Czechoslovak Mathematical Journal

Ľubomír Baňas; Zdzisław Brzeźniak; Mikhail Neklyudov; Martin Ondreját; Andreas Prohl
Ergodicity for a stochastic geodesic equation in the tangent bundle of the 2D sphere

Czechoslovak Mathematical Journal, Vol. 65 (2015), No. 3, 617–657

Persistent URL: `http://dml.cz/dmlcz/144435`

Czechoslovak Mathematical Journal, 65 (140) (2015), 617–657

ERGODICITY FOR A STOCHASTIC GEODESIC EQUATION IN THE TANGENT BUNDLE OF THE 2D SPHERE

Ľubomír Baňas, Bielefeld, Zdzisław Brzeźniak, York,
Mikhail Neklyudov, Pisa, Martin Ondreját, Praha,
Andreas Prohl, Tübingen

(Received November 1, 2013)

Abstract. We study ergodic properties of stochastic geometric wave equations on a particular model with the target being the 2D sphere while considering only solutions which are independent of the space variable. This simplification leads to a degenerate stochastic equation in the tangent bundle of the 2D sphere. Studying this equation, we prove existence and non-uniqueness of invariant probability measures for the original problem and obtain also results on attractivity towards an invariant measure. We also present a structure-preserving numerical scheme to approximate solutions and provide computational experiments to motivate and illustrate the theoretical results.

Keywords: geometric stochastic wave equation; stochastic geodesic equation; ergodicity; attractivity; invariant measure; numerical approximation

MSC 2010: 58J65, 60H10, 60H35, 65C30, 60J60, 65C20, 37A25, 60H15

1. Introduction

Wave equations subject to random excitations have been largely studied in the last fourty years for their applications in physics, relativistic quantum mechanics or oceanography, see e.g. [12]–[15], [18], [19], [27]–[30], [34]–[40]. The mathematical research has paid attention predominantly to stochastic wave equations whose solutions took values in Euclidean spaces, however many physical theories and models in modern physics such as harmonic gauges in general relativity, nonlinear σ-models

The research of the fourth named author was supported by the GAČR Grant number P201/10/0752. This work was partially supported by the Australian Research Council Discovery Project DP120101886.

in particle systems, electro-vacuum Einstein equations or Yang-Mills field theory require the target space of the solutions to be a Riemannian manifold, see e.g. [21] and [41]. Stochastic wave equations with values in Riemannian manifolds were first studied in [11] (see also [10]) where existence and uniqueness of global strong solutions were proved for equations defined on the one-dimensional Minkowski space $\mathbb{R}^{1+1}$ and arbitrary Riemannian manifold. Later, in [9], global existence was proved for equations on a general Minkowski space $\mathbb{R}^{1+d}$ with the target space being restricted to homogeneous spaces (for instance, a sphere) and, in [10], global existence of weak solutions was proved for equations on $\mathbb{R}^{1+1}$ with an arbitrary target. The last two works admitted rougher noises than in [11], but for the price of not dealing with the question of uniqueness and of worse spatial regularity of the solutions.

In the present paper, we intend to open the door and enter into the study of ergodic properties of solutions of these equations. In particular, we are interested in existence and uniqueness (or multitude) of invariant measures of the Markov semigroup associated with solutions of a stochastic geometric equation and we also want to address the questions of ergodic properties and of the rates of convergence to an attracting law, if there is any.

This goal, however, seems to be fairly complicated and too ambitious to achieve at once, hence we will proceed *a minori ad majus* and study just space independent solutions of a damped stochastic geometric wave equation in the 2D sphere. This particular exemplary equation is, in our opinion, quite illustrative to understand what one can expect in the general case. In this way, the stochastic equation will reduce to a degenerate second order stochastic differential equation with values in the tangent bundle $T\mathbb{S}^2$. We will prove that there exist plenty of invariant measures and that the system always converges in total variation to a limit law. If we however restrict the state space to a suitable submanifold in $T\mathbb{S}^2$ then there exists just one unique invariant measure (the normalized surface measure on this submanifold) which attracts every initial distribution in total variation with an exponential rate.

A further goal of this paper is to construct a numerical scheme for solving a class of SDEs on manifolds—the geodesic equation on the sphere $\mathbb{S}^2$ with stochastic forcing. A convergent discretization in space and time for a first order stochastic Landau-Lifshitz-Gilbert equation where solutions take values in $\mathbb{S}^2$ is proposed in [3], [4]; the present case is however very different, and the structure preserving discretization given in Section 6.1 is inspired by the "discrete Lagrange multiplier" strategy developed in [6].

Computational examples for the stochastic geodesic equation on the sphere are provided in Section 6.2 to motivate long-time asymptotics, which is then studied analytically in the later sections.

2. Notation and conventions

If Y is a topological space, we will denote by $B_b(Y)$ the space of real bounded Borel functions on Y, by $C_b(Y)$ the space of real bounded continuous functions on Y, by $\mathscr{B}(Y)$ the Borel σ-algebra over Y. We will work on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ equipped with a filtration $(\mathscr{F}_t)$ such that $\mathscr{F}_0$ contains all $\mathbb{P}$-negligible sets in $\mathscr{F}$ and W will be a standard $(\mathscr{F}_t)$-Wiener process. Throughout this paper, all initial conditions are assumed to be $\mathscr{F}_0$-measurable.

3. The problem

Let M be a compact m-dimensional Riemannian manifold embedded in a Euclidean space $\mathbb{R}^n$. Denote by T_pM the tangent space at $p \in M$, by $N_pM = (T_pM)^\perp$ the normal space at $p \in M$, by $TM = \bigcup_{p\in M} T_pM$ and $T^kM = \bigcup_{p\in M} (T_pM)^k$ the tangent bundle and the k-tangent bundle of M, respectively, by $S_p\colon T_pM \times T_pM \to N_pM$, $p \in M$ the second fundamental form of M in $\mathbb{R}^n$, and let W be, for simplicity, a one-dimensional Wiener process. According to [11], the general Cauchy problem for a stochastic geometric wave equation has the form

$$\mathrm{d}u_t = \left(\Delta u - \sum_{i=1}^m S_u(u_{x_i}, u_{x_i}) + S_u(u_t, u_t) + F_u(Du)\right)\mathrm{d}t + G_u(Du)\,\mathrm{d}W \tag{3.1}$$

$$u \in M, \quad (u(0), u_t(0)) \in TM \tag{3.2}$$

where F is a drift, G a diffusion and Du denotes the $(m+1)$-tuple $(u_t, u_{x_1}, \ldots, u_{x_m})$ in the equation (3.1). For the equation to make sense, it is required that $F\colon T^{m+1}M \to TM$ and $G\colon T^{m+1}M \to TM$ are Borel measurable and that $F_p(X_0, \ldots, X_m)$ and $G_p(X_0, \ldots, X_m)$ belong to the tangent space T_pM for every $p \in M$ and every $X_0, \ldots, X_m \in T_pM$.

In case M is the unit sphere in $\mathbb{R}^3$ then the second fundamental form satisfies $S_p(X, Y) = -\langle X, Y\rangle p$. If we set $F_p(X_0, X_1, X_2) = -X_0/2$, $G_p(X_0, X_1, X_2) = p \times X_0$ then the equation (3.1) with the constraints (3.2) has the form

$$\mathrm{d}u_t = \left(\Delta u + (|\nabla u|^2 - |u_t|^2)u - \frac{1}{2}u_t\right)\mathrm{d}t + u \times u_t\,\mathrm{d}W, \quad |u| = 1,\ u(0) \perp u_t(0). \tag{3.3}$$

If we consider just space independent solutions, i.e. solutions independent of the spatial variables, then (3.3) reduces to an Itô SDE

$$\mathrm{d}u' = \left(-|u'|^2 u - \frac{1}{2}u'\right)\mathrm{d}t + (u \times u')\,\mathrm{d}W, \quad |u| = 1,\ u(0) \perp u'(0), \tag{3.4}$$

or, equivalently, to a Stratonovich SDE

$$\mathrm{d}u' = -|u'|^2 u\,\mathrm{d}t + (u \times u') \circ \mathrm{d}W, \quad |u| = 1,\ u(0) \perp u'(0), \tag{3.5}$$

which is the stochastic geodesic equation for the unit sphere.[1] Let us rewrite (3.5) to two equations of first order equations

$$\mathrm{d}z = f(z)\,\mathrm{d}t + g(z) \circ \mathrm{d}W, \quad z \in T\mathbb{S}^2, \quad z(0) \in T\mathbb{S}^2, \tag{3.6}$$

where $T\mathbb{S}^2 \subseteq \mathbb{R}^6$ is the tangent bundle of $\mathbb{S}^2$, i.e. $T\mathbb{S}^2 = \{(u, v)\colon |u| = 1,\ u \perp v\}$, and

$$z = \begin{pmatrix} u \\ v \end{pmatrix}, \quad f(z) = \begin{pmatrix} v \\ -|v|^2 u \end{pmatrix}, \quad g(z) = \begin{pmatrix} 0 \\ u \times v \end{pmatrix}. \tag{3.7}$$

Remark 3.1. Observe that the restrictions of f and g to $T\mathbb{S}^2$ are vector fields on the manifold $T\mathbb{S}^2$. Hence (3.6) is a correctly defined stochastic differential equation on the manifold $T\mathbb{S}^2$, cf. [26], Chapter V.

The equation (3.4) and its equivalent formulations (3.5), (3.6) will be the *object of study* of the present paper. It is also important to realize while reading the paper that (3.4) is a particular case of the stochastic geometric wave equation (3.1)–(3.2).

4. Basic properties of solutions of the SDE

We will study existence of global solutions, dependence on initial conditions, some further qualitative properties of solutions of the equation (3.6) and the Feller property of the associated Markov semigroup.

4.1. Global existence. The nonlinearities of the equation (3.6) are locally Lipschitz on $\mathbb{R}^6$, hence, by the standard existence result (see e.g. [26], Lemma 2.1), the equation (3.6) considered without the constraint,

$$\mathrm{d}z = f(z)\,\mathrm{d}t + g(z) \circ \mathrm{d}W, \quad z(0) \in T\mathbb{S}^2, \tag{4.1}$$

has a unique local solution z in $\mathbb{R}^6$ defined up to an explosion time $\tau > 0$, i.e.,

$$\limsup_{t \uparrow \tau} |z(t)| = \infty \quad \text{a.s. on } [\tau < \infty]. \tag{4.2}$$

[1] The geodesic equation for the unit sphere has the form $u'' = -|u'|^2 u$, $|u| = 1$, $u'(0) \perp u(0)$.

Proposition 4.1. *The solution to* (4.1) *is unique, global and satisfies* $z = (u, v) \in T\mathbb{S}^2$*, i.e. it is a solution to the equation* (3.6)*. Moreover,* $|v(t)| = |v(0)|$ *for every* $t \geqslant 0$ *a.s.*

Proof. Applying the Itô formula to $|u|^2$, we obtain that $\phi = |u|^2 - 1$ satisfies a.s. on $[0, \tau)$ the ODE

$$\phi'' = -2|v|^2\phi - \frac{1}{2}\phi', \quad \phi(0) = 0, \quad \phi'(0) = 0. \tag{4.3}$$

Hence, by the uniqueness of the solutions to the equation (4.3), we obtain that $\phi = 0$ on $[0, \tau)$, consequently, $|u| = 1$ on $[0, \tau)$ a.s. In particular, differentiating $|u|^2 = 1$, we obtain that $u \perp v = 0$ on $[0, \tau)$ a.s. Now, applying the Itô formula to $|v|^2$, we obtain that $\varphi = |v|^2$ satisfies on $[0, \tau)$ a.s. the equation

$$\varphi' = -(1 + 2\langle u, v\rangle)|v|^2 + |u \times v|^2.$$

The right hand side equals

$$-(1 + 2\langle u, v\rangle)|v|^2 + |u|^2|v|^2 - \langle u, v\rangle^2 = 0$$

as $u \perp v$ and $|u| = 1$ a.s. Hence $|v|$ is pathwise constant. In particular, $\tau = \infty$ a.s. by (4.2). □

4.2. The Markov and the Feller properties. Define $Y = \mathbb{R}^n$. It is well known that if $\widetilde{f}$, $\widetilde{g}$ are C^∞ vector fields on $\mathbb{R}^n$ with a compact support and u^ξ denotes the solution of the equation

$$\mathrm{d}X = \widetilde{f}(X)\,\mathrm{d}t + \widetilde{g}(X) \circ \mathrm{d}W, \quad X(0) = \xi \tag{4.4}$$

for an $\mathscr{F}_0$-measurable Y-valued random variable ξ then the solutions of the equation (4.4) possess the Markov property and define a Feller semigroup[2] on Y by which we mean that

(a) the transition function

$$q_{t,x}(A) = \mathbb{P}[u^x(t) \in A], \quad t \geqslant 0,\ x \in Y,\ A \in \mathscr{B}(Y)$$

is jointly measurable in $(t, x) \in [0, \infty) \times Y$ for every $A \in \mathscr{B}(Y)$,

[2] We allow here a little inaccuracy. More precisely, the semigroup is defined on the space of bounded Borel functions on Y.

(b) the endomorphisms on $B_b(Y)$

$$Q_t\varphi(x) = \mathbb{E}\varphi(u^x(t)), \quad t \geqslant 0,\ x \in Y,\ \varphi \in B_b(Y)$$

possess the semigroup property, i.e., $Q_t \circ Q_s = Q_{t+s}$ for every $t, s \geqslant 0$,

(c) $Q_t\varphi$ is continuous on Y whenever $t \geqslant 0$ and $\varphi \in C_b(Y)$,

(d) $\mathbb{E}[\varphi(u^\xi(t)); \mathscr{F}_s] = (Q_{t-s}\varphi)(u^\xi(s))$ holds a.s. for every $\varphi \in B_b(Y)$, $0 \leqslant s \leqslant t$ and an $\mathscr{F}_0$-measurable Y-valued random variable ξ,

see e.g. [17], Section 9.2.1. In fact, (a) and (c) follow simply from the fact that

$$Q_t\varphi(x) \text{ is jointly continuous in } (t, x) \text{ on } [0, \infty) \times Y \text{ if } \varphi \in C_b(Y), \tag{4.5}$$

see again [17], Section 9.2.1, for the proof of (4.5), and the semigroup property (b) follows from the Markov property (d).

Moreover, if $\varphi \in C^2(Y)$ with derivatives of order $0, 1, 2$ bounded then

$$\varrho(t, x) = Q_t\varphi(x) \text{ belongs to } C^{1,2}([0, \infty) \times Y) \tag{4.6}$$

with ϱ, $\partial\varrho/\partial t$, $\partial\varrho/\partial x_i$, $\partial^2\varrho/\partial x_i\partial x_j$ bounded for every $i, j \in \{1, \dots, n\}$ and it is a solution to the backward Kolmogorov equation

$$\frac{\partial U}{\partial t} = \sum_{i=1}^{n} \widetilde{f}_i \frac{\partial U}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \widetilde{g}_i \frac{\partial}{\partial x_i}\Big(\widetilde{g}_j \frac{\partial U}{\partial x_j}\Big), \quad U(0, x) = \varphi(x) \text{ for every } x \in Y \tag{4.7}$$

unique in the class $C^{1,2}([0, \infty) \times Y)$, see e.g. [17], Section 9.3.

Unfortunately, the coefficients of the equation (3.6) are not compactly supported so we cannot simply conclude that the solutions of (3.6) possess the Markov property and define a Feller semigroup in the sense (a)–(d) above. Yet, it is true, as it will be shown below.

Notation 4.2. From now on, z^ξ denotes the solution of (3.6) with the initial condition ξ, $p_{t,x}(A) = \mathbb{P}[z^x(t) \in A]$ and $P_t\varphi(x) = \mathbb{E}\varphi(z^x(t))$ are defined for $\varphi \in B_b(T\mathbb{S}^2)$, $t \geqslant 0$, $x \in T\mathbb{S}^2$ and $A \in \mathscr{B}(T\mathbb{S}^2)$.

Proposition 4.3. *The solutions of* (3.6) *satisfy the Markov property and define a Feller semigroup on $T\mathbb{S}^2$. In fact, $P_t\varphi(x)$ is jointly continuous in (t, x) on $[0, \infty) \times T\mathbb{S}^2$ for every $\varphi \in C_b(T\mathbb{S}^2)$ and*

$$\mathbb{E}[\varphi(z^\xi(t)); \mathscr{F}_s] = (P_{t-s}\varphi)(z^\xi(s)) \quad \text{a.s.}$$

holds for every $\varphi \in B_b(T\mathbb{S}^2)$, $0 \leqslant s \leqslant t$ and every initial $T\mathbb{S}^2$-valued initial condition ξ.

P r o o f. Let us prove the joint continuity assertion first. Assume that $(t_n, x_n) \to (t, x)$ in $[0,\infty) \times T\mathbb{S}^2$ and let $\sup_n |x_n| \leqslant l$. Let $\widetilde{f}$, $\widetilde{g}$ be compactly supported C^∞ vector fields on $\mathbb{R}^6$ such that $f = \widetilde{f}$ and $g = \widetilde{g}$ on the ball of radius l in $\mathbb{R}^6$. Now $|z^{x_n}(t)| = |x_n| \leqslant l$ and $|z^x(t)| = |x| \leqslant l$ holds for every $t \geqslant 0$ a.s. by Proposition 4.1 and hence z^{x_n}, z^x are also solutions to the equation

$$\mathrm{d}X = \widetilde{f}(X)\,\mathrm{d}t + \widetilde{g}(X) \circ \mathrm{d}W.$$

So, if $\varphi \in C_b(T\mathbb{S}^2)$ and $\widetilde{\varphi} \in C_b(\mathbb{R}^6)$ is any extension of φ (which always exists by the Tietze theorem) then

$$\lim_{n\to\infty} P_{t_n}\varphi(x_n) = \lim_{n\to\infty} \mathbb{E}\widetilde{\varphi}(z_n(t_n)) = \mathbb{E}\widetilde{\varphi}(z(t)) = P_t\varphi(x)$$

by (4.5).

To prove the Markov property, let $\xi = (\xi^1, \xi^2)$ be a $T\mathbb{S}^2$-valued initial condition and define $\xi_k = (\xi^1, \xi^2 \mathbf{1}_{[|\xi^2| \leqslant k]})$. Then ξ_k take values in $T\mathbb{S}^2$ and by Proposition 4.1, $|z^{\xi_k}(t)| = |\xi_k| \leqslant \sqrt{1+k^2}$. Let $\widetilde{f}$, $\widetilde{g}$ be compactly supported C^∞ vector fields on $\mathbb{R}^6$ such that $f = \widetilde{f}$ and $g = \widetilde{g}$ on the ball of radius $\sqrt{1+k^2}$ in $\mathbb{R}^6$ and define $Q_t\phi(y) = \mathbb{E}\phi(u^y(t))$ for $\phi \in B_b(\mathbb{R}^6)$, $y \in \mathbb{R}^6$, $t \geqslant 0$ and u^y the solutions to $\mathrm{d}X = \widetilde{f}(X)\,\mathrm{d}t + \widetilde{g}(X) \circ \mathrm{d}W$, $X(0) = y$. By the first part of the proof, we know that $P_t\varphi(x) = Q_t\widetilde{\varphi}(x)$ holds for every $x \in T\mathbb{S}^2$ such that $|x| \leqslant \sqrt{1+k^2}$, $\varphi \in B_b(T\mathbb{S}^2)$, $\widetilde{\varphi} \in B_b(\mathbb{R}^6)$, $\varphi = \widetilde{\varphi}$ on $T\mathbb{S}^2$ and $t \geqslant 0$.

Now $z^{\xi_k} = u^{\xi_k}$ and if we define $A_k = [|\xi^2| \leqslant k]$ and $\widetilde{\varphi} \in B_b(\mathbb{R}^6)$ extends $\varphi \in B_b(T\mathbb{S}^2)$ then

$$\begin{aligned}
\mathbf{1}_{A_k}\mathbb{E}[\varphi(z^\xi(t)); \mathscr{F}_s] &= \mathbb{E}[\mathbf{1}_{A_k}\varphi(z^\xi(t)); \mathscr{F}_s] \\
&= \mathbb{E}[\mathbf{1}_{A_k}\varphi(z^{\xi_k}(t)); \mathscr{F}_s] = \mathbf{1}_{A_k}\mathbb{E}[\varphi(z^{\xi_k}(t)); \mathscr{F}_s] \\
&= \mathbf{1}_{A_k}\mathbb{E}[\widetilde{\varphi}(u^{\xi_k}(t)); \mathscr{F}_s] = \mathbf{1}_{A_k}(Q_{t-s}\widetilde{\varphi})(u^{\xi_k}(s)) \\
&= \mathbf{1}_{A_k}(P_{t-s}\varphi)(z^{\xi_k}(s)) = \mathbf{1}_{A_k}(P_{t-s}\varphi)(z^\xi(s)) \quad \text{a.s.}
\end{aligned}$$

by the Markov property of solutions of the equation (4.4). To obtain the result, let $k \to \infty$. $\square$

5. Multitude of invariant measures

Now we are ready to prove that the equation (3.6) and, consequently, also the equation (3.3) have many invariant measures due to the geometric nature of the equation.

Definition 5.1. Let Y be a Polish space, $r_{t,x}(\cdot)$ probability measures on $\mathscr{B}(Y)$ indexed by $(t,x) \in [0,\infty) \times Y$ such that $r_{t,x}(A)$ is jointly measurable in (t,x) on $[0,\infty) \times Y$ for every $A \in \mathscr{B}(Y)$ and the operators

$$R_t\varphi(x) = \int_Y \varphi \,\mathrm{d}r_{t,x}, \quad \varphi \in B_b(Y),\ t \geqslant 0$$

possess the semigroup property on $B_b(Y)$. We introduce the adjoint endomorphisms R_t^* acting on the space of probability measures on $\mathscr{B}(Y)$:

$$R_t^*\nu(A) = \int_Y r_{t,x}(A)\,\mathrm{d}\nu(x), \quad t \geqslant 0,\ A \in \mathscr{B}(Y).$$

A probability measure ν on $\mathscr{B}(Y)$ is called *invariant* provided that

$$R_t^*\nu = \nu \quad \text{for all } t \geqslant 0 \text{ and } A \in \mathscr{B}(Y).$$

A probability measure on $\mathscr{B}(Y)$ is called *ergodic* provided that it is an extreme point in the convex set of invariant probability measures.

Remark 5.2. To make the meaning of the above definition clear, apply the Markov property in Proposition 4.3 with $s = 0$. If ξ is an $\mathscr{F}_0$-measurable $T\mathbb{S}^2$-valued random variable with a distribution ν then $P_t^*\nu$ is the law of $z^\xi(t)$.

At this moment, we introduce subsets of the tangent bundle $T\mathbb{S}^2$

$$M_r = \{(u,v) \in T\mathbb{S}^2\colon\ |v| = r\}, \quad r \geqslant 0. \tag{5.1}$$

Remark 5.3 (Invariance). If $r > 0$ and $x \in M_r$ then $z^x(t) \in M_r$ for every $t \geqslant 0$ a.s. If $|u| = 1$ then $z^{(u,0)}(t) = (u,0)$ for every $t \geqslant 0$ a.s. These conclusions follow directly from Proposition 4.1.

Corollary 5.4. *Let $r > 0$. For every $t \geqslant 0$, P_t is an endomorphism on $B_b(M_r)$.*

Corollary 5.5. *Let $x \in M_0$. Then δ_x is an invariant measure.*

We are going to prove that there is more to see than what was disclosed by Corollary 5.5, on the sets M_r as far as invariant measures are concerned.

Remark 5.6. Observe that, for every $r > 0$, the mappings f and g in (3.7) are vector fields on the manifold M_r. In particular, Proposition 4.1 is now a direct consequence of the general result [26], Theorem 1.1, Chapter V.

In view of Remark 5.6, we can introduce the following second order differential operator on M_r.

Definition 5.7. Define the second order differential operator

$$\mathcal{A}\varphi = f(\varphi) + \frac{1}{2}g(g(\varphi)) \tag{5.2}$$

for $\varphi \in C^2(M_r)$ for $r > 0$.

The next result follows from [26], Chapter V, Theorem 3.1, but, rather than checking the assumptions in [26], Chapter V, Section 3, we will give, for our purposes and for the reader's comfort, the short proof here.

Proposition 5.8. *Let $r > 0$ and let $\varphi \in C^2(M_r)$. Then $\varrho(t,x) = P_t\varphi(x)$ belongs to $C^{1,2}([0,\infty) \times M_r)$ and satisfies the backward Kolmogorov equation*

$$\frac{\partial \varrho}{\partial t} = \mathcal{A}\varrho \quad \text{on } [0,\infty) \times M_r, \ \varrho(0,\cdot) = \varphi. \tag{5.3}$$

On the other hand, if $\varrho \in C^{1,2}([0,\infty) \times M_r)$ satisfies (5.3) then $\varrho(t,x) = P_t\varphi(x)$ on $[0,\infty) \times M_r$.

Proof. Let $k \in \mathbb{N}$ and let $\widetilde{f}$ and $\widetilde{g}$ be C^∞ vector fields on $\mathbb{R}^6$ such that $\widetilde{f} = f$ and $\widetilde{g} = g$ on the centered ball in $\mathbb{R}^6$ of radius $R = \sqrt{1+r^2}$. Denote by u^x the solution of $\mathrm{d}X = \widetilde{f}(X)\,\mathrm{d}t + \widetilde{g}(X) \circ \mathrm{d}W$, $X(0) = x$ and let Q_t be the associated Markov operators. Let $\widetilde{\varphi} \in C^2(\mathbb{R}^6)$ be a compactly supported extension of φ. Then $z^x = u^x$ for every $x \in M_r$ by Proposition 4.1, $J(t,x) = Q_t\widetilde{\varphi}(x) \in C^{1,2}([0,\infty) \times \mathbb{R}^6)$ by (4.6), hence $J(t,x) = \varrho(t,x)$ for $(t,x) \in [0,\infty) \times M_r$. In particular, $\varrho \in C^{1,2}([0,\infty) \times M_r)$ and (5.3) holds by (4.7).

To prove the converse assertion, extend ϱ to a function in $C^{1,2}([0,\infty) \times \mathbb{R}^6)$, let $t > 0$ and apply the Itô formula to $\varrho(t-r, z^x(r))$ for $r \in [0,t]$, obtaining

$$\varphi(z^x(t)) = \varrho(0, z^x(t)) = \varrho(t,x) + \int_0^t g(\varrho)(t-r, z^x(r))\,\mathrm{d}W.$$

Taking expectations on both sides yields the claim. □

The next assertion is obvious if $Q \in \mathbb{R}^3 \otimes \mathbb{R}^3$ is a unitary matrix with $\det Q = 1$ due to the invariance of the equation (3.6) for positively oriented unitary matrices. But it also holds if $\det Q = -1$. To prove this, we are going to use the uniqueness of the solutions of the backward Kolmogorov equation.

Corollary 5.9. *Let Q be a 3×3-unitary matrix. Denote $\widetilde{Q} = \operatorname{diag}[Q, Q] \in \mathbb{R}^6 \otimes \mathbb{R}^6$. Then*

$$p(t, \widetilde{Q}x, A) = p(t, x, [\widetilde{Q} \in A])$$

holds for every $(t, x) \in [0, \infty) \times M_r$, every $A \in \mathscr{B}(M_r)$ and every $r > 0$.

Proof. Let $\varphi \in C^2(M_r)$ and define $\varrho(t, x) = P_t\varphi(x)$ for $(t, x) \in [0, \infty) \times M_r$. Then ϱ verifies (5.3). Now define $\varrho(t, x) = \varrho(t, \widetilde{Q}x)$ for $(t, x) \in [0, \infty) \times M_r$, which we can do since $\widetilde{Q}$ is a diffeomorphism on M_r. Then $\varrho \in C^{1,2}([0, \infty) \times M_r)$ and

$$\frac{\partial \varrho}{\partial t}(t, x) - \mathcal{A}\varrho(t, x) = \frac{\partial \varrho}{\partial t}(t, \widetilde{Q}x) - \mathcal{A}\varrho(t, \widetilde{Q}x) = 0 \quad \text{on } [0, \infty) \times M_r, \ \varrho(0, \cdot) = \varphi(\widetilde{Q}\cdot).$$

So, from the uniqueness part of Proposition 5.8, we obtain that

$$P_t\varphi(\widetilde{Q}x) = P_t(\varphi \circ \widetilde{Q})(x) \quad \text{on } [0, \infty) \times M_r. \tag{5.4}$$

By density of $C^2(M_r)$ in $C(M_r)$ we get that (5.4) holds for every $\varphi \in C(M_r)$ and consequently for every $\varphi \in B_b(M_r)$. □

Now we are ready to describe some analytic properties of the Markov semigroup (P_t) on M_r.

Theorem 5.10. *Let $r > 0$. Then (P_t) is a C_0-semigroup on $C(M_r)$, $P_t[C^2(M_r)] \subseteq C^2(M_r)$, $C^2(M_r)$ is contained in the domain of the infinitesimal generator A of (P_t) and $A = \mathcal{A}$ on $C^2(M_r)$.*

Proof. The C_0 property follows from the joint continuity in Proposition 4.3 and the invariance of $C^2(M_r)$ under the mappings P_t, $t \geqslant 0$, from Proposition 5.8. By the Itô formula,

$$P_t\varphi(x) = \varphi(x) + \int_0^t P_s(\mathcal{A}\varphi)(x)\,\mathrm{d}s, \quad t \geqslant 0, \ x \in M_r,$$

so φ belongs to the domain of the infinitesimal generator A of (P_t) and $A\varphi = \mathcal{A}\varphi$. □

Corollary 5.11. *Let $r > 0$. Then there exists an invariant measure with the support in M_r.*

P r o o f. Let θ be a Borel probability measure with a support in M_r. The semigroup (P_t) is Feller on $B_b(T\mathbb{S}^2)$, the average probability measures $T^{-1}\int_0^T P_s^*\theta\,\mathrm{d}s$ are supported in M_r, hence they are tight and therefore any of its weak cluster points is an invariant probability measure according to the Krylov-Bogolyubov theorem, see e.g. Corollary 3.1.2 in [16]. $\square$

We have proved so far that the tangent bundle $T\mathbb{S}^2$ decomposes to invariant sets

$$T\mathbb{S}^2 = \bigcup_{x\in M_0}\{x\} \cup \bigcup_{r>0} M_r,$$

where on each of these sets there exists an invariant measure.

6. Numerical simulations

We present a numerical scheme to approximate problem (3.6). It is the consequent simulations that lead us to conjecture that (P_t^*) restricted to M_r attracts every initial distribution on M_r to the normalized surface measure on M_r. In particular, this would mean that the normalized surface measure on M_r is the unique invariant measure on M_r, cf. Corollary 5.11.

6.1. Numerical approximation. Let $I_k := \{t_n\}_{n=0}^N$ denote an equi-distant mesh of size $k > 0$ covering $[0, T]$. The following Algorithm A gives a non-dissipative, symmetric discretization of (3.4) with solutions $\{(U^n, V^n);\ n \geqslant 0\}$. We denote $d_t\varphi^{n+1} := (\varphi^{n+1} - \varphi^n)/k$. Throughout this section, $C > 0$ denotes a constant which does not depend on k and T.

Algorithm A. Let (U^0, V^0) be such that $(U^0, V^0) = 0$, $|U^0| = 1$, $|V^0| = r$, and define $U^{-1} := U^0 - kV^0$. For every $n \geqslant 0$, find the $\mathbb{R}^{3+3+1}$-valued random variable $(U^{n+1}, V^{n+1}, \lambda^{n+1})$ such that

$$\begin{aligned} V^{n+1} - V^n &= k\frac{\lambda^{n+1}}{2}(U^{n+1} + U^{n-1}) \\ &\quad + \frac{1}{4}(U^{n+1} + U^{n-1}) \times (V^{n+1} + V^n)\Delta_{n+1}W, \\ d_tU^{n+1} &= V^{n+1}, \end{aligned} \tag{6.1}$$

where $\Delta_{n+1}W := W(t_{n+1}) - W(t_n) \sim \mathscr{N}(0,k)$, and

$$
\lambda^{n+1} = \begin{cases} 0 & \text{for } \frac{1}{2}(U^{n+1}+U^{n-1}) = 0, \\ -\dfrac{(V^n, V^{n+1})}{\left|\frac{1}{2}(U^{n+1}+U^{n-1})\right|^2} & \text{for } \frac{1}{2}(U^{n+1}+U^{n-1}) \neq 0 \text{ and } n \geqslant 1, \\ -\dfrac{(V^0,V^1) - \frac{1}{2}|V^0|^2}{\left|\frac{1}{2}(U^1+U^{-1})\right|^2} & \text{for } \frac{1}{2}(U^1+U^{-1}) \neq 0 \text{ and } n = 0. \end{cases} \tag{6.2}
$$

For our simulations, we use $(U^0,V^0) := (u(0),v(0))$. Next, we show the existence of a sequence of triples $\{(U^{n+1},V^{n+1},\lambda^{n+1});\ 0 \leqslant n \leqslant N-1\}$ which solves (6.1)–(6.2). Definition (6.2) of the discrete Lagrange multiplier λ^{n+1} ensures that $|U^{n+1}| = 1$ for $n \geqslant 0$; here, the definition of λ^1 according to $(6.2)_3$ accounts for the fact that $1 \neq |U^{-1}| \leqslant 1 + rk$ in general. Finally, $|V^{n+1}| = |V^1| = |V^0| + Ck$ for all $n \geqslant 0$ is valid, so that the solutions $\{(U^n,V^n);\ 0 \leqslant n \leqslant N\}$ of Algorithm A inherit the properties of the solutions $\{(u(t),v(t));\ t \in [0,T]\}$ of (3.5) stated in Proposition 4.1.

Proposition 6.1. *Let $k \leqslant k_0(r)$ be sufficiently small. For every $n \geqslant 0$ there exists an $\mathbb{R}^{3+3+1}$-valued random variable $(U^{n+1},V^{n+1},\lambda^{n+1})$ which solves (6.1)–(6.2). Moreover, the iterates satisfy $|U^{n+1}| = |U^0|$, and $|V^{n+1}| = |V^1|$ for $1 \leqslant n \leqslant N-1$, where $\big||V^1| - |V^0|\big| \leqslant Ck$.*

The proof is by induction, and uses Brouwer's fixed point argument to show existence, and a proper "testing" of (6.1) in combination with the definition (6.2) to verify the given properties.

Proof. *Induction assumption.* Fix $n \geqslant 1$; for the sake of better presentation, we consider $n = 0$ at the end of the proof. Let $\{(U^l,V^l);\ 0 \leqslant l \leqslant n\}$ be a solution of (6.1)–(6.2) which satisfies $|U^l| = 1$ for $0 \leqslant l \leqslant n$ and $|V^l| = \widetilde{r} := |V^1| \leqslant 2r$ for $1 \leqslant l \leqslant n$. Further, let $k \leqslant k_0(\widetilde{r})$ be such that $k|V^l| \leqslant 1/4$.

Step 1. Construction of $(U^{n+1},V^{n+1},\lambda^{n+1})$. In a preparatory step, we define $A^{n+1} := (U^{n+1}+U^{n-1})/2$ and rewrite the leading term in (6.1) as

$$
V^{n+1} - V^n = \frac{1}{k}(U^{n+1} - 2U^n + U^{n-1}) = \frac{2}{k}(A^{n+1} - U^n). \tag{6.3}
$$

Hence, (6.1) may be rewritten as

$$
\begin{aligned} \frac{2}{k}(A^{n+1} - U^n) &= k\lambda^{n+1}A^{n+1} + \frac{1}{2}A^{n+1} \times (V^{n+1}+V^n)\Delta_{n+1}W \\ &= k\lambda^{n+1}A^{n+1} + A^{n+1} \times \Big(V^n - \frac{1}{k}U^n\Big)\Delta_{n+1}W. \end{aligned} \tag{6.4}
$$

Obviously, we have found (U^{n+1}, V^{n+1}) once A^{n+1} is constructed, which is a zero of the mapping $\mathcal{F}^{\omega}_{0,n}\colon \mathbb{R}^3 \to \mathbb{R}^3$,

$$\mathcal{F}^{\omega}_{0,n}(\widetilde{A}) := \frac{2}{k}(\widetilde{A} - U^n) - k\widetilde{\lambda}_{0,n}\widetilde{A} - \widetilde{A} \times \Big(V^n - \frac{1}{k}U^n\Big)\Delta_{n+1}W, \tag{6.5}$$

where according to (6.3),

$$\widetilde{\lambda}_{0,n} \equiv \widetilde{\lambda}_{0,n}(\widetilde{A}) := -\frac{|V^n|^2 + 2k^{-1}(V^n, \widetilde{A} - U^n)}{|\widetilde{A}|^2}. \tag{6.6}$$

Since $\mathcal{F}^{\omega}_{0,n}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is not a continuous mapping, we consider a modification $\mathcal{F}^{\omega}_{\varepsilon,n}$ with some $1/8 \leqslant \varepsilon \leqslant 1/4$, where $\widetilde{\lambda}_{0,n}$ in $\mathcal{F}^{\omega}_{0,n}$ is replaced by

$$\widetilde{\lambda}_{\varepsilon,n} \equiv \widetilde{\lambda}_{\varepsilon,n}(\widetilde{A}) := -\frac{|V^n|^2 + 2k^{-1}(V^n, \widetilde{A} - U^n)}{\max\{\varepsilon, |\widetilde{A}|^2\}}. \tag{6.7}$$

(a) *Solvability of* $\mathcal{F}^{\omega}_{\varepsilon,n}(\widetilde{A}) = 0$ *for every* $1/8 \leqslant \varepsilon \leqslant 1/4$. The map $\mathcal{F}^{\omega}_{\varepsilon,n}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is continuous. Moreover, by computing

$$\begin{aligned}\frac{2}{k}(\widetilde{A} - U^n, \widetilde{A}) &\geqslant \frac{2}{k}(|\widetilde{A}| - |U^n|)|\widetilde{A}|,\\ -k(\widetilde{\lambda}_{\varepsilon,n}\widetilde{A}, \widetilde{A}) &\geqslant -k|V^n|^2 - 2|V^n|(|\widetilde{A}| + |U^n|),\end{aligned}$$

we may conclude by the induction assumption that there exists a deterministic number $R_n := R_n(\widetilde{r}) > 0$ such that for $k \leqslant \widetilde{k}_0(\widetilde{r})$ we have

$$(\mathcal{F}^{\omega}_{\varepsilon,n}(\widetilde{A}), \widetilde{A}) \geqslant 0 \quad \forall \widetilde{A} \in \{\mathcal{A} \in \mathbb{R}^3\colon\ |\mathcal{A}| \geqslant R_n\}.$$

By Brouwer's fixed point theorem, there exists $\widetilde{A}^*$ such that $\mathcal{F}^{\omega}_{\varepsilon,n}(\widetilde{A}^*) = 0$ where $1/8 \leqslant \varepsilon \leqslant 1/4$. We now show that $\widetilde{A}^*$ also solves $\mathcal{F}^{\omega}_{0,n}(\widetilde{A}^*) = 0$ provided $k \leqslant k_0(r) \leqslant \widetilde{k}_0(\widetilde{r})$. For this purpose, we use the definitions

$$U^{n+1}_{\varepsilon} := 2\widetilde{A}^* - U^{n-1} \quad \text{and} \quad V^{n+1}_{\varepsilon} := \frac{1}{k}(U^{n+1}_{\varepsilon} - U^n) \tag{6.8}$$

to write (see (6.1))

$$V^{n+1}_{\varepsilon} - V^n = k\frac{\lambda^{n+1}_{\varepsilon}}{2}(U^{n+1}_{\varepsilon} + U^{n-1}) + \frac{1}{4}(U^{n+1}_{\varepsilon} + U^{n-1}) \times (V^{n+1}_{\varepsilon} + V^n)\Delta_{n+1}W, \tag{6.9}$$

where

$$\begin{aligned}\lambda^{n+1}_{\varepsilon} &:= -\frac{(V^n, V^{n+1}_{\varepsilon})}{\max\{\varepsilon, |\frac{1}{2}(U^{n+1}_{\varepsilon} + U^{n-1})|^2\}}\\ &= -\frac{1}{k}\frac{(V^n, [U^{n+1}_{\varepsilon} + U^{n-1}] - [U^n + U^{n-1}])}{\max\{\varepsilon, |\frac{1}{2}(U^{n+1}_{\varepsilon} + U^{n-1})|^2\}}.\end{aligned} \tag{6.10}$$

It now suffices to show that $\frac{1}{2} \leqslant |(U_\varepsilon^{n+1} + U^{n-1})/2|^2 = \max\{\varepsilon, |\frac{1}{2}(U_\varepsilon^{n+1} + U^{n-1})|^2\}$, since in this case $\widetilde{\lambda}_{\varepsilon,n}(\widetilde{A}^*) = \widetilde{\lambda}_{0,n}(\widetilde{A}^*)$.

(b) $\widetilde{A}^*$ *also satisfies* $\mathcal{F}^\omega_{0,n}(\widetilde{A}^*) = 0$ *provided* $k \leqslant k_0(\widetilde{r})$. By (6.9), the inverse triangle inequality, the induction assumption, and for $k \leqslant k_0(\widetilde{r})$,

$$
\begin{aligned}
(6.11)\quad \Big|\frac{1}{2}(U_\varepsilon^{n+1} + U^{n-1})\Big| &= \Big|\frac{k}{2}V_\varepsilon^{n+1} + \frac{1}{2}(U^n + U^{n-1})\Big| = \Big|\frac{k}{2}(V_\varepsilon^{n+1} + V^n) + U^{n-1}\Big| \\
&\geqslant |U^{n-1}| - \Big(\frac{k}{2}|V^n| + \frac{k}{2}|V_\varepsilon^{n+1}|\Big) \\
&\geqslant 1 - \frac{1}{4} - \frac{k}{2}|V_\varepsilon^{n+1}|.
\end{aligned}
$$

It remains to show that $k|V_\varepsilon^{n+1}|/2 \leqslant 1/4$. For this purpose, multiply (6.9) with $V_\varepsilon^{n+1} + V^n$ and use the binomial formula to get

$$
\begin{aligned}
(6.12)\qquad |V_\varepsilon^{n+1}|^2 - |V^n|^2 &= k\frac{\lambda_\varepsilon^{n+1}}{2}(U_\varepsilon^{n+1} + U^{n-1}, V_\varepsilon^{n+1} + V^n) \\
&= k\frac{\lambda_\varepsilon^{n+1}}{2}(kV_\varepsilon^{n+1} + U^n + U^{n-1}, V_\varepsilon^{n+1} + V^n).
\end{aligned}
$$

Note that since $|a|/\max\{\varepsilon, |a|^2\} \leqslant 1/\sqrt{\varepsilon}$, we get by $(6.10)_2$

$$
\frac{k}{2}|\lambda_\varepsilon^{n+1}| \leqslant \Big(\frac{1}{\sqrt{\varepsilon}} + \frac{1}{2\varepsilon}(|U^n| + |U^{n-1}|)\Big)|V^n| \leqslant \frac{1}{\sqrt{\varepsilon}}\Big(1 + \frac{1}{\sqrt{\varepsilon}}\Big)\widetilde{r} := C_\varepsilon\widetilde{r},
$$

so that the following bound follows from (6.12):

$$
\begin{aligned}
(6.13)\qquad |V_\varepsilon^{n+1}|^2 \leqslant \widetilde{r}^2 &+ C_\varepsilon\widetilde{r}\Big(k\Big[1 + \frac{1}{2}\Big]|V_\varepsilon^{n+1}|^2 + \frac{k}{2}|V^n|^2\Big) \\
&+ \frac{1}{2}|V_\varepsilon^{n+1}|^2 + \frac{C_\varepsilon^2}{2}\widetilde{r}^2|U^n + U^{n-1}|^2 \\
&+ \frac{C_\varepsilon}{2}\widetilde{r}(|U^n + U^{n-1}|^2 + |V^n|^2).
\end{aligned}
$$

Consequently, by induction assumption, for $k \leqslant k_0(\widetilde{r})$, and since $1/8 \leqslant \varepsilon \leqslant 1/4$,

$$
(6.14)\qquad \frac{1}{4}|V_\varepsilon^{n+1}|^2 \leqslant C(1 + \widetilde{r})^2.
$$

Therefore, we may choose $k \leqslant k_0(\widetilde{r})$ sufficiently small to validate $k|V_\varepsilon^{n+1}|/2 \leqslant 1/4$. By the arguments given before, this settles the existence of a triple $(U^{n+1}, V^{n+1}, \lambda^{n+1})$ which solves (6.1)–(6.2) for the index $n + 1$.

Step 2. Properties of (U^{n+1}, V^{n+1}). We start with showing $|U^{n+1}| = 1$. Taking the scalar product of $(6.1)_1$ with $(U^{n+1} + U^{n-1})/(2k)$, using $(6.1)_2$, the binomial formula, and elementary calculations lead to

$$
\begin{aligned}
(6.15)\ \lambda^{n+1}\Big|\frac{1}{2}(U^{n+1}+U^{n-1})\Big|^2 &= \frac{1}{2}(d_t V^{n+1}, U^{n+1}+U^{n-1}) \\
&= \frac{1}{2k^2}(|U^{n+1}|^2 + 2(U^{n+1},U^{n-1}) - 2(U^{n+1},U^n) - 2(U^n,U^{n-1}) + |U^{n-1}|^2) \\
&= \frac{1}{2k^2}(|U^{n+1}|^2 - 2k(U^{n+1},V^n) - 2(U^n,U^{n-1}) + |U^{n-1}|^2).
\end{aligned}
$$

By induction assumption, the last term may be replaced by the identity $|U^{n-1}|^2 = |U^n|^2 = 1$. Hence, (6.15) is equal to

$$
\begin{aligned}
&= \frac{1}{2k^2}(|U^{n+1}|^2 - |U^n|^2 - 2k(U^{n+1},V^n) - 2(U^n,U^{n-1}) + 2|U^n|^2) \\
&= \frac{1}{2k^2}(|U^{n+1}|^2 - |U^n|^2 - 2k(U^{n+1},V^n) + 2k(U^n,V^n)) \\
&= \frac{1}{2k^2}(|U^{n+1}|^2 - 1 - 2k^2(V^{n+1},V^n)).
\end{aligned}
$$

The definition of λ^{n+1} in (6.2) then implies $|U^{n+1}| = 1$.

In order to verify $|V^{n+1}| = \widetilde{r}$, we take the scalar product of $(6.1)_1$ with $V^{n+1} + V^n = (U^{n+1} - U^{n-1})/k$ and use the binomial formula:

$$
(6.16) \qquad |V^{n+1}|^2 - |V^n|^2 = \frac{\lambda^{n+1}}{2}(|U^{n+1}|^2 - |U^{n-1}|^2) = 0.
$$

This settles the inductive argument for $n \geqslant 1$.

Modifications for $n = 0$.

Step 1′. In order to construct a triple (U^1, V^1, λ^1), we proceed as in Step 1, with the following exceptions in (6.6), (6.7), (6.10):

$$
\widetilde{\lambda}_{\varepsilon,0} := -\frac{\frac{1}{2}|V^0|^2 + 2k^{-1}(V^0, \widetilde{A} - U^0)}{\max\{\varepsilon, |\widetilde{A}|^2\}}, \quad \lambda^1_\varepsilon := -\frac{(V^0, V^1_\varepsilon) - \frac{1}{2}|V^0|^2}{\max\{\varepsilon, |\frac{1}{2}(U^1_\varepsilon + U^{-1})|^2\}}.
$$

The estimate of $|V^1_\varepsilon|^2 \leqslant C(1+r)^2 \leqslant C(1+\widetilde{r})^2$ in (6.14) follows accordingly since the additional term $-|V^0|^2/2$ in the nominator of λ^1_ε has modulus $r^2/2$. The remaining arguments from Step 1 now apply to establish the existence of the triple (U^1, V^1, λ^1). Note, in particular, that according to (6.11) we have

$$
(6.17) \qquad \Big|\frac{1}{2}(U^1 + U^{-1})\Big| \geqslant \frac{1}{2}.
$$

Step 2′. A slightly modified version of (6.15) leads to the calculation[3]

$$\begin{aligned}
\lambda^1\Big|\frac{1}{2}(U^1+U^{-1})\Big|^2 &= \frac{1}{2}(d_tV^1, U^1+U^{-1})\\
&= \frac{1}{2k^2}(|U^1|^2 - 2k(U^1,V^0) - 2(U^0,U^{-1}) + |U^{-1}|^2)\\
&= \frac{1}{2k^2}(|U^1|^2 - 2k(U^1,V^0) - 2|U^0|^2 + 2k(U^0,V^0) + |U^{-1}|^2)\\
&= \frac{1}{2k^2}((|U^1|^2 - |U^0|^2) - 2k^2(V^1,V^0) - |U^0|^2 + |U^{-1}|^2)\\
&= \frac{1}{2k^2}((|U^1|^2 - 1) - 2k^2(V^1,V^0) - |U^0|^2 + |U^0|^2\\
&\qquad - 2k(V^0,U^0) + k^2|V^0|^2)\\
&= \frac{1}{2k^2}((|U^1|^2 - 1) - 2k^2(V^1,V^0) + k^2|V^0|^2),
\end{aligned}$$

so that $|U^1| = 1$ now follows from $(6.2)_3$.

Next, we proceed as in (6.16) to bound $|V^1|$. By the definition of U^{-1}, and $(U^0, V^0) = 0$, we have

$$\begin{aligned}
(6.18)\qquad |U^{-1}|^2 &= (U^0 - kV^0, U^0 - kV^0)\\
&= |U^0|^2 - 2k(V^0,U^0) + k^2|V^0|^2 = |U^0|^2 + k^2|V^0|^2.
\end{aligned}$$

Taking the scalar product of $(6.1)_1$ with $V^1 + V^0 = k^{-1}(U^1 - U^{-1})$ and employing (6.18), and $|U^1| = |U^0| = 1$, we obtain

$$(6.19)\qquad |V^1|^2 - |V^0|^2 = \frac{\lambda^1}{2}(|U^1|^2 - |U^{-1}|^2) = k^2\frac{\lambda^1}{2}|V^0|^2.$$

In order to bound $|\lambda^1|$ we use (6.17) and the triangle and Young's inequalities:

$$(6.20)\qquad |\lambda^1| \leqslant 4\Big(|V^0||V^1| + \frac{1}{2}|V^0|^2\Big) \leqslant (2|V^1|^2 + 4|V^0|^2).$$

Using (6.20) we get from (6.19) that

$$(6.21)\qquad |V^1|^2 \leqslant |V^0|^2 + k^2|V^0|^2(|V^1|^2 + 2|V^0|^2).$$

Since $|V^0| = r$, it follows from (6.21) that for $k \leqslant k_0(r)$

$$(6.22)\qquad |V^1|^2 \leqslant |V^0|^2 + k^2\frac{3r^4}{1 - r^2k_0^2} \leqslant \left(|V^0| + k\sqrt{\frac{3r^4}{1 - r^2k_0^2}}\right)^2.$$

Hence, it follows from (6.22) that $\big||V^1| - |V^0|\big| \leqslant C(r)k$. The modulus of $|V^1|$ is then exactly preserved for $n > 0$, see (6.16). $\square$

[3] Note that $|U^{-1}|^2 = |kV^0 - U^0|^2$ need not be 1. By the binomial formula, and since $(U^0, V^0) = 0$.

6.2. Numerical experiments. We use Algorithm A to provide simulations for (3.4) in the form

$$\mathrm{d}\dot{u} = -|\dot{u}|^2 u \,\mathrm{d}t + \sqrt{D}(u \times \dot{u}) \circ \mathrm{d}W,$$

where D is a fixed constant that controls the intensity of the noise term. Instead of (6.2), we use an equivalent form

$$\lambda^{n+1} = \frac{-k^{-1}(V^n, U^{n+1} + U^{n-1}) + \frac{1}{2}k^{-2}(1 - |U^{n-1}|^2)}{|\frac{1}{2}(U^{n+1} + U^{n-1})|^2}, \quad n \geqslant 0. \tag{6.23}$$

The above formula is equivalent to the formulation (6.2); since $|U^l|^2 = 1$, $l \geqslant 0$ we obtain for $n > 0$ that $-(V^n, U^{n+1} + U^{n-1})/k + (1 - |U^{n-1}|^2)/(2k^2) = -(V^n, U^{n+1} + U^{n-1})/k = -(V^n, V^{n+1}) + (V^n, U^n + U^{n-1})/k = -(V^n, V^{n+1})$. The equivalence for $n = 0$ follows similarly by recalling that $(U^0, V^0) = 0$. The formulation (6.23) is more convenient for numerical computations, since in this reformulation the round off errors and errors due to inexact solution of the nonlinear system (6.1) do not accumulate over time in the constraint $|U^n| = 1$. The solution of the nonlinear scheme (6.1)–(6.23) is obtained up to machine accuracy by a simple fixed-point algorithm, cf. [5].

The stochastic process $\{(U^n, V^n),\ n \geqslant 0\}$ is computed by the classical Monte-Carlo sampling algorithm; we denote by N_{mc} the number of simulated sample paths of the corresponding stochastic process. In order to obtain an approximation of the marginal probability density function of the stochastic process $\{U^n,\ n \geqslant 0\}$, the unit sphere $\mathbb{S}^2$ is divided into segments $\omega_{ij} \subset \mathbb{S}^2$ associated with points

$$x_{ij} = \Big(\sin\Big(\frac{i\pi}{16}\Big)\cos\Big(\frac{j\pi}{16}\Big),\ \sin\Big(\frac{i\pi}{16}\Big)\sin\Big(\frac{j\pi}{16}\Big),\ \cos\Big(\frac{i\pi}{16}\Big)\Big),$$

$i = 0, \ldots, 16$, $j = 0, \ldots, 31$ such that $\omega_{ij} = \{x \in \mathbb{S}^2;\ x_{ij} = \arg\min\limits_{x_{lm}} |x - x_{lm}|\}$. For the above partition of the sphere, at a fixed time level $t_n = nk$, we construct a piecewise constant empirical probability density function $\hat{f}^n(x)\colon\ \mathbb{S}^2 \to \mathbb{R}$ of $U^n \in \mathbb{S}^2$ as

$$\hat{f}^n(x)|_{\omega_{ij}} = \hat{f}^n(x_{ij}) = \frac{\#\{l;\ U^{n,l} \in \omega_{ij}\}}{|\omega_{ij}|N}$$

for $i = 0, \ldots, 16$, $j = 0, \ldots, 31$, where $\#\Omega$ denotes the cardinality of the set Ω and $\{U^{n,l},\ n \geqslant 0\}$ is the l-th realization (sample path) of the stochastic process $\{U^n,\ n \geqslant 0\}$.

The marginal probability density function $\hat{f}^n$ of $\{U^n,\ n \geqslant 0\}$ was constructed as the average of $N_{\mathrm{mc}} = 20000$ sample paths. For all computations in this section we take the time step size $k = 0.001$ and the initial conditions $U^0 = (0,1,0)$, $V^0 = (1,0,0)$. The marginal probability density function $\hat{f}^0$ associated with the above initial conditions is a Dirac delta function concentrated at U^0.

In Figure 1 we display the computed probability density $\hat{f}^n$ for $D = 1$, $T = 60$ at different time levels. Initially the probability density function is advected in the direction of the initial velocity V^0 and is simultaneously being diffused. For early times, the diffusion seems to act predominantly in the direction perpendicular to the initial velocity. In Figure 2 we display the time averaged marginal probability density function $\bar{f}$ over the last 100 time levels (i.e., we compute $\bar{f}(x) = \frac{1}{100}\sum\limits_{T/k-100}^{T/k} \hat{f}^n(x)$), the function $t_n \to \mathbb{E}[U^n]$ and a zoom at $t_n \to \mathbb{E}[U^n]$ near the center of the sphere.

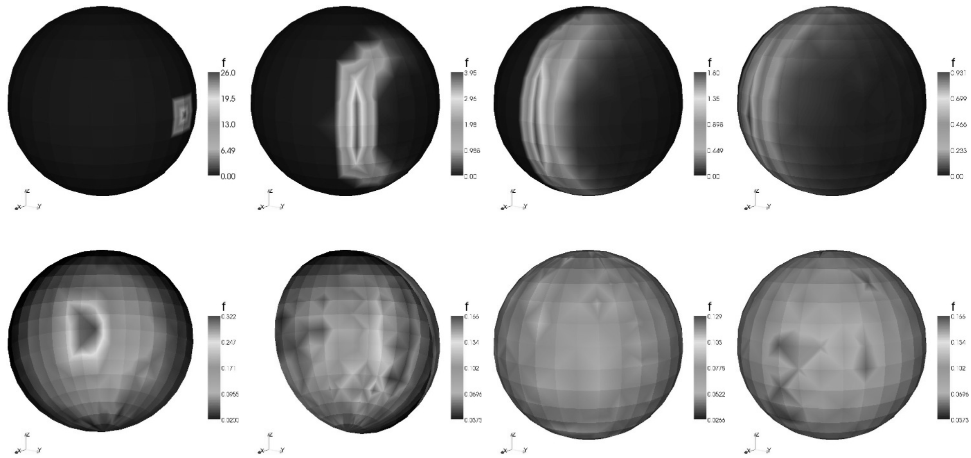

Figure 1. Approximate marginal probability density $\hat{f}^n$ of $\{U^n,\ n \geqslant 0\}$ for $D = 1$ at times $t_n = 0, 1, 1.5, 2.1, 4.3, 5.5, 10, 60$.

The evolution of the probability density $\{\hat{f}^n,\ n \geqslant 0\}$ for $D = 10$, $T = 60$ is shown in Figure 3. Similarly to the previous experiment the probability density function diffuses and becomes uniform for large times. Some advection in the direction of the initial velocity can still be observed, however, the overall process has a predominantly diffuse character. We observe a damping effect which is due to the effects of the random forcing term, see Figure 7. In Figure 4 we display the time averaged probability density function $\bar{f}$, the function $t_n \to \mathbb{E}[U^n]$ and a zoom at $t_n \to \mathbb{E}[U^n]$ for $n \geqslant 0$ near the center.

Figure 6 contains the computed functions of $t_n \to \mathbb{E}[U^n]$ for $D = 0.1$ and $D = 100$. The respective probability densities asymptotically converge towards the uniform distribution for large times.

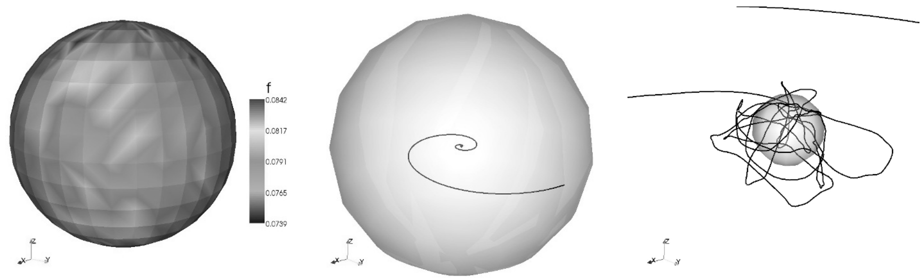

Figure 2. Time averaged marginal probability density $\bar{f}$ of $\{U^n,\ n \geqslant 0\}$ (left), $t_n \to \mathbb{E}[U^n]$ (middle), and a zoom at $\mathbb{E}[U^n]$ with a sphere with radius 0.01 (right), $D = 1$.

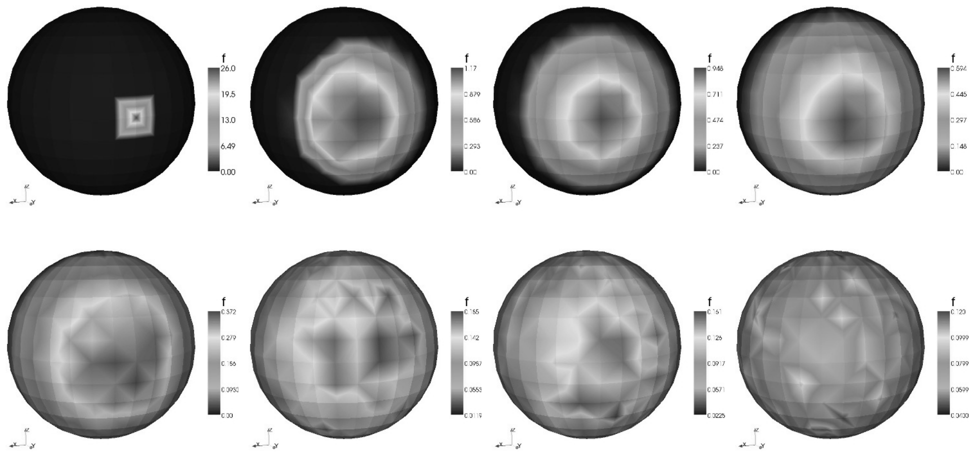

Figure 3. Approximate probability density $\hat{f}^n$ of $\{U^n,\ n \geqslant 0\}$ for $D = 10$ at $t_n = 0, 0.9, 1.2, 2, 3.1, 8, 10, 60$.

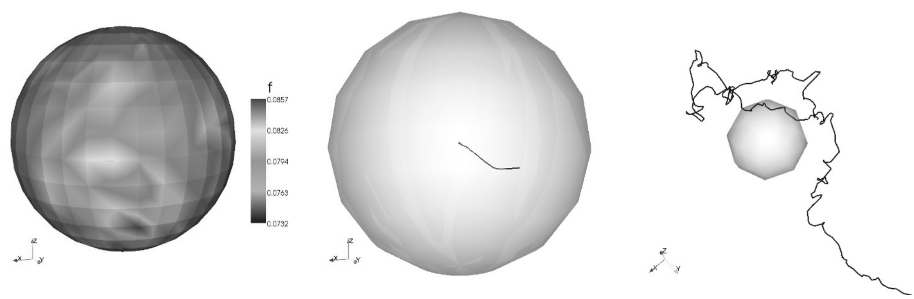

Figure 4. Time averaged marginal probability density $\bar{f}$ of $\{U^n,\ n \geqslant 0\}$ (left), $t_n \to \mathbb{E}[U^n]$ (middle), and a zoom at $t_n \to \mathbb{E}[U^n]$ for $n \geqslant 0$ with a sphere with radius 0.01 (right), $D = 10$.

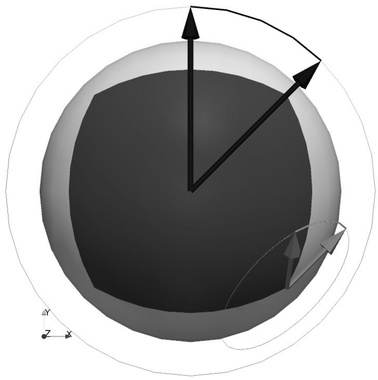

Figure 5. The partition of the submanifold M_1 of $T\mathbb{S}^2$: ω_i in light gray, a segment γ_i^j in black, the dark gray arc indicates the elements of M_i^j starting from a point in the down-right corner of ω_i.

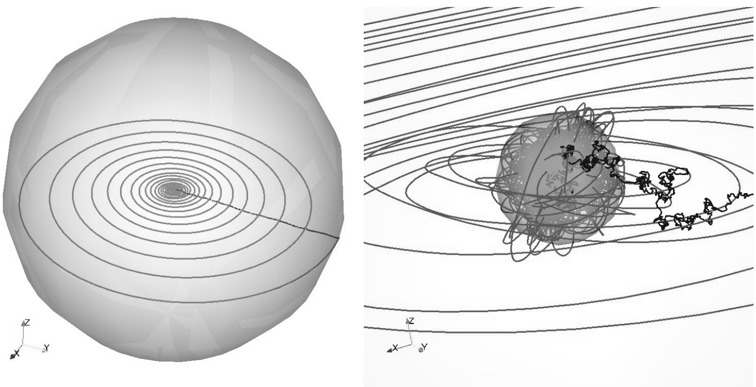

Figure 6. The function $t_n \to \mathbb{E}[U^n]$ (left), and a zoom near the center with a sphere with radius 0.01 (right) for $D = 100$ (black line), $D = 0.1$ (gray line).

In Figure 7 we show the graphs of the time evolution of the approximate error $\mathcal{E}^n_{\max}\colon\ t_n \to \max_{x\in\mathbb{S}^2} \left|\hat{f}^n(x) - f^{\mathbb{S}^2}\right|$ for $D = 0.01, 0.1, 1, 10, 100$ with $f^{\mathbb{S}^2}$ being the uniform distribution on the unit sphere. The quantity $\mathcal{E}^n_{\max}$ serves as a measure of the speed of convergence towards the uniform probability distribution $f^{\mathbb{S}^2}$. Note that the oscillations in the error graphs are due to the approximation of the probability density. The numerical experiments provide evidence that the probability densities for all D converge towards the uniform probability density $f^{\mathbb{S}}$ for $t \to \infty$. The probability density evolutions for decreasing values of D have an increasingly "advective" character, and the evolutions for increasing values have an increasingly "diffusive" character. It is also interesting to note that the convergence in time towards the

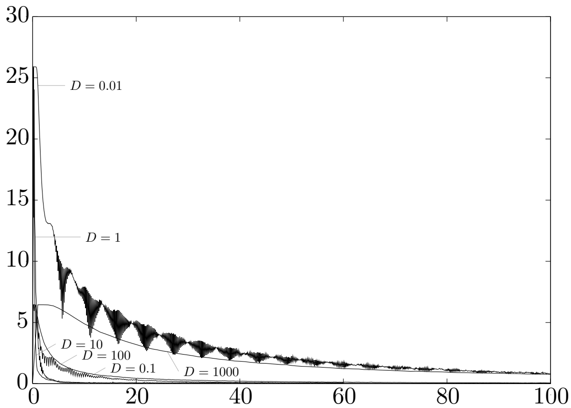

Figure 7. Evolution of $\mathcal{E}^n_{\max}$, $n \geqslant 0$ for different values of the coefficient D.

uniform distribution becomes slower for increasing and decreasing values of D when compared with the fastest converging evolutions for $D = 1$ or $D = 10$.

In the final experiment we study the long time behavior of the pair $\{(U^n, V^n), n \geqslant 0\}$ for $D = 1$, $N_{\mathrm{mc}} = 20000$. To this end, we introduce a partition of the manifold M_1 defined in (5.1). First, we consider a partition of the unit sphere into segments $\{\omega_i,\ i = 1, \ldots, 6\}$ associated with the points $\widetilde{x}_i = (\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)$ in such a way that $x \in \mathbb{S}^2$ belongs to ω_i if and only if $|x - \widetilde{x}_i| = \min_{1 \leqslant j \leqslant 6} |x - \widetilde{x}_j|$. Next, we denote by T_i the tangent planes to points $\widetilde{x}_i$. Fixing an $i \in \{1, \ldots, 6\}$, the orthogonal projections of vectors $\{\widetilde{x}_1, \ldots, \widetilde{x}_6\}$ onto the tangent plane T_i delimit 4 sectors on T_i. We subsequently halve each sector obtaining thus 8 equiangular sectors $\gamma_i^1, \ldots, \gamma_i^8$ in T_i. Now we introduce the following partition of M_1 into 6×8 segments (see Figure 5): a point $(p, \xi) \in T\mathbb{S}^2$ belongs to M_i^j if $p \in \omega_i$ and the orthogonal projection of ξ onto the tangent plane T_i belongs to the sector γ_i^j. The approximate probability density function $\hat{f}^n_{M_1}$ of $\{(U^n, V^n),\ n \geqslant 0\}$ is computed analogously to the marginal probability density function $\hat{f}^n$ of $\{U^n,\ n \geqslant 0\}$. It can be verified by symmetries of this partition that the normalized surface volume of each M_i^j is equal to $1/48$. For $n = 60000$ (i.e., at time $t_n = 60$) we have for $i = 1, \ldots, 6$, $j = 1, \ldots, 8$ that $\#\{l;\ U^{n,l} \in \omega_i\} \in (3380, 3260) \approx N_{\mathrm{mc}}/6 = 3333$ and that $\#\{l;\ (U^{n,l}, V^{n,l}) \in M_i^j\} \in (386, 455) \approx N_{\mathrm{mc}}/6/8 = 417$, see Figure 8 left and Figure 8 right, respectively. This result indicates that the pointwise probability measure for (U^n, V^n), $n \geqslant 0$ converges to the invariant measure $\overline{\nu}$ which is the uniform measure on the set M_1. Figure 9 reveals that the (suitably rescaled) approximate error $\mathcal{E}^n_{M_1,\max} = \left|\hat{f}^n_{M_1} - \overline{\nu}\right|$ for $\{(U^n, V^n),\ n \geqslant 0\}$ has evolution similar to the corresponding error $\mathcal{E}^n_{\max}$ for U^n. Moreover, it seems that the convergence of the error in time is exponential, see Figure 9.

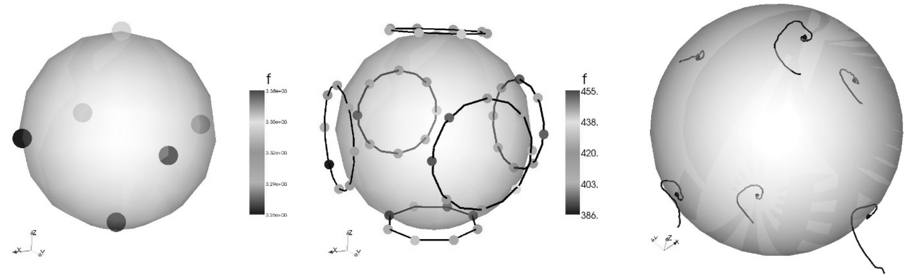

Figure 8. Probability density function $\hat{f}^n_{M_1}$ of $\{(U^n, V^n),\ n \geqslant 0\}$ at time $T = 60$ (left and middle), and the evolution of $t_n \to \int_{\omega_i} \mathbb{E}[V^n]$, $i = 1, \dots, 6$, $n \geqslant 0$ (right).

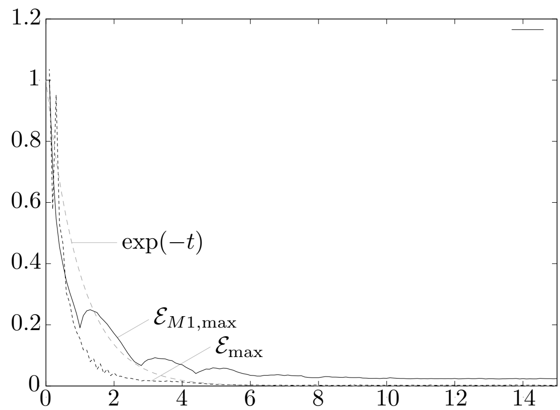

Figure 9. Time evolutions of $\mathcal{E}^n_{M_1,\max}$ (rescaled) and $\mathcal{E}^n_{\max}$, $n \geqslant 0$.

7. Invariant measures on M_r, $r > 0$

It is known that equations on manifolds with non-degenerate diffusions have a unique invariant probability law, that this invariant measure is absolutely continuous with respect to the surface measure, and the density is C^∞-smooth and strictly positive, see e.g. [2] or [26], Proposition 4.5. Unfortunately, the equation (3.6) on M_r has a degenerate diffusion — there is just one vector field g in the diffusion but M_r is a 3-dimensional manifold. In other words, there is not enough noise in the equation in order the above cited results on the nice ergodic behavior could be applied in our case. We must therefore proceed in another way to confirm the conjectures of Section 6.

Convention 7.1. In the present section, we restrict the operators (P_t) and (P^*_t) to the invariant space M_r where $r > 0$ is fixed. More precisely, the operators (P_t) are understood as endomorphisms on $B_b(M_r)$ and (P^*_t) are endomorphisms on the space of probability measures on $\mathscr{B}(M_r)$, cf. Theorem 5.10. Also M_r is understood as a submanifold in $\mathbb{R}^6$.

Notation 7.2. We denote by λ_r the normalized surface (Riemannian) measure on M_r.

7.1. Uniqueness. We are going to prove, using the geometric version of the Hörmander theorem A.3, that λ_r is the unique invariant measure on M_r. But let us first, before we proceed with the study of the qualitative properties of the adjoint Markov semigroup (P_t^*), establish some further geometric properties of the drift and the diffusion vector fields f and g defined in (3.7).

Lemma 7.3. *M_r is a connected 3-dimensional submanifold in $\mathbb{R}^6$ and the vector fields f and g on M_r satisfy*

$$[g,f] = \begin{pmatrix} u \times v \\ 0 \end{pmatrix}, \quad [f,[g,f]] = r^2 g, \quad [g,[g,f]] = -f, \quad \operatorname{div} f = \operatorname{div} g = \operatorname{div}[g,f] = 0$$

where $[\cdot,\cdot]$ is the Jacobi bracket.

Proof. Obviously, any (p,ξ_1) and (p,ξ_2) in M_r can be connected by a rotation curve in the circle $\{(p,\xi)\colon\ \xi \perp p,\ |\xi| = r\}$ and if $|p| = |q| = 1$ and $\gamma\colon\ [a,b] \to \mathbb{S}^2$ is a curve connecting p and q with $|\dot\gamma| = r$ then $\Gamma = (\gamma,\dot\gamma)$ is a curve connecting $(p,\dot\gamma(a))$ and $(q,\dot\gamma(b))$ in M_r. Altogether, any two points in M_r can be connected by an at most two times broken curve.

Observe that f, g and $[g,f]$ are orthogonal tangent vector fields on M_r. If we define $E_1 = f/(r^2+r^4)^{1/2}$, $E_2 = g/r$, $E_3 = [G,F]/r$ then $\{E_1,E_2,E_3\}$ is an othonormal frame on M_r and

$$\operatorname{div} Y = \sum_{j=1}^{3} \langle d_{E_j} Y, E_j\rangle_{\mathbb{R}^6} = 0, \quad Y \in \{f, g, [f,g]\}$$

where $d_X Y(p) = \lim\limits_{t\to 0} t^{-1}[Y(p+tX) - Y(p)]$. $\square$

Notation 7.4. Let $S^1,\dots,S^m$ be vector fields on a manifold M. Denote by $(S^1,\dots,S^m)$ the smallest algebra for the Jacobi bracket $[X,Y] = XY - YX$ that contains $\{S^1,\dots,S^m\}$ and denote

$$\mathscr{L}(S^1,\dots,S^m)(p) = \{S_p\colon\ S \in \mathscr{L}(S^1,\dots,S^m)\} \subseteq T_pM, \quad p \in M.$$

Corollary 7.5. *$\mathscr{L}(f,g)(z) = T_zM_r$ holds for every $z \in M_r$.*

The following result is known[4] but we can give its straight analytic proof in few lines now.

Proposition 7.6. *A probability measure ν on $\mathscr{B}(M_r)$ is invariant if and only if*

$$\int_{M_r} \mathcal{A}h \,\mathrm{d}\nu = 0 \quad \text{for every } h \in C^2(M_r) \tag{7.1}$$

where the operator $\mathcal{A}$ was defined in (5.2).

P r o o f. This is an immediate consequence of the C_0-semigroup property of (P_t) on $C(M_r)$, the invariance of $C^2(M_r)$ under the mappings (P_t), $t \geqslant 0$, the fact that $P_t \circ \mathcal{A} = \mathcal{A} \circ P_t$ on $C^2(M_r)$ for every $t \geqslant 0$, and the density of $C^2(M_r)$ in $B_b(M_r)$ which was all proved in Theorem 5.10. □

Proposition 7.7. *Let $R \in C^2(M_r)$. Then the measure $\mathrm{d}\nu = R\,\mathrm{d}\lambda_r$ satisfies* (7.1) *if and only if R is constant on M_r.*

P r o o f. Using the standard formulae

$$\int_{M_r} Yh \,\mathrm{d}\lambda_r = -\int_{M_r} h \operatorname{div} Y \,\mathrm{d}\lambda_r, \quad \operatorname{div}(hY) = Y(h) + h \operatorname{div} Y$$

that hold for any smooth vector field Y on M_r and any smooth function h on M_r, applying Lemma 7.3 and Proposition 7.6 and using the fact that $C^2(M_r)$ is dense in $L^1(M_r, \lambda_r)$, we get that ν satisfies (7.1) if and only if the identity

$$fR = \frac{1}{2} g(gR) \tag{7.2}$$

holds on M_r. But

$$\int_{M_r} R\Big(fR - \frac{1}{2}g(gR)\Big) \,\mathrm{d}\lambda_r = \frac{1}{2}\int_{M_r} |gR|^2 \,\mathrm{d}\lambda_r$$

as f and g are divergence-free, so we conclude that (7.2) holds if and only if $gR = fR = 0$. If R is constant, this equality surely holds. For the converse implication, by definition of the Lie bracket, $[g, f]R = 0$ holds. Since f_z, g_z and $[g, f]_z$ span $T_z M_r$ for every $z \in M_r$ by Lemma 7.3, we obtain that R is locally constant. But M_r is connected by Lemma 7.3, hence R is constant. □

[4] See e.g. (4.58) on page 292 in [26].

Theorem 7.8. *λ_r is the unique invariant probability measure on M_r.*

P r o o f. Let ν be an invariant measure. Since (7.1) holds and the geometric version of the Hörmander theorem A.3 is applicable due to Corollary 7.5, we conclude that ν has a smooth density R with respect to λ_r. But then $R = 1$ on M_r by Proposition 7.7. $\square$

8. The transition probabilities on M_r, $r > 0$

In this section, we continue the study of the Markov semigroup (P_t) and its adjoint semigroup (P_t^*) restricted to M_r as set forth in Convention 7.1, with $r > 0$ fixed. We are going to show that the transition probabilities $p_{t,x}$ restricted to $\mathscr{B}(M_r)$ for $x \in M_r$ are absolutely continuous with respect to the normalized surface measure λ_r on M_r for every $(t, x) \in (0, \infty) \times M_r$ and that the density $p(t, x, \cdot)$ satisfies $p \in C^\infty((0, \infty) \times M_r \times M_r)$. The density $p(t, x, \cdot)$ should be denoted by $p_r(t, x, \cdot)$ to indicate the dependence on $r > 0$ but we will not use this notation since r is fixed in this section and we will not use the densities elsewhere in this paper.

An expert could be simply advised to apply the abstract results based on the geometric Hörmander theorem in [24], Theorem 3, but we prefer to guide the reader through, to explain the actual structure of the problem better.

For, let us define the adjoint operator

$$\mathcal{A}^* h = -f(h) + \frac{1}{2} g(g(h)), \quad h \in C^2(M_r) \tag{8.1}$$

to the operator $\mathcal{A}$ defined in (5.2). Indeed, by Lemma 7.3,

$$\int_{M_r} (\mathcal{A} h_1) h_2 \, \mathrm{d}\lambda_r = \int_{M_r} h_1 \mathcal{A}^* h_2 \, \mathrm{d}\lambda_r, \quad h_1, h_2 \in C^2(M_r) \tag{8.2}$$

as f and g are divergence-free on M_r.

Theorem 8.1. *The transition probabilities $p_{t,x}$ are absolutely continuous with respect to the normalized surface measure λ_r on M_r for every $(t, x) \in (0, \infty) \times M_r$, and the density $p(t, x, \cdot)$ satisfies $p \in C^\infty((0, \infty) \times M_r \times M_r)$.*

P r o o f. Consider the Riemannian manifold $N = (0, \infty) \times M_r \times M_r$ and define the Radon measure

$$\begin{aligned}\Gamma(A) &= \int_0^\infty \int_{M_r} \int_{M_r} \mathbf{1}_A(t, x, z) \, \mathrm{d}p_{t,x}(z) \, \mathrm{d}\lambda_r(x) \, \mathrm{d}t \\ &= \mathbb{E} \int_0^\infty \int_{M_r} 1_A(t, x, z^x(t)) \, \mathrm{d}t \, \mathrm{d}\lambda_r(x), \quad A \in \mathscr{B}(N).\end{aligned}$$

Every function $h \in C^\infty(N)$ has variables (t,x,z) and we are going to indicate by $\mathcal{A}_z$ that the operator $\mathcal{A}$ is applied to the variable z and by $\mathcal{A}_x^*$ that the operator $\mathcal{A}^*$ is applied to the variable x of the function $h(t,x,z)$.

By the Itô formula,

$$\int_0^\infty \int_{M_r} \Big(\frac{\partial H}{\partial t} + \mathcal{A}H\Big)\,\mathrm{d}p_{t,x}\,\mathrm{d}t = 0, \quad H \in C^\infty_{\mathrm{comp}}((0,\infty)\times M_r) \tag{8.3}$$

holds for every $x \in M_r$, hence

$$\int_N \Big(\frac{\partial h}{\partial t} + \mathcal{A}_z h\Big)\,\mathrm{d}\Gamma = 0, \quad h \in C^\infty_{\mathrm{comp}}(N). \tag{8.4}$$

Let $h_1 \in C^\infty_{\mathrm{comp}}(0,\infty)$, $h_2, h_3 \in C^\infty(M_r)$ and define $H(t,x) = h_1(t)h_2(x)$, $h(t,x,z) = h_1(t)h_2(x)h_3(z)$ and $v(t,x) = P_t h_3(x)$. Then

$$\begin{aligned}\int_N \Big(\frac{\partial h}{\partial t} + \mathcal{A}_x^* h\Big)\,\mathrm{d}\Gamma &= \int_0^\infty \int_{M_r} \Big(\frac{\partial H}{\partial t} + \mathcal{A}^* H\Big) v\,\mathrm{d}\lambda_r\,\mathrm{d}t \\ &= \int_0^\infty \int_{M_r} H\Big(-\frac{\partial v}{\partial t} + \mathcal{A}v\Big)\,\mathrm{d}\lambda_r\,\mathrm{d}t = 0\end{aligned}$$

by (5.3) and the duality (8.2). In fact,

$$\int_N \Big(\frac{\partial h}{\partial t} + \mathcal{A}_x^* h\Big)\,\mathrm{d}\Gamma = 0, \quad h \in C^\infty_{\mathrm{comp}}(N) \tag{8.5}$$

by a density argument as shown in Proposition B.1.

Altogether we have obtained that

$$\int_N \Big(2\frac{\partial h}{\partial t} + \mathcal{A}_x^* h + \mathcal{A}_z h\Big)\,\mathrm{d}\Gamma = 0, \quad h \in C^\infty_{\mathrm{comp}}(N).$$

In order to apply the geometric Hörmander theorem A.3, we define the vector fields

$$Y(t,x,z) = \begin{pmatrix} 2 \\ -f(x) \\ f(z) \end{pmatrix}, \quad X^1(t,x,z) = \begin{pmatrix} 0 \\ g(x) \\ 0 \end{pmatrix}, \quad X^2(t,x,z) = \begin{pmatrix} 0 \\ 0 \\ g(z) \end{pmatrix}$$

where the vector field Y corresponds to the operator $h \mapsto 2\partial h/\partial t - f_x(h) + f_z(h)$, the vector field X^1 to the operator $h \mapsto g_x(h)$ and the vector field X^2 to the operator

$h \mapsto g_z(h)$. Defining also $h = [g, f]$ on M_r, we get by Lemma 7.3 that

$$[Y, X^1] = \begin{pmatrix} 0 \\ h(x) \\ 0 \end{pmatrix}, \quad [Y, X^2] = -\begin{pmatrix} 0 \\ 0 \\ h(z) \end{pmatrix}, \quad [X^1, X^2] = 0, \quad [Y, [Y, X^1]] = -r^2 X^1,$$

$$[Y, [Y, X^2]] = -r^2 X^2, \quad [X^1, [Y, X^1]] = -\begin{pmatrix} 0 \\ f(x) \\ 0 \end{pmatrix}, \quad [X^1, [Y, X^2]] = 0,$$

$$[X^2, [Y, X^1]] = 0, \quad [X^2, [Y, X^2]] = \begin{pmatrix} 0 \\ 0 \\ f(z) \end{pmatrix}, \quad [[Y, X^1], [Y, X^2]] = 0.$$

At this stage we see that

$$\mathscr{L}(Y, X^1, X^2)(t, x, z) = \mathbb{R} \times T_x M_r \times T_z M_r = T_{(t,x,z)} N, \quad (t, x, z) \in N,$$

so the geometric Hörmander theorem A.3 is applicable and Γ has a smooth density $p \in C^\infty(N)$ with respect to $\mathrm{d}t \otimes \lambda_r \otimes \lambda_r$.

Let $\varphi \in C(M_r)$. Then, by the standard measure theoretical properties of integrals,

$$P_t \varphi(x) = \int_{M_r} \varphi(z) p(t, x, z) \, \mathrm{d}\lambda_r(z) \tag{8.6}$$

holds for $\mathrm{d}t \otimes \lambda_r$ a.e. (t, x). But since both sides are continuous in (t, x) (the right hand side by Theorem 5.10), the identity (8.6) holds for every $(t, x) \in (0, \infty) \times M_r$. By standard procedure, we extend (8.6) to hold for every $\varphi \in B_b(M_r)$ and every $(t, x) \in (0, \infty) \times M_r$. □

The following result recasts Corollary 5.9 in terms of the transition densities.

Corollary 8.2. *Let Q be a 3×3-unitary matrix. Denote $\widetilde{Q} = \operatorname{diag}[Q, Q]$. Then*

$$p(t, x, y) = p(t, \widetilde{Q}x, \widetilde{Q}y)$$

holds for every $(t, x, y) \in (0, \infty) \times M_r \times M_r$.

Proof. We just realize that $\widetilde{Q}$ is a measure preserving diffeomorphism on M_r (as a restriction of an isometry on $\mathbb{R}^6$) and then apply Corollary 5.9. □

9. Controlability in M_r, $r > 0$

In this section, we are going to examine the supports of the probability measures $p_{t,x}$ on $\mathscr{B}(M_r)$ for $x \in M_r$. Again, in this section, the Markov semigroup (P_t) and its adjoint semigroup (P_t^*) are restricted to M_r as in Convention 7.1, with $r > 0$ fixed.

Theorem 9.1. *Let $t \geqslant 2\pi/r$. Then* $\operatorname{supp} p_{t,x} = M_r$ *holds for every* $x \in M_r$.

9.1. General support result. Let $x \in M_r$ and denote by $V^{x,a}$ the solutions of the ordinary differential equation

$$X' = f(X) + a(t)g(X), \quad X(0) = x \tag{9.1}$$

on M_r where $a \in L^1_{\mathrm{loc}}([0,\infty))$ and f and g are defined in (3.7).

Remark 9.2. It can be checked analogously as in the proof of Proposition 4.1 that the solutions $V^{x,a}$ take values in M_r and are therefore global.

The next lemma tells us that, to describe the support of the probabilities $p_{t,x}$ for $x \in M_r$, it is sufficient and necessary to study solutions of the ordinary differential equation (9.1).

Lemma 9.3. *Let $t > 0$ and $x \in M_r$. Then*

$$\operatorname{supp} p_{t,x} = \overline{\{V^{x,a}(t)\colon\ a \in L^1(0,t)\}}^{M_r}. \tag{9.2}$$

P r o o f. Let $\widetilde{f}$, $\widetilde{g}$ be smooth compactly supported vector fields on $\mathbb{R}^6$ and denote by μ the law of the solution of the equation

$$\mathrm{d}X = \widetilde{f}(X) + \widetilde{g}(X) \circ \mathrm{d}W, \quad X(0) = x \tag{9.3}$$

on $\mathscr{B}(C([0,t];\mathbb{R}^6))$. Let also $a \in L^1(0,t)$ and denote by v^a the solution of

$$X' = \widetilde{f}(X) + a(t)\widetilde{g}(X), \quad X(0) = x. \tag{9.4}$$

Then, according to the Support theorem of Stroock and Varadhan [42] (see also [1], [7], [8], [22], [31] for generalizations or shorter proofs),

$$\operatorname{supp} \mu = \overline{\{v^a\colon\ a \in L^1(0,t)\}}$$

where the closure and the support are taken in $C([0,t];\mathbb{R}^6)$. Since $v^{a_n} \to v^a$ uniformly on $[0,t]$ if $a_n \to a$ in $L^1(0,t)$ and A is a dense subset in $L^1(0,t)$, it also holds that

$$\operatorname{supp} \mu = \overline{\{v^a\colon\ a \in A\}}.$$

To get back to our problem (3.6), let $\widetilde{f}$, $\widetilde{g}$ be smooth compactly supported vector fields on $\mathbb{R}^6$ such that $\widetilde{f} = f$ and $\widetilde{g} = g$ on the centered ball in $\mathbb{R}^6$ of the radius

$R = |x|$. Then the solution X coincides with z^x being the solution of (3.6) with $z^x(0) = x$. Also, by uniqueness, $V^{x,a} = v^a$. Thus we conclude that

$$\operatorname{supp}(\operatorname{Law} z^x) = \overline{\{V^{x,a}\colon\ a \in L^1(0,t)\}} \tag{9.5}$$

where both the support and the closure are taken in $C([0,t];M_r)$ being a closed subset of $C([0,t];\mathbb{R}^6)$.

Now consider the projection $\pi_t\colon\ C([0,t];M_r) \to M_r\colon\ \xi \mapsto \xi(t)$. Since π_t is continuous,

$$\overline{\pi_t[\operatorname{supp}(\operatorname{Law} z^x)]} = \operatorname{supp}(\operatorname{Law} \pi_t(z^x)) = \operatorname{supp} p_{t,x},$$

and by continuity of π_t and (9.5), we also have

$$\begin{aligned}\overline{\pi_t[\operatorname{supp}(\operatorname{Law} z^x)]} &= \overline{\pi_t[\overline{\{V^{x,a}\colon\ a \in L^1(0,t)\}}]} = \overline{\pi_t[\{V^{x,a}\colon\ a \in L^1(0,t)\}]}\\ &= \overline{\{V^{x,a}(t)\colon\ a \in L^1(0,t)\}}.\end{aligned}$$

□

9.2. The control problem. In view of Lemma 9.3, it remains to prove that the ordinary differential equation (9.1) can be controlled to hit every point in M_r after time $2\pi/r$. It turns out that it is necessary to enter deeper into the geometry of the 2D sphere.

For, consider the equation (9.1) with a constant control $a \in \mathbb{R}$

$$w'' = -|w'|^2 w + a w \times w' \tag{9.6}$$

and with the initial condition $w(0) = p$, $w'(0) = \xi$ for $x = (p,\xi) \in M_r$. It can be guessed (and consequently checked) from the rotational symmetries of (9.6) that the unique solution has the form

$$\begin{gathered} w^{x,a}(t) = \frac{a}{b}E_1^{x,a} + \frac{r}{b}E_2^{x,a}\cos(bt) + \frac{r}{b}E_3^{x,a}\sin(bt),\\ E_1^{x,a} = \frac{a}{b}p + \frac{1}{b}p \times \xi, \quad E_2^{x,a} = \frac{r}{b}p - \frac{a}{rb}p \times \xi, \quad E_3^{x,a} = \frac{1}{r}\xi \end{gathered} \tag{9.7}$$

where $b = \sqrt{r^2 + a^2}$. Since $\{E_1^{x,a}, E_2^{x,a}, E_3^{x,a}\}$ is orthonormal with $\det(E_1^{x,a}, E_2^{x,a}, E_3^{x,a}) = 1$, we deduce that $w^{x,a}$ is a parametrization of a circle on $\mathbb{S}^2$ with the derivative of constant length r.

Lemma 9.4. *A C^2-smooth curve such that $|w|_{\mathbb{R}^3} = 1$ and $|w'|_{\mathbb{R}^3} = r$ satisfies the equation* (9.6) *for some control $a \in \mathbb{R}$ if and only if it parametrizes a non-degenerate circle*[5] *on $\mathbb{S}^2$.*

Hence, solutions of (9.6) can be regarded as oriented circles in $\mathbb{S}^2$.

Notation 9.5. In the sequel, we are going to consider pairs (K, Y) where K is a non-degenerate circle in $\mathbb{S}^2$, i.e. K is understood as a one-dimensional submanifold in $\mathbb{S}^2$, and Y is a vector field on the manifold K with $|Y_p| = r$ for every $p \in K$, i.e. Y determines an orientation of the manifold K. Such pairs are going to be called "oriented circles" in $\mathbb{S}^2$ for simplicity.

Remark 9.6. Any non-degenerate circle K in $\mathbb{S}^2$ can be described in a unique way as $K = (v + P) \cap \mathbb{S}^2$ where P is a two-dimensional subspace in $\mathbb{R}^3$, $v \in \mathbb{R}^3$ is perpendicular to P and $|v| < 1$. Here the vector v is the center of the circle K and P is the plane of the circle. Obviously, if $\bar{v} \in \mathbb{R}^3$ then $K = (\bar{v} + P) \cap \mathbb{S}^2$ if and only if $\bar{v} - v \in P$. Also

$$T_z K = \{p \in P\colon\ p \perp z\} = \{p \in P\colon\ p \perp (z - v)\}, \quad z \in K.$$

If we define $\theta = \sqrt{1 - |v|^2}$, $\{p_1, p_2\}$ is an orthonormal basis in P and

$$Y_z = \frac{r}{\theta}[-\langle z, p_1\rangle p_2 + \langle z, p_2\rangle p_1], \quad z \in K$$

then $\{Y, -Y\}$ are the only two vector fields on K of length r.

Lemma 9.7. *Let $x = (p, \xi) \in M_r$ and define the circle on $\mathbb{S}^2$*

$$K = (p + \operatorname{span}\{E_2^{x,a}, E_3^{x,a}\}) \cap \mathbb{S}^2$$

in the notation of (9.7) *and the vector field on K of length r*

$$Y(z) = -b\langle z, E_3^{x,a}\rangle E_2^{x,a} + b\langle z, E_2^{x,a}\rangle E_3^{x,a}, \quad z \in K^{x,a}$$

where $b = \sqrt{r^2 + a^2}$. Then K is the orbit of $w^{x,a}$ and $Y(w^{x,a}) = (w^{x,a})'$ holds on $\mathbb{R}$.

The following technical result tells us that we can move continuously from one element in M_r to another, along two oriented circles in $\mathbb{S}^2$ with just one "switch" from one circle to the other.

[5] Here "non-degenerate" means that the radius of the circle is strictly positive.

Proposition 9.8. *In the terminology of Definition* 9.5, *let* (K, Y) *be an oriented circle in* $\mathbb{S}^2$ *and let* $(p, \xi) \in M_r$ *satisfy* $p \notin K$. *Then there exist* $z \in K$ *and an oriented circle* (T, B) *in* $\mathbb{S}^2$ *such that* $z, p \in T$, $B_z = Y_z$ *and* $B_p = \xi$.

Proof. Denote by Q_z the vector space generated by $\{p - z, Y_z\}$ for $z \in K$. Since $p - z$ and Y_z are linearly independent, Q_z is two-dimensional. Now $T_z = (p + Q_z) \cap \mathbb{S}^2$ is a non-degenerate circle in $\mathbb{S}^2$ as it contains two distinct points $p, z \in \mathbb{S}^2$. Fixing $z \in K$, we are going to show that there exists a vector field B of length r on T_z such that $B_z = Y_z$. For, if we define

$$R_z = r^2(p - z) - \langle p - z, Y_z\rangle Y_z, \quad z \in K$$

then $R_z \neq 0$ by the linear independence of $\{p - z, Y_z\}$ and we can set $V_z = R_z/|R_z|$. So $\{V_z, r^{-1}Y_z\}$ is an orthonormal basis in Q_z. Let q_z be the orthogonal projection of p onto Q_z and define $p_z = p - q_z$, $\theta_z = \sqrt{1 - |p_z|^2}$. So $T = (p_z + Q_z) \cap \mathbb{S}^2$. According to Remark 9.6,

$$B_z(\tau) = \frac{1}{\theta_z}[\langle \tau, Y_z\rangle V_z - \langle \tau, V_z\rangle Y_z], \quad \tau \in T_z$$

is a vector field of length r on T_z. Since $z - p$ and $z - p_z$ belong to Q_z and $p_z \perp Q_z$, we have $z = p_z + \langle z, V_z\rangle V_z + r^{-2}\langle z, Y_z\rangle Y_z = p_z + \langle z, V_z\rangle V_z$ as $z \perp Y_z$, hence

$$1 = |z|^2 = |p_z|^2 + \langle z, V_z\rangle^2, \quad \theta_z = |\langle z, V_z\rangle|.$$

But

$$|R_z|\langle z, V_z\rangle = \langle z, R_z\rangle = r^2\langle z, p - z\rangle = r^2(\langle z, p\rangle - 1) \leqslant 0$$

so we conclude that $\theta_z = -\langle z, V_z\rangle$. From this we obtain that $B_z(z) = -\theta_z^{-1}\langle z, V_z\rangle \times Y_z = Y_z$. Eventually, $B_z(p) = [\langle p, Y_z\rangle V_z - \langle p, V_z\rangle Y_z]/\theta_z$. It remains to prove that the mapping $L\colon K \to \{\zeta \in T_p\mathbb{S}^2 \colon |\zeta| = r\}$ defined by $L(z) = B_z(p)$ is a surjection. Since K and $\{\zeta \in T_p\mathbb{S}^2 \colon |\zeta| = r\}$ are homeomorphic with $\mathbb{S}^1$ and L is continuous, it is sufficient to prove that L is locally injective by Proposition C.1. Here we can easily see that L_z spans the one-dimensional vector space $Q_z \cap \{p\}^\perp$.

So let us study the injectivity of L. Let $K = (v + U) \cap \mathbb{S}^2$ where U is a two-dimensional subspace in $\mathbb{R}^3$, $v \perp U$ and $|v| < 1$. Let $z_1 \in K$. Then there exists an orthonormal basis u_1, u_2 in U such that $z_1 = v + \xi u_1$ where $1 = |v|^2 + \xi^2$ and $Y(z_1) = ru_2$. If $z_2 \in K$ satisfies $z_1 \neq z_2$ then there exists a unique $\Delta \in (-\pi, \pi] \setminus \{0\}$ such that

$$z_2 = v + \xi u_1 \cos\Delta + \xi u_2 \sin\Delta$$

and, consequently,

$$Y(z_2) = r[-u_1 \sin\Delta + u_2 \cos\Delta].$$

Then $Q_{z_1} \cap Q_{z_2}$ is a one-dimensional space spanned by

$$A = (z_1 - p)\sin\Delta + \frac{\xi}{r}(1-\cos\Delta)Y(z_1) = (z_2 - p)\sin\Delta - \frac{\xi}{r}(1-\cos\Delta)Y(z_2).$$

Obviously, the vector A belongs also to $\{p\}^\perp$ if and only if

$$\psi(\Delta) := \frac{\sin\Delta}{1-\cos\Delta} = \frac{\xi\langle p, u_2\rangle}{1-\langle p, z_1\rangle}. \tag{9.8}$$

Now $\psi\colon (-\pi, \pi] \setminus \{0\} \to \mathbb{R}$ is a bijection and the right hand side of (9.8) is bounded by a constant $C_{p,K}$ irrespective of z_1, z_2, u_1 or u_2, as $p \notin K$. So Δ satisfying the identity (9.8) must verify $|\Delta| \geqslant \varepsilon_{p,K} > 0$ and, consequently, $|z_1 - z_2| \geqslant \varepsilon'_{p,K} > 0$. In particular, L is locally injective and, consequently, L is surjective. The identity (9.8) then also implies that

$$\{z \in K \setminus \{z_1\}\colon\ L(z) \in \{-L(z_1), L(z_1)\}\} = \{z \in K \setminus \{z_1\}\colon\ \dim Q_{z_1} \cap Q_z \cap \{p\}^\perp = 1\}$$

contains exactly one element z_2 which, by surjectivity of L, must satisfy $L(z_1) = -L(z_2)$. In particular, L is injective. □

9.3. Proof of Theorem 9.1. Let $(p_1, \xi_1), (p_3, \xi_3) \in M_r$. We are going to show that, choosing a suitable piecewise constant control a in the equation (9.1), we can reach (p_3, ξ_3) from (p_1, ξ_1) by the solution (9.1) with this control a in any time $T \geqslant 2\pi/r$. We are going to proceed in steps.

First let $a_1 = 0$ and move (p_1, ξ_1) along the solution of (9.6) with the constant control a_1 to some (p_2, ξ_2) in a very short time just to arrange $p_2 \neq p_3$.

Next let a_2 be an extremely large constant control so that the orbit K_2 of the solution $w^{(p_2,\xi_2),a_2}$ does not contain p_3. This can be done by choosing a large control a as the diameter of the orbit is $2r/\sqrt{r^2+a^2}$ by (9.7). This solution defines an oriented circle (K_2, Y_2) in $\mathbb{S}^2$ and $p_3 \notin K_2$. Hence, by Proposition 9.8, there exists an oriented circle (K_3, Y_3) in $\mathbb{S}^2$ such that $z \in K_2 \cap K_3$, $p_3 \in K_3$, $Y_2(z) = Y_3(z)$ and $Y_3(p_3) = \xi_3$. This circle K_3 is associated with a control $a_3 \in \mathbb{R}$.

Let a be the piecewise constant control with steps a_1, a_2 and a_3 at times τ_1, τ_2 and τ_3 so that the solution X to (9.1) with this control satisfies $X(0) = (p_1, \xi_1)$, $X(\tau_1) = (p_2, \xi_2)$, $X(\tau_2) = (z, Y_2(z)) = (z, Y_3(z))$ and $X(\tau_3) = (p_3, \xi_3)$. Now τ_1 was as small as we wanted, $\tau_2 - \tau_1$ too because a_2 was large and the periodicity of the solutions to (9.1) with a constant control a is $2\pi/\sqrt{r^2+a^2}$ by (9.7). Hence $\tau_3 - \tau_2$ is not larger that $2\pi/r$ since we do not let the solution run the full period. Altogether, $\tau_3 < T$.

Let $a_4 \in \mathbb{R}$ be a control such that $T - \tau_3 \in \{2\pi k/\sqrt{r^2+a_4^2}\colon\ k \geqslant 0\}$ and let $a = a_4$ on $(\tau_3, T]$. Then $X(T) = X(\tau_3) = (p_3, \xi_3)$. In other words, we let the solution revolve to wait for the time T, to wind up at the point of the departure (p_3, ξ_3). □

In this section, again, we consider the Markov semigroup (P_t) and its adjoint semigroup (P_t^*) restricted to M_r as in Convention 7.1, with $r > 0$ fixed. We are going to prove the exponential convergence to the invariant measure λ_r in total variation via the Doeblin theorem and a minorization condition due to [32] and [33].

Lemma 10.1. *The transition densities satisfy $p > 0$ on $(2\pi/r, \infty) \times M_r \times M_r$.*

P r o o f. We develop the idea of [33], Section 5.2, and the proof of [32], Lemma 2.3. According to Theorem 8.1, the transition densities $p(t, x, \cdot)$ are smooth in all three variables. Let $t_1 > 2\pi/r$ and $t_2 > 0$ satisfy $t = t_1 + t_2$. Let also $x_0, y_0 \in M_r$ be such that $p(t_2, \cdot, \cdot) \geqslant \varepsilon$ on a neighbourhood $O_{x_0} \times O_{y_0}$ for some $\varepsilon > 0$. Then, from the Chapman-Kolmogorov identity,

$$\begin{aligned} p(t,x,y) &= \int_{M_r} p(t_1,x,z)p(t_2,z,y)\,\mathrm{d}\lambda_r(z) \\ &\geqslant \varepsilon p(t_1,x,O_{x_0}) > 0, \quad x \in M_r,\ y \in O_{y_0} \end{aligned}$$

since the support of $p_{t_1,x}$ is M_r by Theorem 9.1. Now if $p(t, x_1, y_1) = 0$ for some $x_1, y_1 \in M_r$, let $Q \in \mathbb{R}^3 \otimes \mathbb{R}^3$ be one of the two unitary matrices for which $\widetilde{Q} = Q \otimes Q = \operatorname{diag}[Q,Q]$ satisfies $\widetilde{Q} y_1 = y_0$. Then $0 = p(t, x_1, y_1) = p(t, \widetilde{Q} x_1, y_0)$ by Corollary 8.2, which is a contradiction. $\square$

Theorem 10.2. *There exist positive constants c_r, α_r such that*

$$\|P_t^*\nu - \lambda_r\| \leqslant c_r \mathrm{e}^{-\alpha_r t} \|\nu - \lambda_r\|, \quad t \geqslant 0 \tag{10.1}$$

holds for every probability measure ν on $\mathscr{B}(M_r)$, where $\|\cdot\|$ is the norm in total variation on M_r.

P r o o f. Set $\tau = 4\pi/r$. According to Lemma 10.1, there exists $\varepsilon > 0$ such that $p_{\tau,x}(A) \geqslant \varepsilon \lambda_r(A)$ holds for every $x \in M_r$ and every $A \in \mathscr{B}(M_r)$. Hence, by the Doeblin theorem[6], (P_t^*) has a unique invariant probability measure μ on M_r and there exist positive constants c_r and α_r such that

$$\|P_t^*\nu - \mu\| \leqslant c_r \mathrm{e}^{-\alpha_r t} \|\nu - \mu\|, \quad t \geqslant 0$$

holds for every probability measure ν on $\mathscr{B}(M_r)$. But λ_r is the unique invariant probability measure on M_r by Theorem 7.8. $\square$

[6] See e.g. [20], Theorem 4, for a particularly simple proof of the Doeblin theorem.

11. Invariant measures and attractivity on $T\mathbb{S}^2$

In this last section, we are going to study the global dynamics on the full target space $T\mathbb{S}^2$. We will identify the set of invariant probability measures on $\mathscr{B}(T\mathbb{S}^2)$, the set of ergodic probability measures on $\mathscr{B}(T\mathbb{S}^2)$, and it will be shown that the dual Markov semigroup is always attractive.

Notation 11.1. Extend λ_r from $\mathscr{B}(M_r)$ to $\mathscr{B}(T\mathbb{S}^2)$ in the unique way to obtain a probability measure on $\mathscr{B}(T\mathbb{S}^2)$, i.e. $A \mapsto \lambda_r(A \cap M_r)$. Let us denote this extension still by λ_r.

Definition 11.2. If ν is a probability measure on $T\mathbb{S}^2$, we define the probability measures

$$\nu_*(U) = \nu\{(p,\xi) \in T\mathbb{S}^2 \colon\ |\xi| \in U\}, \quad U \in \mathscr{B}([0,\infty)),$$
$$\bar{\nu}(A) = \nu(A \cap M_0) + \int_{(0,\infty)} \lambda_r(A \cap M_r)\,\mathrm{d}\nu_*, \quad A \in \mathscr{B}(T\mathbb{S}^2)$$

in the notation of (5.1).

Remark 11.3. One can check by the definition of λ_r that the mapping $r \mapsto \lambda_r(A \cap M_r)$ is Borel measurable on $(0,\infty)$ for every $A \in \mathscr{B}(T\mathbb{S}^2)$ by the Fubini theorem.

Theorem 11.4. *Let z be a solution of* (3.6) *on $T\mathbb{S}^2$ with an initial distribution ν on $\mathscr{B}(T\mathbb{S}^2)$. Then the laws of $z(t)$ converge in total variation on $T\mathbb{S}^2$ to $\bar{\nu}$ as $t \to \infty$. Moreover, ν is invariant for* (3.6) *if and only if $\nu = \bar{\nu}$ and $\{\delta_x, \lambda_r \colon\ x \in M_0,\ r > 0\}$ is the set of ergodic probability measures for* (3.6).

P r o o f. Let $F\colon\ [0,\infty) \times \mathscr{B}(T\mathbb{S}^2) \to [0,1]$ be a regular version of a conditional probability measure $\nu(\cdot;\ |\xi| = r)$ on $\mathscr{B}(T\mathbb{S}^2)$ for $r \geqslant 0$, i.e., $F(r,\cdot)$ is a probability measure on $\mathscr{B}(T\mathbb{S}^2)$ for every $r \geqslant 0$, $F(\cdot, A)$ is Borel measurable on $[0,\infty)$ for every $A \in \mathscr{B}(T\mathbb{S}^2)$ and

$$\nu(A \cap \{(p,\xi) \colon\ |\xi| \in U\}) = \int_U F(r,A)\,\mathrm{d}\nu_*(r) \tag{11.1}$$

holds for every $A \in \mathscr{B}(T\mathbb{S}^2)$ and $U \in \mathscr{B}[0,\infty)$. The equality (11.1) implies that

$$\int_{T\mathbb{S}^2} h(|\xi|, p, \xi)\,\mathrm{d}\nu(p,\xi) = \int_{[0,\infty)} \Big(\int_{T\mathbb{S}^2} h(r,y)\,\mathrm{d}F_r(y) \Big)\,\mathrm{d}\nu_*(r) \tag{11.2}$$

holds for every bounded measurable $h\colon [0,\infty)\times T\mathbb{S}^2 \to \mathbb{R}$. In particular, setting $h(r,p,\xi) = \mathbf{1}_{[r=|\xi|]}$, we obtain that $\nu_*(O) = 1$ where $O = \{r \geqslant 0\colon F(r, M_r) = 1\}$. So (11.2) implies that

$$
\begin{aligned}
(P_t^*\nu)(A) &= \int_{T\mathbb{S}^2} p(t,x,A)\,\mathrm{d}\nu = \int_O \left(\int_{M_r} p(t,x,A)\,\mathrm{d}F_r(x)\right)\mathrm{d}\nu_*(r)\\
&= \int_O (P_t^* F_r)(A\cap M_r)\,\mathrm{d}\nu_*(r)\\
&= \nu(A\cap M_0) + \int_{O\cap(0,\infty)} (P_t^* F_r)(A\cap M_r)\,\mathrm{d}\nu_*(r)
\end{aligned}
$$

holds for every $t \geqslant 0$ and $A \in \mathscr{B}(T\mathbb{S}^2)$. By a contradiction argument, we get that $P_t^*\nu$ converge in total variation on $T\mathbb{S}^2$ to $\bar\nu$, by Theorem 10.2.

To prove the invariance part of the claim, realize that

$$
\int_{T\mathbb{S}^2} h\,\mathrm{d}\bar\nu = \int_{M_0} h\,\mathrm{d}\nu + \int_{(0,\infty)} \left(\int_{M_r} h\,\mathrm{d}\lambda_r\right)\mathrm{d}\nu_*
$$

holds for every bounded measurable $h\colon T\mathbb{S}^2 \to \mathbb{R}$ by the definition of the measure $\bar\nu$. Hence, setting $h(x) = p(t,x,A)$, we get that

$$
(P_t^*\bar\nu)(A) = \nu(A\cap M_0) + \int_{(0,\infty)} \lambda_r(A\cap M_r)\,\mathrm{d}\nu_* = \bar\nu(A)
$$

holds for every $A \in \mathscr{B}(T\mathbb{S}^2)$ by Theorem 7.8. In particular, $\bar\nu$ is invariant. If ν is invariant then $\nu = \lim_{t\to\infty} P_t^*\nu = \bar\nu$ by the first part of the proof.

Concerning the ergodic measures (according to Definition 5.1), the probability measures $\{\delta_x, \lambda_r\colon\ x \in M_0,\ r > 0\}$ are invariant by the second part of the proof and ergodicity follows from Remark 11.5 as ergodic probability measures are the extremal points of the set of all invariant probability measures (see e.g. Proposition 3.2.7 in [16]). Indeed, the probability measure ν_a is ergodic for (3.6) if and only if a is an extremal point in the convex set of probability measures on $\mathscr{B}(M_0 \,\dot\cup\, (0,\infty))$. This occurs if and only if a is a Dirac measure, i.e. either $a = \delta_x$ for some $x \in M_0$ (hence $\nu_a = \delta_x$) or $a = \delta_r$ for some $r > 0$ (hence $\nu_a = \lambda_r$). $\square$

Remark 11.5. Invariant measures for (3.6) can be uniquely described as measures

$$
\nu_a(A) = a(A\cap M_0) + \int_{(0,\infty)} \lambda_r(A\cap M_r)\,\mathrm{d}a, \quad A \in \mathscr{B}(T\mathbb{S}^2)
$$

where a is a Borel probability measure on the Polish space[7] $X = M_0 \,\dot\cup\, (0,\infty)$, i.e., $G \subseteq X$ is open if and only if $G\cap M_0$ is open in M_0 and $G\cap(0,\infty)$ is open in $(0,\infty)$.

[7] Topological spaces that can be metrized by a complete separable metric are called Polish spaces.

X is Polish as so are M_0 and $(0,\infty)$. The assignment $a \mapsto \nu_a$ is a bijection onto the set of invariant probability measures.

Appendix A: Lie algebra

Let U be an open set on a C^∞-manifold.

- ▷ The set $\mathscr{L}$ of all smooth tangent vector fields on U is a vector space with the Jacobi bracket. Any vector subspace of $\mathscr{L}$ closed under the Jacobi bracket is called a *Lie algebra*.
- ▷ If $\mathcal{X}$ is a set of smooth tangent vector fields on U, then we denote by $\mathscr{L}(\mathcal{X})$ the smallest Lie algebra containing $\mathcal{X}$.
- ▷ If $\mathcal{A} \subseteq \mathscr{L}$ and $p \in U$, then we define $\mathcal{A}(p) = \{A_p\colon\ A \in \mathcal{A}\}$.

Proposition A.1. *Define* $L_0 = \operatorname{span}\{\mathcal{X}\}$ *and* $L_n = \operatorname{span}\{L_{n-1} \cup \{[A,B]\colon\ A, B \in L_{n-1}\}\}$. *Then* $\bigcup L_n = \mathscr{L}(\mathcal{X})$.

Proposition A.2. *Let* $X_1, \dots, X_m, Y \in \mathscr{L}$ *and let* $f_i \in C^\infty(U)$. *Then*

$$\mathscr{L}(X_1, \dots, X_m, Y)(p) = \mathscr{L}\biggl(X_1, \dots, X_m, Y + \sum_{j=1}^m f_j X_j\biggr)(p), \quad p \in U.$$

P r o o f. Let us write $\mathcal{A}^1 = \{X_1, \dots, X_m, Y\}$, $\mathcal{A}^2 = \Bigl\{X_1, \dots, X_m, Y + \sum\limits_{j=1}^m f_j X_j\Bigr\}$,

$$\mathscr{C}^i = \biggl\{\sum_{k=1}^K h_k L_k\colon\ h_k \in C^\infty(U),\ L_k \in \mathscr{L}(\mathcal{A}^i),\ K \in \mathbb{N}\biggr\}.$$

Apparently, $\mathscr{C}^i$ is a Lie algebra for $i \in \{1,2\}$, $\mathcal{A}^i \subseteq \mathscr{C}^j$ whenever $\{i,j\} = \{1,2\}$, hence $\mathscr{L}(\mathcal{A}^i) \subseteq \mathscr{C}^j$ whenever $\{i,j\} = \{1,2\}$. But then

$$\mathscr{L}(\mathcal{A}^i)(p) \subseteq \mathscr{C}^j(p) \subseteq \mathscr{L}(\mathcal{A}^j)(p).$$

□

Theorem A.3 (Hörmander)**.** *Let M be a Riemannian manifold with a countable topological basis, let $X^1, \dots, X^m, Y$ be smooth vector fields on M, let Z be a smooth function on M and let μ be a Radon measure on $\mathscr{B}(M)$ such that*

$$\int_M \Big\{ Zh + Y(h) + \sum_{i=1}^m X^i(X^i(h)) \Big\} \, \mathrm{d}\mu = 0, \quad h \in C^\infty_{\mathrm{comp}}(M) \tag{11.3}$$

and

$$\operatorname{span}\{L_p \colon\ L \in \mathscr{L}(X^1, \dots, X^m, Y)\} = T_pM, \quad p \in M.$$

Then μ has a C^∞-smooth density with respect to the Riemannian measure on M.

Proof. Let $\varphi\colon O \to U$ be a diffeomorphism from an open set $O \subseteq \mathbb{R}^d$ onto an open set $U \subseteq M$, denote by ϕ the inverse of φ, define $\theta(A) = \mu[\varphi[A]]$ for $A \in \mathscr{B}(O)$, decompose $X^i_\varphi = \sum_{j=1}^d x^i_j \partial^j_\varphi$, $Y_\varphi = \sum_{j=1}^d y_j \partial^j_\varphi$ on O and define $z = Z(\varphi)$ and

$$Q = -y + 2\sum_{i=1}^m (\operatorname{div} x^i) x^i, \quad S = z - \operatorname{div} y + \sum_{i=1}^m \operatorname{div}[(\operatorname{div} x^i) x^i].$$

Then (A.1) implies for smooth functions h with compact support in U (which always satisfy the identity $h = \Phi \circ \phi$ on U for some $\Phi \in C^\infty_{\mathrm{comp}}(O)$) that

$$\int_O \Big\{ z\Phi + \sum_{j=1}^d y_j \frac{\partial \Phi}{\partial z_j} + \sum_{i=1}^m \sum_{j=1}^d \sum_{k=1}^d x^i_j \frac{\partial}{\partial z_j}\Big(x^i_k \frac{\partial \Phi}{\partial z_k}\Big) \Big\} \, \mathrm{d}\theta = 0, \quad \Phi \in C^\infty_{\mathrm{comp}}(O), \tag{11.4}$$

i.e.,

$$S\theta + Q(\theta) + \sum_{i=1}^m x^i(x^i(\theta)) = 0$$

holds in the sense of distributions on O. According to Proposition A.2,

$$\begin{aligned} &\operatorname{span}\{L_z \colon\ L \in \mathscr{L}(x^1, \dots, x^m, y)\} \\ &\qquad = \operatorname{span}\{L_z \colon\ L \in \mathscr{L}(x^1, \dots, x^m, Q)\} = \mathbb{R}^d, \quad z \in O. \end{aligned}$$

Hence, by the Hörmander theorem [23], θ is absolutely continuous with respect to the Lebesgue measure and the density ϱ belongs to $C^\infty(O)$. If we define $L = \sqrt{\det g_{ij}}$ on U then

$$\nu(B) = \int_O \mathbf{1}_B(\varphi) \varrho \, \mathrm{d}x = \int_B \frac{\varrho(\phi)}{L} \, \mathrm{d}x, \quad B \in \mathscr{B}(U).$$

By a localization argument, we obtain that μ has a density $R \in C^\infty(M)$ with respect to $\mathrm{d}x$. $\square$

Appendix B: Density of product functions

Proposition B.1. *Let M be a compact submanifold in $\mathbb{R}^m$. Then*

$$\mathcal{P} = \operatorname{span}\{h_1(t)h_2(x)h_3(z)\colon\ h_1 \in C^\infty_{\mathrm{comp}}(0,\infty),\ h_2, h_3 \in C^\infty(M)\}$$

is dense in the space $C^\infty_{\mathrm{comp}}((0,\infty) \times M \times M)$ in the following sense. Let $h \in C^\infty_{\mathrm{comp}}((0,\infty) \times M \times M)$. Then there exist $\chi_n \in \mathcal{P}$ such that

$$\chi_n \to h \quad \text{and} \quad X_m \dots X_1 \chi_n \to X_m \dots X_1 h$$

uniformly on $(0,\infty) \times M \times M$ for every vector fields $X_1, \dots, X_m$ on $(0,\infty) \times M \times M$.

P r o o f. Let $0 < a < b$ be such that the support of h is contained in $(a,b) \times M \times M$ and extend h to a smooth compactly supported function in $\mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m$. This can be done by standard methods of local extensions and a partition of unity as M is assumed to be compact. Denote by h_1 such an extension. The support of h_1 fits in some large cube $Q = (-N,N)^{1+m+m}$ and we can replicate h_1 to each cube $2Nk + Q$ for $k \in \mathbb{Z}^{1+m+m}$ to obtain a smooth $2N$-periodic function h_2 such that $h_1 = h_2$ in Q. Now we can apply Fejér's theorem on Fourier series to find a sequence of functions

$$\xi_n \in \operatorname{span}\{v_1(t)v_2(x)v_3(z)\colon\ v_1 \in C^\infty_{2N\text{-per}}(0,\infty),\ h_2, h_3 \in C^\infty_{2N\text{-per}}(\mathbb{R}^m)\}$$

such that $\xi_n \to h_2$ in $C^\infty(\mathbb{R}^{1+m+m})$. If $\varrho \in C^\infty(\mathbb{R})$ has support in $(0,\infty)$ and $\varrho = 1$ on $[a,b]$ then we can define $\chi_n(t,x,z) = \varrho(t)\xi_n(t,x,z)$. The restrictions of χ_n to $(0,\infty) \times M \times M$ belong to $\mathcal{P}$ and approximate h in the asserted sense. $\square$

Appendix C: Continuous surjections between circles

Proposition C.1. *Let $f\colon \mathbb{S}^1 \to \mathbb{S}^1$ be continuous and locally injective. Then f is a surjection.*

P r o o f. Since $\mathbb{S}^1$ is compact and f is continuous, $f[\mathbb{S}^1]$ is also a compact. But local injectivity of f implies that $f[\mathbb{S}^1]$ is open. Hence f is a surjection as $\mathbb{S}^1$ is connected. $\square$

Acknowledgement. The authors wish to thank Jan Seidler for valuable discussions and for pointing out the works [24] and [25] to us, and the referee whose comments led to an improvement of the paper.

[1] *S. Aida, S. Kusuoka, D. Stroock*: On the support of Wiener functionals. Asymptotic Problems in Probability Theory: Wiener Functionals and Asymptotics. Proc. Conf., Sanda and Kyoto, Japan, 1990 (K. D. Elworthy et al., eds.). Pitman Res. Notes Math. Ser. 284, Longman Scientific & Technical, Harlow, Essex; John Wiley & Sons, New York, 1993, pp. 3–34.

[2] *L. Arnold, W. Kliemann*: On unique ergodicity for degenerate diffusions. Stochastics *21* (1987), 41–61.

[3] *Ľ. Baňas, Z. Brzeźniak, M. Neklyudov, A. Prohl*: A convergent finite-element-based discretization of the stochastic Landau-Lifshitz-Gilbert equation. IMA J. Numer. Anal. *34* (2014), 502–549.

[4] *Ľ. Baňas, Z. Brzeźniak, M. Neklyudov, A. Prohl*: Stochastic Ferromagnetism. Analysis and Numerics. De Gruyter Studies in Mathematics 58, De Gruyter, Berlin, 2014.

[5] *Ľ. Baňas, A. Prohl, R. Schätzle*: Finite element approximations of harmonic map heat flows and wave maps into spheres of nonconstant radii. Numer. Math. *115* (2010), 395–432.

[6] *S. Bartels, C. Lubich, A. Prohl*: Convergent discretization of heat and wave map flows to spheres using approximate discrete Lagrange multipliers. Math. Comput. *78* (2009), 1269–1292.

[7] *G. Ben Arous, M. Grădinaru*: Hölder norms and the support theorem for diffusions. C. R. Acad. Sci., Paris, Sér. I *316* (1993), 283–286. (In French.)

[8] *G. Ben Arous, M. Grădinaru, M. Ledoux*: Hölder norms and the support theorem for diffusions. Ann. Inst. Henri Poincaré, Probab. Stat. *30* (1994), 415–436.

[9] *Z. Brzeźniak, M. Ondreját*: Stochastic geometric wave equations with values in compact Riemannian homogeneous spaces. Ann. Probab. *41* (2013), 1938–1977.

[10] *Z. Brzeźniak, M. Ondreját*: Weak solutions to stochastic wave equations with values in Riemannian manifolds. Commun. Partial Differ. Equations *36* (2011), 1624–1653.

[11] *Z. Brzeźniak, M. Ondreját*: Strong solutions to stochastic wave equations with values in Riemannian manifolds. J. Funct. Anal. *253* (2007), 449–481.

[12] *E. M. Cabaña*: On barrier problems for the vibrating string. Z. Wahrscheinlichkeitstheor. Verw. Geb. *22* (1972), 13–24.

[13] *R. Carmona, D. Nualart*: Random nonlinear wave equations: propagation of singularities. Ann. Probab. *16* (1988), 730–751.

[14] *R. Carmona, D. Nualart*: Random nonlinear wave equations: smoothness of the solutions. Probab. Theory Relat. Fields *79* (1988), 469–508.

[15] *P.-L. Chow*: Stochastic wave equations with polynomial nonlinearity. Ann. Appl. Probab. *12* (2002), 361–381.

[16] *G. Da Prato, J. Zabczyk*: Ergodicity for Infinite-Dimensional Systems. London Mathematical Society Lecture Note Series 229, Cambridge Univ. Press, Cambridge, 1996.

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

[18] *R. C. Dalang, N. E. Frangos*: The stochastic wave equation in two spatial dimensions. Ann. Probab. *26* (1998), 187–212.

[19] *R. C. Dalang, O. Lévêque*: Second-order linear hyperbolic SPDEs driven by isotropic Gaussian noise on a sphere. Ann. Probab. *32* (2004), 1068–1099.

[20] *P. Diaconis, D. Freedman*: On the hit and run process. University of California Berkeley, Statistics Technical Report no. 497, `http://stat-reports.lib.berkeley.edu/accessPages/497.html` (1997).

[21] *J. Ginibre, G. Velo*: The Cauchy problem for the $\mathrm{O}(N)$, $\mathbb{C}\mathrm{P}(N-1)$, and $G_{\mathbb{C}}(N, p)$ models. Ann. Phys. *142* (1982), 393–415.

[22] *I. Gyöngy, T. Pröhle*: On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem. Comput. Math. Appl. *19* (1990), 65–70.

[23] *L. Hörmander*: Hypoelliptic second order differential equations. Acta Math. *119* (1967), 147–171.

[24] *K. Ichihara, H. Kunita*: A classification of the second order degenerate elliptic operators and its probabilistic characterization. Z. Wahrscheinlichkeitstheor. Verw. Geb. *30* (1974), 235–254.

[25] *K. Ichihara, H. Kunita*: Supplements and corrections to the paper: "A classification of the second order degenerate elliptic operators and its probabilistic characterization". Z. Wahrscheinlichkeitstheor. Verw. Geb. *39* (1977), 81–84.

[26] *N. Ikeda, S. Watanabe*: Stochastic Differential Equations and Diffusion Processes. North-Holland Mathematical Library 24, North-Holland; Publishing Co. Tokyo: Kodansha, Amsterdam, 1989.

[27] *A. Karczewska, J. Zabczyk*: Stochastic PDE's with function-valued solutions. Infinite Dimensional Stochastic Analysis. Proceedings of the Colloquium, Amsterdam, Netherlands, 1999 (P. Clément et al., eds.). Verh. Afd. Natuurkd. 1. Reeks. K. Ned. Akad. Wet. 52, Royal Netherlands Academy of Arts and Sciences, Amsterdam, 2000, pp. 197–216.

[28] *A. Karczewska, J. Zabczyk*: A note on stochastic wave equations. Evolution Equations and Their Applications in Physical and Life Sciences. Proc. Conf., Germany, 1999 (G. Lumer et al., eds.). Lecture Notes in Pure and Appl. Math. 215, Marcel Dekker, New York, 2001, pp. 501–511.

[29] *M. Marcus, V. J. Mizel*: Stochastic hyperbolic systems and the wave equation. Stochastics Stochastics Rep. *36* (1991), 225–244.

[30] *B. Maslowski, J. Seidler, I. Vrkoč*: Integral continuity and stability for stochastic hyperbolic equations. Differ. Integral Equ. *6* (1993), 355–382.

[31] *V. Matskyavichyus*: The support of the solution of a stochastic differential equation. Lith. Math. J. *26* (1986), 91–98 (In Russian.); English translation, Lith. Math. J. *26* (1986), 57–62.

[32] *J. C. Mattingly, A. M. Stuart, D. J. Higham*: Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. Stochastic Processes Appl. *101* (2002), 185–232.

[33] *S. Meyn, R. L. Tweedie*: Markov Chains and Stochastic Stability. With a prologue by Peter W. Glynn, Cambridge University Press, Cambridge, 2009.

[34] *A. Millet, P.-L. Morien*: On a nonlinear stochastic wave equation in the plane: existence and uniqueness of the solution. Ann. Appl. Probab. *11* (2001), 922–951.

[35] *A. Millet, M. Sanz-Solé*: A stochastic wave equation in two space dimension: smoothness of the law. Ann. Probab. *27* (1999), 803–844.

[36] *M. Ondreját*: Existence of global martingale solutions to stochastic hyperbolic equations driven by a spatially homogeneous Wiener process. Stoch. Dyn. *6* (2006), 23–52.

[37] *M. Ondreját*: Existence of global mild and strong solutions to stochastic hyperbolic evolution equations driven by a spatially homogeneous Wiener process. J. Evol. Equ. *4* (2004), 169–191.

[38] *S. Peszat*: The Cauchy problem for a nonlinear stochastic wave equation in any dimension. J. Evol. Equ. *2* (2002), 383–394.

[39] *S. Peszat, J. Zabczyk*: Nonlinear stochastic wave and heat equations. Probab. Theory Relat. Fields *116* (2000), 421–443.

[40] *S. Peszat, J. Zabczyk*: Stochastic evolution equations with a spatially homogeneous Wiener process. Stochastic Processes Appl. *72* (1997), 187–204.

[41] *J. Shatah, M. Struwe*: Geometric Wave Equations. Courant Lecture Notes in Mathematics 2, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, 1998.

[42] *D. W. Stroock, S. R. S. Varadhan*: On the support of diffusion processes with applications to the strong maximum principle. Proc. Conf. Berkeley, Calififornia, 1970/1971, Vol. III: Probability Theory (L. M. Le Cam et al., eds.). Univ. California Press, Berkeley, Calififornia, 1972, pp. 333–359.

Authors' addresses: Ľubomír Baňas, Fakultät für Mathematik, Universität Bielefeld, Postfach 100 131, 33501 Bielefeld, Germany, e-mail: banas@math.uni-bielefeld.de; Zdzisław Brzeźniak, Department of Mathematics, The University of York, Heslington, York YO10 5DD, United Kingdom, e-mail: zb500@york.ac.uk; Mikhail Neklyudov, Department of Mathematics, University of Pisa, Largo Bruno Pontecorvo 5, Pisa 56127, Italy, e-mail: misha.neklyudov@gmail.com; Martin Ondreját, The Institute of Information Theory and Automation of the Czech Academy of Sciences, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic, e-mail: ondrejat@utia.cas.cz; Andreas Prohl, Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany, e-mail: prohl@na.uni-tuebingen.de.