. To define $(\nabla^E)^2$ a torsion free connection on M is used.
For $E = TM$ with M Riemannian and $\nabla^E = \nabla$ the Levi-Civita connection, we write $R = R^{\nabla^E}$ as before. It is important to note that with the induced Levi-Civita connection on the cotangent bundle

$$R^{\nabla^{T^*M}}(u,v)\ell = -\ell \circ R(u,v) \quad \text{ for } \ell \in T^*_x M.$$

This can be seen by computing the Hessian of the function $\phi(W(-))$ for ϕ a one-form and W a vector field, and using its symmetry. More generally, if S is a section of $(\bigotimes^p TM)^*$ then

$$\begin{aligned}\nabla^2 S(u,v)(w^1\otimes ...\otimes w^p) - \nabla^2 S(v,u)(w^1\otimes ...\otimes w^p) &= \\ -S(x)\left(R(u,v)w^1\otimes w^2\otimes ... + ... + w^1\otimes w^2\otimes ...\otimes R(u,v)w^p\right) & \quad (64)\end{aligned}$$

for $u, v, w^1, ...w^p \in T_xM$.

Lemma 5.3 For $u, v, w \in T_xM$ and $f : M \to \mathbf{R}$

$$\nabla^2 df(u,v,w) = \nabla^2 df(u\odot v\odot w) + \frac{1}{3} df\,(R(v,u)w + R(w,u)v)\,. \quad (65)$$

Proof. By the symmetry of Hessians, $\nabla^2 df(u,v,w)$ is symmetric in v, w. Therefore

$$\nabla^2 df(u\odot v\odot w) = \frac{1}{3!}\nabla^2 df\,(2u\otimes v\otimes w + 2v\otimes u\otimes w + 2w\otimes u\otimes v)\,.$$

Taking $S = df$ in (64)

$$\nabla^2 df(v\otimes u\otimes w) = \nabla^2 df(u\otimes v\otimes w) - df(R(v,u)w)$$

and

$$\nabla^2 df(w\otimes u\otimes v) = \nabla^2 df(u\otimes w\otimes v) - df(R(w,u)v),$$

giving the result by the symmetry of $\nabla^2 df$ in the last two variables. □

As an example

Example 5.4

$$d\triangle f(x_0)(u) = \nabla^2 df(u\otimes \Xi) = \nabla^2 df(u\odot \Xi) - \frac{2}{3} df\left(\mathrm{Ric}^\sharp(u)\right) \quad (66)$$

which enables us to rewrite (61) as

$$\begin{aligned}&\nabla^3(P_t f)(U_0) = e^{-\frac{3}{2}(n+1)t}\mathbf{E}\{\nabla^3 f(U_t)\} + \\ &\quad \tfrac{e^{-\frac{1}{2}(n-1)t}}{n+2}(1-e^{-(n+2)t})\mathbf{E}\left\{\,\left(d\triangle f(\mathcal{C}(U_t)) + \tfrac{2n-2}{3} df(\mathcal{C}(U_t))\right)\right\}.\end{aligned}$$

Theorem 5.5 For $u_0, v_0, w_0 \in T_{x_0}S^n$ write $U_t = /\!/_t u_0 \otimes /\!/_t v_0 \otimes /\!/_t w_0$ for parallel translation along a Brownian motion from x_0. Then for a C^3 function $f : S^n \to \mathbf{R}$

$$\nabla^3 P_t f(U_0) = \mathbf{E}\big\{e^{-\frac{3}{2}(n+1)t}\nabla^3 f(U_t)$$
$$+ \tfrac{1}{n+2} e^{-\frac{1}{2}(n-1)t}(1 - e^{-(n+2)t}) d\triangle f(\mathcal{C}(U_t))$$
$$+ e^{-\frac{1}{2}(n-1)t}\left(1 - e^{-(n+2)t}\right) df\big(\tfrac{n}{n+2}\mathcal{C}(U_t) - \langle v_0, w_0\rangle /\!\!/_t u_0\big)\big\}.$$

In particular if u_0, v_0, w_0 are mutually perpendicular,

$$\nabla^3 P_t f(U_0) = e^{-\frac{3}{2}(n+1)t}\mathbf{E}\big\{\nabla^3 f(U_t)\big\}. \tag{67}$$

Proof. By formula (65)

$$\nabla^2 dP_t f(u_0, v_0, w_0) = \nabla^2 dP_t f(u_0 \odot v_0 \odot w_0) + \frac{1}{3} dP_t f(z_0)$$

for

$$\begin{aligned} z_0 &= R(v_0, u_0)w_0 + R(w_0, u_0)v_0 \\ &= \langle u_0, w_0\rangle v_0 + \langle u_0, v_0\rangle w_0 - 2\langle v_0, w_0\rangle u_0 \\ &= \mathcal{C}(U_0^{\odot}) - 3\langle v_0, w_0\rangle u_0 \end{aligned}$$

for our sphere.

Now $d(P_t f)(z_0) = \mathbf{E}\{e^{-\frac{1}{2}(n-1)t} df(/\!\!/_t z_0)\}$
and
$$\tfrac{1}{3} /\!\!/_t z_0 + \tfrac{2(n-1)}{3(n+2)}(1 - e^{-(n+2)t})\mathcal{C}(U_t^{\odot})$$
$$= -\langle v_0, w_0\rangle /\!\!/_t u_0 + \left(\tfrac{n}{n+2} - \tfrac{2(n-1)e^{-(n+2)t}}{3(n+2)}\right)\mathcal{C}(U_t^{\odot}).$$

The result follows by using the formula in Example 5.4 with $U_0 = u_0 \odot v_0 \odot w_0$, together with (65) to go back again to our non-symmetrised U_t.

□

Remark 5.6 Note that for all $p = 2, 3, ..n$, for a sphere S^n, formula (64) can be applied inductively to show that if $u^1, ..., u^p$ are mutually orthogonal then for any C^p function f

$$\nabla^p f(u^1, u^2, ..., u^p) = \nabla^p df(u^1 \odot u^2 ... \odot u^p)$$

For $y \in S^n$ set

$$||\nabla^p f||_{o.n.}(y) = \sup\big\{|\nabla^p f(v^1, ..., v^p)|, \text{orthonormal } v^1, ..., v^p \in T_y S^n\big\}.$$

Applying (60) we obtain the pointwise semi-group domination,

$$||\nabla^p P_t f||_{o.n.} \le e^{-\frac{p}{2}(n+p-2)t} P_t(||\nabla^p f||_{o.n.}) \tag{68}$$

5.2 More general diffusion semi-groups

5.2.1 Heat semigroups on functions and forms on compact Riemannian symmetric spaces

The method described here for spheres should go over directly for the heat semi-group of a compact Riemannian symmetric space. The representation theory involved may be more complicated. It should extend similarly to formulae for derivatives of heat semigroups for forms. See also the alternative approach suggested in Section 4.1.

An additional first order term can be included by combining this method with the more standard method of filtering out redundant noise, but the formulae will be more complicated unless the term comes from a Killing vector field.

5.2.2 General diffusions on manifolds; derivatives of induced semigroups on functions, forms, jets etc

For a heat equations on a general compact Riemannian manifold M a similar approach can be followed but replacing our bundle $p : SO(n+1) \to S^n$ by a bundle $p : \mathrm{Diff}(M) \to M$ where $\mathrm{Diff}(M)$ is a suitable group of diffeomorphisms of M and p the evaluation map at a base point $x_0 \in M$. This can be considered as a principal bundle with group those elements $\mathrm{Diff}_{x_0}(M)$ of $\mathrm{Diff}(M)$ which fix x_0. Our stochastic flow can be considered as a process on $\mathrm{Diff}(M)$ and has a skew product decomposition generalising that described in Section 3.1. See [11], [13]. This gives rise to formulae like (19) and its higher order analogues, but with lower order derivative terms because the second derivative of the flow will not generally vanish: for symmetric spaces it vanished because we had a flow of isometries.

In this case we should look at k-jets of $P_t f$ rather than k-th order covariant derivatives. However the parallel translation in our formula may now not be metric preserving for any metric on the bundle of k-th order tangent vectors (essentially k-th order differential operators), or its dual, the k-jets. This makes uniform estimates difficult to obtain. For tensor bundles, associated to the frame bundle of M, such as $\bigotimes^k TM$, an SDE can be chosen for our diffusion so that its conditioned flow determines Levi-Civita parallel translation, and so is metric preserving, see [10]. For k-th order tangent vectors the question is crucial but open.

A result by Mendes & Redeschi, [24], shows that we cannot have the Levi-Civita connection on tensor bundles induced by an SDE for Brownian motion which has a solution flow of isometries except in the case we have been discussing for symmetric spaces. They call an SDE for Brownian motion which induces the Levi-Civita connection a *virtual immersion*. Ming Liao [22] has somewhat related negative results; in particular there are no isometric stochastic flows of Brownian motions on a Riemannian symmetric space of non-compact type. However Liao shows in [23] that for $n > 3$ there are a continuum of them on S^n.

In [14], in preparation, this set up is extended to a wide class of semigroups induced on sections of natural bundles, such as jet bundles, by a sum of squares representation of a diffusion operator on M. The operator need not be elliptic or hypo-elliptic but for the method to work smoothly it should be *cohesive* in the sense of [13]. However, the crucial question of finding stochastic flows inducing metric connections on natural bundles remains open, to my knowledge.

5.3 Questions

- [Berger's spheres.] It would be interesting to see how the derivatives of the heat semigroup change as the sphere gets smoothly deformed, for example for Berger's spheres which still retain a lot of symmetry; see [15] and for a stochastic analytical discussion and more references [21].
- [Different symmetric space structures.] The same Riemannian manifold can have different symmetric space structures. For example the 3-sphere is also a Lie group and so has the symmetric space structure with group $S^3 \times S^3$ acting by $(g^1, g^2).a = g^1 a (g^2)^{-1}$ for $a \in S^3$. Can such different structures give different derivative formulae?
- [Non-compact type] Are there corresponding formulae for symmetric spaces of non-compact type? In particular for hyperbolic space. As remarked above, [22] and [24] imply that the use of isometric flows as here does not go over immediately to the non-compact case.

References

[1] D. Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 43–75. World Sci. Publ., River Edge, NJ, 1997.

[2] D. Bakry and Michel Émery. Propaganda for Γ_2. In *From local times to global geometry, control and physics (Coventry, 1984/85)*, volume 150 of *Pitman Res. Notes Math. Ser.*, pages 39–46. Longman Sci. Tech., Harlow, 1986.

[3] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[4] Peter Baxendale. Brownian motions in the diffeomorphism group. I. *Compositio Math.*, 53(1):19–50, 1984.

[5] Arthur L. Besse. *Einstein manifolds*. Classics in Mathematics. Springer-Verlag, Berlin, 2008. Reprint of the 1987 edition.

[6] Renato G. Bettiol and Ricardo A. E. Mendes. Sectional curvature and Weitzenböck formulae. *arXiv e-prints*, page arXiv:1708.09033, Aug 2017.

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

[8] K. D. Elworthy. Generalised Weitzenböck formulae for differential operators in Hörmander form. In *Stochastic partial differential equations and related fields*, volume 229 of *Springer Proc. Math. Stat.*, pages 319–331. Springer, Cham, 2018.

[9] K. D. Elworthy and W. S. Kendall. Factorization of harmonic maps and Brownian motions. In *From local times to global geometry, control and physics (Coventry, 1984/85), Pitman Res. Notes Math. Ser., 150*, pages 75–83. Longman Sci. Tech., Harlow, 1986.

[10] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows, Lecture Notes in Mathematics 1720*. Springer, 1999.

[11] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li. Equivariant diffusions on principal bundles. In *Stochastic analysis and related topics in Kyoto*, volume 41 of *Adv. Stud. Pure Math.*, pages 31–47. Math. Soc. Japan, Tokyo, 2004.

[12] 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.

[13] K. David Elworthy, Yves Le Jan, and Xue-Mei Li. *The geometry of filtering*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2010.

[14] K.D. Elworthy. Generalised weitzenböck formulae for differential operators in hörmander form ii: Natural bundles and higher order derivative formulae. *In preparation*, 2019.

[15] P. M. Gadea and J. A. Oubiña. Homogeneous Riemannian structures on Berger 3-spheres. *Proc. Edinb. Math. Soc. (2)*, 48(2):375–387, 2005.

[16] Roe Goodman and Nolan R. Wallach. *Representations and invariants of the classical groups*, volume 68 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1998.

[17] J. E. Humphreys. *Introduction to Lie algebras and representation theory*, volume 9 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1978. Second printing, revised.

[18] S. Kobayashi and K. Nomizu. *Foundations of differential geometry, Vol. II*. Interscience Publishers, 1969.

[19] Xue-Mei Li. Hessian formulas and estimates for parabolic Schrödinger operators. *arXiv e-prints*, page arXiv:1610.09538, Oct 2016.

[20] Xue-Mei Li. Doubly damped stochastic parallel translations and Hessian formulas. In *Stochastic partial differential equations and related fields*, volume 229 of *Springer Proc. Math. Stat.*, pages 345–357. Springer, Cham, 2018.

[21] Xue-Mei Li. Homogenisation on homogeneous spaces. *J. Math. Soc. Japan*, 70(2):519–572, 2018. With an appendix by Dmitriy Rumynin.

[22] Ming Liao. The existence of isometric stochastic flows for Riemannian Brownian motions. In *Diffusion processes and related problems in analysis, Vol. II (Charlotte, NC, 1990)*, volume 27 of *Progr. Probab.*, pages 95–109. Birkhäuser Boston, Boston, MA, 1992.

[23] Ming Liao. Isometric stochastic flows on spheres. *J. Theoret. Probab.*, 12(2):475–488, 1999.

[24] Ricardo A. E. Mendes and Marco Radeschi. Virtual immersions and a characterization of symmetric spaces. *Ann. Global Anal. Geom.*, 55(1):43–53, 2019.

[25] J. R. Norris. Path integral formulae for heat kernels and their derivatives. *Probab. Theory Related Fields*, 94(4):525–541, 1993.

[26] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[27] James Thompson. Approximation of Riemannian measures by Steins method. *In preparation*, 2019.