STOCHASTIC LIE GROUP INTEGRATORS

SIMON J.A. MALHAM* AND ANKE WIESE*

Abstract. We present Lie group integrators for nonlinear stochastic differential equations with non-commutative vector fields whose solution evolves on a smooth finite dimensional manifold. Given a Lie group action that generates transport along the manifold, we pull back the stochastic flow on the manifold to the Lie group via the action, and subsequently pull back the flow to the corresponding Lie algebra via the exponential map. We construct an approximation to the stochastic flow in the Lie algebra via closed operations and then push back to the Lie group and then to the manifold, thus ensuring our approximation lies in the manifold. We call such schemes stochastic Munthe-Kaas methods after their deterministic counterparts. We also present stochastic Lie group integration schemes based on Castell–Gaines methods. These involve using an underlying ordinary differential integrator to approximate the flow generated by a truncated stochastic exponential Lie series. They become stochastic Lie group integrator schemes if we use Munthe-Kaas methods as the underlying ordinary differential integrator. Further, we show that some Castell–Gaines methods are uniformly more accurate than the corresponding stochastic Taylor schemes. Lastly we demonstrate our methods by simulating the dynamics of a free rigid body such as a satellite and an autonomous underwater vehicle both perturbed by two independent multiplicative stochastic noise processes.

Key words. stochastic Lie group integrators, stochastic differential equations on manifolds

AMS subject classifications. 60H10, 60H35, 93E20

1. Introduction. We are interested in designing Lie group numerical schemes for the strong approximation of nonlinear Stratonovich stochastic differential equations of the form

$$y_t = y_0 + \sum_{i=0}^{d} \int_0^t V_i(y_\tau, \tau) \, \mathrm{d}W_\tau^i. \tag{1.1}$$

Here $W^1, \ldots, W^d$ are d independent scalar Wiener processes and $W_t^0 \equiv t$. We suppose that the solution y evolves on a smooth n-dimensional submanifold $\mathcal{M}$ of $\mathbb{R}^N$ with $n \leq N$ and $V_i \colon \mathcal{M} \times \mathbb{R}_+ \to T\mathcal{M}$, $i = 0, 1, \ldots, d$, are smooth vector fields which in local coordinates are $V_i = \sum_{j=1}^{n} V_i^j \partial_{y_j}$. The flow-map $\varphi_t \colon \mathcal{M} \to \mathcal{M}$ of the integral equation (1.1) is defined as the map taking the initial data y_0 to the solution y_t at time t, i.e. $y_t = \varphi_t \circ y_0$.

Our goal in this paper is to show how the Lie group integration methods developed by Munthe-Kaas and co-authors can be extended to stochastic differential equations on smooth manifolds (see Crouch and Grossman [8] and Munthe-Kaas [40]). Suppose we know that the exact solution of a given system of stochastic differential equations evolves on a smooth manifold $\mathcal{M}$ (see Malliavin [36] or Emery [14]), but we can only find the solution pathwise numerically. How can we ensure that our approximate numerical solution also lies in the manifold?

Suppose we are given a finite dimensional Lie group $\mathcal{G}$ and Lie group action Λ_{y_0} that generates transport across the manifold $\mathcal{M}$ from the starting point $y_0 \in \mathcal{M}$ via elements of $\mathcal{G}$. Then with any given elements ξ in the Lie algebra $\mathfrak{g}$ corresponding to the Lie group $\mathcal{G}$, we can associate the infinitesimal action λ_ξ using the Lie group action Λ_{y_0}. The map $\xi \mapsto \lambda_\xi$ is a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathfrak{X}(\mathcal{M})$, the Lie algebra

of vector fields over the manifold $\mathcal{M}$. Further the Lie subalgebra $\{\lambda_\xi \in \mathfrak{X}(\mathcal{M}) \colon \xi \in \mathfrak{g}\}$ is isomorphic to a finite dimensional Lie algebra with the same structure constants (see Olver [42], p. 56).

Conversely, suppose we know that the Lie algebra generated by the set of governing vector fields V_i, $i = 0, 1, \ldots, d$, on $\mathcal{M}$ is finite dimensional, call this $\mathfrak{X}_F(\mathcal{M})$. Then we know there exists a finite dimensional Lie group $\mathcal{G}$ whose Lie algebra $\mathfrak{g}$ has the same structure constants as $\mathfrak{X}_F(\mathcal{M})$ relative to some basis, and there is a Lie group action Λ_{y_0} such that $V_i = \lambda_{\xi_i}$, $i = 0, 1, \ldots, d$, for some $\xi_i \in \mathfrak{g}$ (see Olver [42], p. 56 or Kunita [30], p. 194). The choice of group and action is not unique.

In this paper we assume that there is a finite dimensional Lie group $\mathcal{G}$ and action Λ_{y_0} such that our set of governing vector fields V_i, $i = 0, 1, \ldots, d$, are each infinitesimal Lie group actions generated by some element in $\mathfrak{g}$ via Λ_{y_0}, i.e. $V_i = \lambda_{\xi_i}$ for some $\xi_i \in \mathfrak{g}$, $i = 0, 1, \ldots, d$. They are said to be fundamental vector fields. This means that we can write down the set of governing vector fields X_{ξ_i} for a system of stochastic differential equations on the Lie group $\mathcal{G}$ that, via the Lie group action Λ_{y_0}, generates the flow governed by the set of vector fields V_i on the manifold. The vector fields V_i on $\mathcal{M}$ are simply the push forward of the vector fields X_{ξ_i} on $\mathcal{G}$ via the Lie group action Λ_{y_0}. Typically the flow on the Lie group also needs to be computed numerically. We thus want the approximation to remain in the Lie group so that the Lie group action takes us back to the manifold.

To achieve this, we pull back the set of governing vector fields X_{ξ_i} on $\mathcal{G}$ to the set of governing vector fields v_{ξ_i} on $\mathfrak{g}$, via the exponential map 'exp' from $\mathfrak{g}$ to $\mathcal{G}$. Thus the stochastic flow generated on $\mathfrak{g}$ by the vector fields v_{ξ_i} generates the stochastic flow on $\mathcal{G}$ generated by the X_{ξ_i}. The set of governing vector fields on $\mathfrak{g}$ are for each $\sigma \in \mathfrak{g}$:

$$v_{\xi_i} \circ \sigma \equiv \sum_{k=0}^{\infty} \frac{B_k}{k!} (\mathrm{ad}_\sigma)^k \circ \xi_i \,, \tag{1.2}$$

where B_k is the kth Bernoulli number and the adjoint operator ad_σ is a closed operator on $\mathfrak{g}$, in fact $\mathrm{ad}_\sigma \circ \zeta = [\sigma, \zeta]$, the Lie bracket on $\mathfrak{g}$. Now the essential point is that $\xi_i \in \mathfrak{g}$ and so the series on the right or any truncation of it is closed in $\mathfrak{g}$. Hence if we construct an approximation to our stochastic differential equation on $\mathfrak{g}$ using the vector fields v_{ξ_i} or an approximation of them achieved by truncating the series representation, then that approximation must reside in the Lie algebra $\mathfrak{g}$. We can then push the approximation in the Lie algebra forward onto the Lie group and then onto the manifold. Provided we compute the exponential map and action appropriately, our approximate solution lies in the manifold (to within machine accuracy). In summary, for a given $\xi \in \mathfrak{g}$ and any $y_0 \in \mathcal{M}$ we have the following commutative diagram:

$$
\begin{array}{ccccc}
\mathfrak{g} & \xrightarrow{\ \exp_* \ } & \mathfrak{X}(\mathcal{G}) & \xrightarrow{\ (\Lambda_{y_0})_* \ } & \mathfrak{X}(\mathcal{M}) \\[4pt]
{\scriptstyle v_\xi}\Big\uparrow & & \Big\uparrow{\scriptstyle X_\xi} & & \Big\uparrow{\scriptstyle \lambda_\xi} \\[4pt]
\mathfrak{g} & \xrightarrow{\ \exp \ } & \mathcal{G} & \xrightarrow{\ \Lambda_{y_0} \ } & \mathcal{M}
\end{array}
$$

We have implicitly separated the governing set of vector fields V_i, $i = 0, 1, \ldots, d$, from the driving path process $w \equiv (W^1, \ldots, W^d)$. Together they generate the unique solution process $y \in \mathcal{M}$ to the stochastic differential equation (1.1). When there is only one driving Wiener process ($d = 1$) the Itô map $w \mapsto y$ is continuous in the topology of uniform convergence. When there are two or more driving processes

($d \geq 2$) the Universal Limit Theorem tells us that the Itô map $w \mapsto y$ is continuous in the p-variation topology, in particular for $2 \leq p < 3$ (see Lyons [32], Lyons and Qian [33] and Malliavin [36]). A Wiener path with $d \geq 2$ has finite p-variation for $p > 2$. This means that from a pathwise perspective, approximations to y constructed using successively refined approximations to w are only guaranteed to converge to the correct solution y, if we include information about the Lévy chordal areas of the driving path process. Note however that the L^2-norm of the 2-variation of a Wiener process is finite. In the Lie group integration procedure prescribed above we must solve a stochastic differential system on the Lie algebra $\mathfrak{g}$ defined by the set of governing vector fields v_{ξ_i} and the driving path process $w \equiv (W^1, \dots, W^d)$. In light of the Universal Limit Theorem and with stepsize adaptivity in mind in future (see Gaines and Lyons [20]), we for instance use in our examples order 1 stochastic numerical methods—that include the Lévy chordal area—to solve for the flow on the Lie algebra $\mathfrak{g}$.

We have thus explained the idea behind Munthe-Kaas methods and how they can be generalized to the stochastic setting. The first half of this paper formalizes this procedure.

In the second half of this paper, we consider autonomous vector fields and construct stochastic Lie group integration schemes using Castell–Gaines methods. This approach proceeds as follows. We truncate the stochastic exponential Lie series expansion corresponding to the flow φ_t of the solution process y to the stochastic differential equation (1.1). We then approximate the driving path process $w \equiv (W^1, \dots, W^d)$ by replacing it by a suitable nearby piecewise smooth path in the appropriate variation topology. An approximation to the solution y_t requires the exponentiation of the approximate truncated exponential Lie series. This can be achieved by solving the system of ordinary differential equations driven by the vector field that is the approximate truncated exponential Lie series. If we use ordinary Munthe-Kaas methods as the underlying ordinary differential integrator the Castell–Gaines method becomes a stochastic Lie group integrator.

Further, based on the Castell–Gaines approach we then present *uniformly accurate exponential Lie series integrators* that are globally more accurate than their stochastic Taylor counterpart schemes (these are investigated in detail in Lord, Malham and Wiese [31] for linear stochastic differential equations). They require the assumption that a sufficiently accurate underlying ordinary differential integrator is used; that integrator could for example be an ordinary Lie group Munthe-Kaas method. In the case of two driving Wiener processes we derive the order $1/2$, and in the case of one driving Wiener process the order 1 uniformly accurate exponential Lie series integrators. As a consequence we confirm the asymptotic efficiency properties for both these schemes proved by Castell and Gaines [8] (see Newton [41] for more details on the concept of asymptotic efficiency). We also present in the case of one driving Wiener process a new order $3/2$ uniformly accurate exponential Lie series integrator (also see Lord, Malham and Wiese [31]).

We present two physical applications that demonstrate the advantage of using stochastic Munthe-Kaas methods. First we consider a free rigid body which for example could model the dynamics of a satellite. We suppose that it is perturbed by two independent multiplicative stochastic noise processes. The governing vector fields are non-commutative and the corresponding exact stochastic flow evolves on the unit sphere. We show that the stochastic Munthe-Kaas method, with an order 1 stochastic Taylor integrator used to progress along the corresponding Lie algebra, preserves the

approximate solution in the unit sphere manifold to within machine error. However when an order 1 stochastic Taylor integrator is used directly, the solution leaves the unit sphere. The contrast between these two methods is more emphatically demonstrated in our second application. Here we consider an autonomous underwater vehicle that is also perturbed by two independent multiplicative stochastic noise processes. The exact stochastic flow evolves on the manifold which is the dual of the Euclidean Lie algebra $\mathfrak{se}(3)$; two independent Casimirs are conserved by the exact flow. Again the stochastic Munthe-Kass method preserves the Casimirs to within machine error. However the order 1 stochastic Taylor integrator is not only unstable for large stepsizes, but the approximation drifts off the manifold and makes a dramatic excursion off to infinity in the embedding space $\mathbb{R}^6$.

Preserving the approximate flow on the manifold of the exact dynamics may be a required property for physical or financial systems driven by smooth or rough paths—for general references see Iserles, Munthe-Kaas, Nørsett and Zanna [25], Hairer, Lubich and Wanner [22], Elworthy [13], Lyons and Qian [33] and Milstein and Tretyakov [38]. Stochastic Lie group integrators in the form of Magnus integrators for linear stochastic differential equations were investigated by Burrage and Burrage [5]. They were also used in the guise of Möbius schemes (see Schiff and Shnider [43]) to solve stochastic Riccati equations by Lord, Malham and Wiese [31] where they outperformed direct stochastic Taylor methods. Further applications where they might be applied include: backward stochastic Riccati equations arising in optimal stochastic linear-quadratic control (Kohlmann and Tang [28]); jump diffusion processes on matrix Lie groups for Bayesian inference (Srivastava, Miller and Grenander [44]); fractional Brownian motions on Lie groups (Baudoin and Coutin [3]) and stochastic dynamics triggered by DNA damage (Chickarmane, Ray, Sauro and Nadim [10]).

Our paper is outlined as follows. In Section 2 we present the basic geometric setup, *sans* stochasticity. In particular we present a generalized right translation vector field on a Lie group that forms the basis of our subsequent transformation from the Lie group to the manifold. Using a Lie group action, this vector field pushes forward to an infinitesimal Lie group action vector field that generates a flow on the smooth manifold. In Section 3 we specialize to the case of a matrix Lie group and using the exponential map, derive the pullback of the generalized right translation vector field on the Lie group to the corresponding vector field on the Lie algebra. To help give some context to our overall scheme, we provide in Section 4 illustrative examples of manifolds and natural choices for associated Lie groups and actions that generate flows on those manifolds. Then in Section 5 we show how a flow on a smooth manifold corresponding to a stochastic differential equation can be generated by a stochastic flow on a Lie algebra via a Lie algebra action. We explicitly present stochastic Munthe-Kaas Lie group integration methods in Section 6. We start the second half of our paper by reviewing the exponential Lie series for stochastic differential equations in Section 7. We show in Section 8 how to construct geometric stochastic Castell–Gaines numerical methods. In particular we also present uniformly accurate exponential Lie series numerical schemes that not only can be used as geometric stochastic integrators, but also are always more accurate than stochastic Taylor numerical schemes of the corresponding order. In Section 9 we present our concrete numerical examples. Finally in Section 10 we conclude and present some further future applications and directions.

2. Lie group actions. Suppose $\mathcal{M}$ is a smooth finite n-dimensional submanifold of $\mathbb{R}^N$ with $n \leq N$. We use $\mathfrak{X}(\mathcal{M})$ to denote the Lie algebra of vector fields on the manifold $\mathcal{M}$, equipped with the Lie–Jacobi bracket $[U, V] \equiv U \cdot \nabla V - V \cdot \nabla U$, for all

$U, V \in \mathfrak{X}(\mathcal{M})$. Let $\mathcal{G}$ denote a finite dimensional Lie group.

DEFINITION 2.1 (Lie group action). *A left Lie group action of a Lie group $\mathcal{G}$ on a manifold $\mathcal{M}$ is a smooth map $\Lambda \colon \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ satisfying for all $y \in \mathcal{M}$ and $R, S \in \mathcal{G}$: (1) $\Lambda(\mathsf{id}, y) = y$; (2) $\Lambda(R, \Lambda(S, y)) = \Lambda(RS, y)$. We denote $\Lambda_y \circ S \equiv \Lambda(S, y)$.*

Hereafter we suppose $y_0 \in \mathcal{M}$ is fixed and focus on the action map $\Lambda_{y_0} \colon \mathcal{G} \to \mathcal{M}$. We assume that the Lie group action Λ is *transitive*, i.e. transport across the manifold from any point $y_0 \in \mathcal{M}$ to any other point $y \in \mathcal{M}$ can always be achieved via a group element $S \in \mathcal{G}$ with $y = \Lambda_{y_0} \circ S$ (Marsden and Ratiu [37], p. 310).

We define the Lie algebra $\mathfrak{g}$ associated with the Lie group $\mathcal{G}$ to be the *vector space of all right invariant vector fields on $\mathcal{G}$*. By standard construction this is isomorphic to the tangent space to $\mathcal{G}$ at the identity $\mathsf{id} \equiv \mathsf{id}_{\mathcal{G}}$ (see Olver [42], p. 48 or Marsden and Ratiu [37], p. 269).

DEFINITION 2.2 (Generalized right translation vector field). *Suppose we are given a smooth map $\xi \colon \mathcal{M} \to \mathfrak{g}$. With each such map ξ we associate a vector field $X_\xi \colon \mathcal{G} \to \mathfrak{X}(\mathcal{G})$ defined as follows*

$$X_\xi \circ S \equiv \partial_\tau \exp\big(\tau\, \xi(\Lambda_{y_0} \circ S)\big)\, S\big|_{\tau=0}$$

for $S \in \mathcal{G}$, where 'exp' is the usual local diffeomorphism $\exp \colon \mathfrak{g} \to \mathcal{G}$ from a neighbourhood of the zero element $\mathsf{o} \in \mathfrak{g}$ to a neighbourhood of $\mathsf{id} \in \mathcal{G}$.

DEFINITION 2.3 (Infinitesimal Lie group action). *We associate with each vector field $X_\xi \colon \mathcal{G} \to \mathfrak{X}(\mathcal{G})$ a vector field $\lambda_\xi \colon \mathcal{M} \to \mathfrak{X}(\mathcal{M})$ as the push forward of X_ξ from $\mathcal{G}$ to $\mathcal{M}$ by Λ_{y_0}, i.e. $\lambda_\xi \equiv \big(\Lambda_{y_0}\big)_* X_\xi$, so that if $S \in \mathcal{G}$ and $y = \Lambda_{y_0} \circ S \in \mathcal{M}$, then*

$$\lambda_\xi \circ y \equiv \partial_\tau \Lambda_{y_0} \circ \gamma(\tau)\big|_{\tau=0}\,,$$

where $\gamma(t) \in \mathcal{G}$, $\gamma(0) = S$ and $\partial_\tau \gamma(\tau) = X_\xi \circ \gamma(\tau)$ (the flow generated on $\mathcal{G}$ by the vector field X_ξ starting at $S \in \mathcal{G}$). Naturally, as a vector field λ_ξ is linear, and also

$$\lambda_\xi \circ y \equiv \mathcal{L}_{X_\xi} \circ \Lambda_{y_0} \circ S\,,$$

the Lie derivative of Λ_{y_0} along X_ξ at $S \in \mathcal{G}$.

Remarks.

1. The map $\Lambda(S) \colon \mathcal{M} \to \mathcal{M}$ defined by $y \mapsto \Lambda(S) \circ y \equiv \Lambda_y \circ S$ represents a flow on $\mathcal{M}$. Hence if $y = \Lambda(S) \circ y_0$, the push forward of λ_ξ by $\Lambda(S)$ is given by $\big(\Lambda(S)\big)_* \lambda_\xi \equiv \lambda_{\mathrm{Ad}_S \xi}$ (Marsden and Ratiu [37], p. 317).

2. We define the *isotropy subgroup* at $y_0 \in \mathcal{M}$ by $\mathcal{G}_{y_0} \equiv \{S \in \mathcal{G} \colon \Lambda_{y_0} \circ S = y_0\}$; it is a closed subgroup of $\mathcal{G}$ (see Helgason [23], p. 121 or Warner [48], p. 123). We define the *global isotropy subgroup* by $\mathcal{G}_{\mathcal{M}} \equiv \cap_{y_0 \in \mathcal{M}} \mathcal{G}_{y_0} \equiv \{S \in \mathcal{G} \colon \Lambda_{y_0} \circ S = y_0,\ \forall y_0 \in \mathcal{M}\}$; it is a normal subgroup of $\mathcal{G}$ (see Olver [42], p. 38).

3. A Lie group action is said to be is *effective/faithful* if the map $S \mapsto \Lambda(S)$ from $\mathcal{G}$ to $\mathrm{Diff}(\mathcal{M})$, the group of diffeomorphisms on $\mathcal{M}$, is one-to-one. This is equivalent to the condition that different group elements have different actions, i.e. $\mathcal{G}_{\mathcal{M}} \equiv \{\mathsf{id}_{\mathcal{G}}\}$. A Lie group action is said to be *free* if $\mathcal{G}_{y_0} = \{\mathsf{id}_{\mathcal{G}}\}$ for all $y_0 \in \mathcal{M}$, i.e. Λ_{y_0} is a diffeomorphism from $\mathcal{G}$ to $\mathcal{M}$. For more details see Marsden and Ratiu [37], p. 310 and Olver [42], p. 38.

4. The map $\gamma \colon \mathcal{G}/\mathcal{G}_{y_0} \to \mathcal{M}$ defined by $\gamma \colon S \cdot \mathcal{G}_{y_0} \mapsto \Lambda_{y_0} \circ S$ is a diffeomorphism, i.e. $\mathcal{M} \cong \mathcal{G}/\mathcal{G}_{y_0}$ for any $y_0 \in \mathcal{M}$ (a manifold $\mathcal{M}$ with a Lie group action $\Lambda \colon \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ defined over it is thus diffeomorphic to a *homogeneous manifold*; see Warner [48], p. 123 or Olver [42], p. 40). Further, the induced action of $\mathcal{G}/\mathcal{G}_{\mathcal{M}}$ on $\mathcal{M}$ is effective. Hence if Λ is not an effective action of $\mathcal{G}$, we can replace it (without loss of generality) by the induced action of $\mathcal{G}/\mathcal{G}_{\mathcal{M}}$ (see Olver [42], p. 38).

5. Our definition for the generalized right translation vector field X_ξ on $\mathcal{G}$ is motivated by the standard right translation vector field used to identify $\mathfrak{g}$, the vector space of right invariant vector fields on $\mathcal{G}$, with $T_{\mathrm{id}}\mathcal{G}$, the tangent space to $\mathcal{G}$ at the identity. When $\xi \in \mathfrak{g}$ is constant, $X_\xi \in \mathfrak{X}(\mathcal{G})$ is right invariant and a Lie bracket on $T_{\mathrm{id}}\mathcal{G}$ can be defined via right extension by the corresponding Lie–Jacobi bracket for the vector fields X_ξ on $\mathfrak{X}(\mathcal{G})$. Unless $\xi \in \mathfrak{g}$ is constant, X_ξ is not in general right invariant. For further details see Varadarajan [47], Olver [42], or Marsden and Ratiu [37].

6. The infinitesimal generator map $\xi \mapsto \lambda_\xi$ from $\mathfrak{g}$ to $\mathfrak{X}(\mathcal{M})$ is a Lie algebra homomorphism. If we identify $\mathfrak{g}$ as the vector space of left invariant vector fields on $\mathcal{G}$ this map becomes an anti-homomorphism. The Lie–Jacobi bracket as defined above gives the right (rather than left) Lie algebra stucture over the group of diffeomorphisms on $\mathcal{M}$. If in addition we take the Lie–Jacobi bracket to be minus that defined above—associated with the left Lie algebra structure—then the infinitesimal generator map becomes a homomorphism again. See for example Marsden and Ratiu [37], p. 324 or Munthe-Kaas [40].

7. The image of $\mathfrak{g}$ under the infinitesimal generator map $\xi \mapsto \lambda_\xi$ forms a finite dimensional Lie algebra of vector fields on $\mathcal{M}$ which is isomorphic to the Lie algebra of the effectively acting quotient group $\mathcal{G}/\mathcal{G}_{\mathcal{M}}$ (see Olver [42], p. 56). Thus the tangent space to $\mathcal{M}$ at any point is $\mathfrak{g}$ and $\mathcal{M}$ inherents a connection from $\mathcal{G}/\mathcal{G}_{\mathcal{M}}$. Connections are necessary to define martingales on manifolds, but not for defining semimartingales (our focus here); see Malliavin [36] and Emery [14].

8. A comprehensive study of the systematic construction of symmetry Lie groups from given vector fields can be found in Olver [42].

9. We assumed above that the vector fields X_ξ and λ_ξ are autonomous. However all results in this and subsequent sections up to Section 7 can be straightforwardly extended to non-autonomous vector fields generated by $\xi\colon \mathcal{M} \times \mathbb{R} \to \mathfrak{g}$ with $(y,t) \mapsto \xi(y,t)$ for all $y \in \mathcal{M}$ and $t \in \mathbb{R}$.

10. For full generality we want to suspend reference to embedding spaces as far as possible. However in subsequent sections to be concise we will more explicitly reclaim this context.

3. Pull back to the Lie algebra. For ease of presentation, we will assume in this section that $\mathcal{G}$ is a matrix Lie group. Recall that the exponential map $\exp\colon \mathfrak{g} \to \mathcal{G}$ is a local diffeomorphism from a neighbourhood of $\mathfrak{o} \in \mathfrak{g}$ to a neighbourhood of $\mathrm{id} \in \mathcal{G}$. Let $v_\xi\colon \mathfrak{g} \to \mathfrak{g}$ be the pull back of the vector field $X_\xi\colon \mathcal{G} \to \mathfrak{X}(\mathcal{G})$ from $\mathcal{G}$ to $\mathfrak{g}$ via the exponential mapping $\exp\colon \mathfrak{g} \to \mathcal{G}$, i.e. $v_\xi \circ \sigma \equiv \exp^* X_\xi \circ \sigma$. If $\sigma \in \mathfrak{g}$ then

$$v_\xi \circ \sigma = \mathrm{dexp}_\sigma^{-1} \circ \xi\big(\Lambda_{y_0} \circ \exp \sigma\big). \tag{3.1}$$

Here $\mathrm{dexp}_\sigma^{-1}\colon \mathfrak{g} \to \mathfrak{g}$ is the inverse of the right-trivialized tangent map of the exponential $\mathrm{dexp}_\sigma\colon \mathfrak{g} \to \mathfrak{g}$ defined as follows. If $\beta(\tau)$ is a curve in $\mathfrak{g}$ such that $\beta(0) = \sigma$ and $\beta'(0) = \eta \in \mathfrak{g}$ then $\mathrm{dexp}\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is the local smooth map (Varadarajan [47], p. 108)

$$\mathrm{dexp}_\sigma \circ \eta \equiv \partial_\tau \exp \beta(\tau)\big|_{\tau=0} \exp(-\sigma)$$
$$= \left(\frac{\exp(\mathrm{ad}_\sigma) - \mathrm{id}}{\mathrm{ad}_\sigma}\right) \circ \eta.$$

Note that as a tangent map $\mathrm{dexp}_\sigma\colon \mathfrak{g} \to \mathfrak{g}$ is linear. The inverse operator $\mathrm{dexp}_\sigma^{-1}$ is the operator series (1.2) generated by considering the reciprocal of dexp_σ.

To show that (3.1) is true, if $\exp\colon \mathfrak{g}\to\mathcal{G}$ with $\sigma\mapsto S=\exp\sigma$, and $\beta(\tau)\in\mathfrak{g}$ with $\beta(0)=\sigma$ and $\partial_\tau\beta(\tau)=v_\xi\circ\beta(\tau)$, then:

$$
\begin{aligned}
\exp_* v_\xi\circ S &= \left.\partial_\tau\exp\beta(\tau)\right|_{\tau=0}\\
&= \left(\mathrm{dexp}_\sigma\circ v_\xi\circ\sigma\right)\exp(\sigma)\\
&\equiv X_\xi\circ S\,.
\end{aligned}
$$

Since 'exp' is a diffeomorphism in a neighbourhood of $\mathsf{o}\in\mathfrak{g}$, this push forward calculation establishes the pull back (3.1) for all $\sigma\in\mathfrak{g}$ in that neighbourhood.

4. Illustrative examples. Suppose the vector field $V\colon \mathcal{M}\times\mathbb{R}\to\mathfrak{X}(\mathcal{M})$ generates a flow solution $y_t\in\mathcal{M}$ starting from $y_0\in\mathcal{M}$. Then assume there exists a:
  1. Lie group $\mathcal{G}$ with corresponding Lie algebra $\mathfrak{g}$;
  2. Lie group action $\Lambda_{y_0}\colon \mathcal{G}\to\mathcal{M}$ for which a starting point $y_0\in\mathcal{M}$ is fixed;
  3. Vector field $\lambda_\xi\colon \mathcal{M}\times\mathbb{R}\to\mathfrak{X}(\mathcal{M})$ such that: $V\equiv\lambda_\xi$, i.e. V is a fundamental vector field corresponding to the action Λ_{y_0}.

Let us suppose $\mathcal{G}$ is a matrix Lie group (or can be embedded into a matrix Lie group, for example the Euclidean group $SE(3)$ is naturally embedded into the special linear group $SL(4;\mathbb{R})$). We have for all $S\in\mathcal{G}$ and $t\in\mathbb{R}$,

$$
X_\xi(S,t)\equiv\xi\big(\Lambda_{y_0}(S),t\big)\,S\,. \tag{4.1}
$$

If $V=\lambda_\xi$ for some $\xi\colon \mathcal{M}\to\mathfrak{g}$, some Lie group $\mathcal{G}$ and corresponding action Λ_{y_0}, then the flow generated by X_ξ on $\mathcal{G}$ drives the flow generated by V on $\mathcal{M}$. In each of the examples below, given the manifold $\mathcal{M}$, we present a natural Lie group and action associated with the manifold structure, and identify vector fields which generate flows on the manifold via the Lie group.

Stiefel manifold $\mathbb{V}_{n,k}$. Suppose $\mathcal{M}=\mathbb{V}_{n,k}\equiv\{y\in\mathbb{R}^{n\times k}\colon y^{\mathrm{T}}y=I\}$. Take $\mathcal{G}=SO(n)$, the special orthogonal group, and $\Lambda_{y_0}(S)\equiv Sy_0$, the action of left multiplication. The corresponding Lie algebra $\mathfrak{g}=\mathfrak{so}(n)$. Then by direct calculation $\lambda_\xi(y)=\xi(y,t)\,y$. Hence if the given vector field $V(y,t)=\xi(y,t)\,y$, then the push forward of the flow generated by $X_\xi(S,t)$ on $\mathcal{G}$ in (4.1) is the flow generated by V on $\mathcal{M}$. Note that the unit sphere $\mathbb{S}^2\cong\mathbb{V}_{3,1}$, i.e. $\mathbb{S}^2$ is just a particular Stiefel manifold. In Section 9 as an application, we consider rigid body dynamics evolving on $\mathbb{S}^2$.

Isospectral manifold $\mathbb{S}_n$. Suppose $\mathcal{M}=\mathbb{S}_n=\{y\in\mathbb{R}^{n\times n}\colon y^{\mathrm{T}}=y\}$, the set of $n\times n$ real symmetric matrices. Take $\mathcal{G}=O(n)$, the orthogonal group and $\Lambda_{y_0}(S)\equiv Sy_0S^{\mathrm{T}}$, which is an isospectral action (Munthe-Kaas [40]). The corresponding Lie algebra is $\mathfrak{g}=\mathfrak{so}(n)$. Again, by direct calculation $\lambda_\xi(y)=\xi(y,t)\,y-y\,\xi(y,t)$. Hence if the given vector field $V(y,t)=\xi(y,t)\,y-y\,\xi(y,t)$, then the push forward of the flow generated by $X_\xi(S,t)$ on $\mathcal{G}$ in (4.1) is the flow generated by V on $\mathcal{M}$.

Dual of the Euclidean algebra $\mathfrak{se}(3)^*$. Suppose $\mathcal{M}=\mathfrak{se}(3)^*\cong\mathbb{R}^3$, the dual of the Euclidean algebra $\mathfrak{se}(3)$ of the Euclidean group $SE(3)=\big\{(s,\rho)\in SE(3)\colon s\in SO(3),\ \rho\in\mathbb{R}^3\big\}$. Take $\mathcal{G}=SE(3)$ so $\mathfrak{g}=\mathfrak{se}(3)$ and $\Lambda\equiv\mathrm{Ad}^*\colon \mathcal{G}\times\mathfrak{g}^*\to\mathfrak{g}^*$, the coadjoint action of $\mathcal{G}$ on $\mathfrak{g}^*$. Then by direct calculation $\lambda_\xi(y)=-\mathrm{ad}^*_\xi(y)$. Since $\lambda_\xi(y)$ in linear in ξ and $-\lambda_\xi(y)\equiv\lambda_{-\xi}(y)$, it follows that if $V(y)=\mathrm{ad}^*_\xi(y)$, then the push forward of the flow generated by $X_{-\xi}(S,t)=-\xi\big(\Lambda_{y_0}(S),t\big)\,S$ on $\mathcal{G}$ is the flow generated by V on $\mathcal{M}$. For more details see Section 9 where we investigate the dynamics of an autonomous underwater vehicle evolving on $\mathfrak{se}(3)^*$.

Grassmannian manifold $\mathrm{Gr}(k,n)$**.** The Grassmannian manifold $\mathcal{M} = \mathrm{Gr}(k,n)$ is the space of k-dimensional subspaces of $\mathbb{R}^n$. Take $\mathcal{G} = \mathrm{GL}(n)$, the general linear matrix group, where if $S \in \mathrm{GL}(n)$, we identify

$$S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

where the block matrices α, β, γ and δ are sizes $k \times k$, $k \times (n-k)$, $(n-k) \times k$ and $(n-k) \times (n-k)$, respectively (see Schiff and Shnider [43]; Munthe-Kaas [40]). We choose the action of $\mathrm{GL}(n)$ on $\mathrm{Gr}(k,n)$ to be the generalized Möbius transformation $\Lambda_{y_0}(S) = (\alpha y_0 + \beta)(\gamma y_0 + \delta)^{-1}$. Hence if

$$\xi(t) = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix},$$

then direct calculation reveals that $\lambda_\xi(y) = a(t)y + b(t) - yc(t)y - yd(t)$. Hence if the given vector field $V(y) = a(t)y + b(t) - yc(t)y - yd(t)$, then the push forward of the flow generated by $X_\xi(S,t) = \xi(t)S$ on $\mathcal{G}$ is the flow generated by V on $\mathrm{Gr}(k,n)$.

5. Stochastic Lie group integration. We show that if a Lie group action $\Lambda \colon \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ exists, then for $y_0 \in \mathcal{M}$ fixed, the Lie algebra action $\Lambda_{y_0} \circ \exp \colon \mathfrak{g} \to \mathcal{M}$ carries a flow on $\mathfrak{g}$ to a flow on $\mathcal{M}$.

THEOREM 5.1. *Suppose there exists a Lie group action* $\Lambda \colon \mathcal{G} \times \mathcal{M} \to \mathcal{M}$*. Then if there exists a process* $\sigma \in \mathfrak{g}$ *and a stopping time* T_* *such that on* $[0, T_*)$*,* σ *satisfies the Stratonovich stochastic differential equation*

$$\sigma_t = \sum_{i=0}^{d} \int_0^t v_{\xi_i} \circ \sigma_\tau \, \mathrm{d}W_\tau^i, \tag{5.1}$$

then the process $y = \Lambda_{y_0} \circ \exp \sigma \in \mathcal{M}$ *satisfies the Stratonovich stochastic differential equation on* $[0, T_*)$*:*

$$y_t = y_0 + \sum_{i=0}^{d} \int_0^t \lambda_{\xi_i} \circ y_\tau \, \mathrm{d}W_\tau^i. \tag{5.2}$$

Proof. Using Itô's lemma, if $\sigma_t \in \mathfrak{g}$ satisfies (5.1) then $\Lambda_{y_0} \circ \exp \sigma_t$ satisfies

$$\Lambda_{y_0} \circ \exp \sigma_t = \Lambda_{y_0} \circ \exp \mathbf{o} + \sum_{i=0}^{d} \int_0^t \mathcal{L}_{v_{\xi_i}} \circ \Lambda_{y_0} \circ \exp \sigma_\tau \, \mathrm{d}W_\tau^i.$$

Now recall that for each $i = 0, 1, \ldots, d$, X_{ξ_i} is the push forward of v_{ξ_i} from $\mathfrak{g}$ to $\mathcal{G}$ via the exponential map, and that λ_{ξ_i} is the push forward of X_{ξ_i} from $\mathcal{G}$ to $\mathcal{M}$ via Λ_{y_0} and so the Lie derivative

$$\mathcal{L}_{v_{\xi_i}} \circ \Lambda_{y_0} \circ \exp \sigma_t \equiv \lambda_{\xi_i} \circ y_t.$$

Then since $y_t = \Lambda_{y_0} \circ \exp \sigma_t$, we conclude that $y \in \mathcal{M}$ is a process satisfying the stochastic differential equation (5.2). □

COROLLARY 5.2. *Suppose that for each* $i = 0, 1, \ldots, d$ *there exists* $\xi_i \colon \mathcal{M} \to \mathfrak{g}$ *such that the vector field* $V_i \colon \mathcal{M} \to \mathfrak{X}(\mathcal{M})$ *and* $\lambda_{\xi_i} \colon \mathcal{M} \to \mathfrak{X}(\mathcal{M})$ *can be identified, i.e.*

$$V_i \equiv \lambda_{\xi_i}. \tag{5.3}$$

Then the push forward by '$\Lambda_{y_0} \circ \exp$' of the flow on the Lie algebra manifold $\mathfrak{g}$ generated by the stochastic differential equation (5.1) is the flow on the smooth manifold $\mathcal{M}$ generated by the stochastic differential equation (5.2), whose solution can be expressed in the form $y_t = \Lambda_{y_0} \circ \exp \sigma_t$.

Remark. If the action is free then '$\Lambda_{y_0} \circ \exp$' is a diffeomorphism from a neighbourhood of $\mathsf{o} \in \mathfrak{g}$ to a neighbourhood of $y_0 \in \mathcal{M}$.

6. Stochastic Munthe-Kaas methods. Assuming that the vector fields in our original stochastic differential equation (1.1) are fundamental and satisfy (5.3), then stochastic Munthe-Kaas methods are constructed as follows:

1. Subdivide the global interval of integration $[0, T]$ into subintervals $[t_n, t_{n+1}]$.
2. Starting with $t_0 = 0$, repeat the next two steps over successive intervals $[t_n, t_{n+1}]$ until $t_{n+1} = T$.
3. Compute an approximate solution $\hat{\sigma}_{t_n, t_{n+1}}$ to (5.1) across $[t_n, t_{n+1}]$ using a stochastic Taylor, stochastic Runge–Kutta or Castell–Gaines method.
4. Compute the approximate solution $y_{t_{n+1}} \approx \Lambda_{y_{t_n}} \circ \exp \hat{\sigma}_{t_n, t_{n+1}}$.

Note that by construction $\hat{\sigma}_{t_n, t_{n+1}} \in \mathfrak{g}$ because the stochastic differential equation (5.1) (or any stochastic Taylor or other sensible approximation) evolves the solution locally on the Lie algebra $\mathfrak{g}$ via the vector fields $v_{\xi_i} \colon \mathfrak{g} \to \mathfrak{g}$. Suitable methods for approximating the exponential map to ensure it maps $\mathfrak{g}$ to $\mathcal{G}$ appropriately can be found in Iserles and Zanna [26]. Then by construction $y_{t_{n+1}} \in \mathcal{M}$.

For example, with two Wiener processes and autonomous vector fields $v_{\xi_i} \circ \sigma$, an order 1 stochastic Taylor Munthe-Kaas method is based on

$$\hat{\sigma}_{t_n, t_{n+1}} = \left(J_0 v_{\xi_0} + J_1 v_{\xi_1} + J_2 v_{\xi_2} + \tfrac{1}{2} J_1^2 v_{\xi_1}^2 + J_{12} v_{\xi_1} v_{\xi_2} + J_{21} v_{\xi_2} v_{\xi_1} + \tfrac{1}{2} J_2^2 v_{\xi_2}^2 \right) \circ \mathsf{o}, \quad (6.1)$$

evaluated at the zero element $\mathsf{o} \in \mathfrak{g}$. Typically '$\mathrm{dexp}_\sigma^{-1}$' is truncated to only include the necessary low order terms to maintain the order of the numerical scheme.

Remark. It is natural to invoke Ado's Theorem (see for example Olver [42] p. 54): any finite dimensional Lie algebra is isomorphic to a Lie subalgebra of $\mathfrak{gl}(n)$ (the general linear algebra) for some $n \in \mathbb{N}$. However as Munthe-Kaas [40] points out, directly using a matrix representation for the given Lie group might not lead to the optimal computational implementation (other data structures might do so).

7. Exponential Lie series. The stochastic Taylor series is known in different contexts as the *Neumann series*, *Peano–Baker series* or *Feynman–Dyson path ordered exponential*. If the vector fields in the stochastic differential equation (1.1) are autonomous (which we assume henceforth), i.e. for all $i = 0, 1, \ldots, d$, $V_i = V_i(y)$ only, then the stochastic Taylor series for the flow is

$$\varphi_t = \sum_{m=0}^{\infty} \sum_{\alpha \in \mathbb{P}_m} J_{\alpha_1 \cdots \alpha_m}(t) \, V_{\alpha_1} \cdots V_{\alpha_m}.$$

Here $\mathbb{P}_m$ is the set of all combinations of multi-indices $\alpha = (\alpha_1, \ldots, \alpha_m)$ of length m with $\alpha_i \in \{0, 1, \ldots, d\}$ and

$$J_{\alpha_1 \cdots \alpha_m}(t) \equiv \int_0^t \cdots \int_0^{\tau_{m-1}} \mathrm{d}W_{\tau_m}^{\alpha_1} \cdots \mathrm{d}W_{\tau_1}^{\alpha_m}$$

are multiple Stratonovich integrals.

The logarithm of φ_t is the *exponential Lie series*, *Magnus expansion* (Magnus [34]) or *Chen–Strichartz formula* (Chen [9], Strichartz [45]). In other words we can express the flow map in the form $\varphi_t = \exp \psi_t$, where

$$\psi_t = \sum_{i=0}^{d} J_i(t) V_i + \sum_{j>i=0}^{d} \tfrac{1}{2}(J_{ij} - J_{ji})(t)[V_i, V_j] + \cdots$$

is the exponential Lie series for our system, and $[\cdot\,,\cdot]$ is the Lie–Jacobi bracket on $\mathfrak{X}(\mathcal{M})$. See Yamato [49], Kunita [29], Ben Arous [1] and Castell [7] for the derivation and convergence of the exponential Lie series expansion in the stochastic context; Strichartz [45] for the full explicit expansion; Sussmann [46] for a related product expansion and Lyons [32] for extensions to rough paths.

Let us denote the truncated exponential Lie series by

$$\hat{\psi}_t = \sum_{\alpha \in \mathbb{Q}_m} J_\alpha \, c_\alpha \,, \tag{7.1}$$

where $\mathbb{Q}_m$ denotes the finite set of multi-indices α for which $\|J_\alpha\|_{L^2}$ is of order up to and including t^m, where $m = 1/2, 1, 3/2, \ldots$. The terms c_α are linear combinations of finitely many (length α) products of the smooth vector fields V_i, $i = 0, 1, \ldots, d$. The following asymptotic convergence result can be established along the lines of the proof for linear stochastic differential equations in Lord, Malham and Wiese [31]; we provide a proof in Appendix A.

THEOREM 7.1. *Assume the vector fields V_i have $2m+1$ uniformly bounded derivatives, for all $i = 0, 1, \ldots, d$. Then for $t \leq 1$, the flow $\exp \hat{\psi}_t \circ y_0$ is square-integrable, where $\hat{\psi}_t$ is the truncated Lie series (7.1). Further, if y is the solution of the stochastic differential equation (1.1), there exists a constant $C(m, \|y_0\|_2)$ such that*

$$\left\| y_t - \exp \hat{\psi}_t \circ y_0 \right\|_{L^2} \leq C(m, \|y_0\|_2) \, t^{m+1/2} \,. \tag{7.2}$$

8. Geometric Castell–Gaines methods. Consider the truncated exponential Lie series $\hat{\psi}_{t_n, t_{n+1}}$ across the interval $[t_n, t_{n+1}]$. We approximate higher order multiple Stratonovich integrals across each time-step by their expectations conditioned on the increments of the Wiener processes on suitable subdivisions (Gaines and Lyons [20]). An approximation to the solution of the stochastic differential equation (1.1) across the interval $[t_n, t_{n+1}]$ is given by the flow generated by the truncated and conditioned exponential Lie series $\hat{\psi}_{t_n, t_{n+1}}$ via

$$y_{t_{n+1}} \approx \exp(\hat{\psi}_{t_n, t_{n+1}}) \circ y_{t_n} \,.$$

Hence the solution to the stochastic differential equation (1.1) can be approximately computed by solving the ordinary differential system (see Castell and Gaines [8]; Misawa [39])

$$u'(\tau) = \hat{\psi}_{t_n, t_{n+1}} \circ u(\tau) \tag{8.1}$$

across the interval $\tau \in [0, 1]$. Then if $u(0) = y_{t_n}$ we will get $u(1) \approx y_{t_{n+1}}$. We must choose a sufficiently accurate ordinary differential integrator to solve (8.1)—we implicitly assume this henceforth.

The set of governing vector fields V_i, $i = 0, 1, \ldots, d$, prescribes a map from the driving path process $w \equiv (W^1, \ldots, W^d)$ to the unique solution process $y \in \mathcal{M}$ to the stochastic differential equation (1.1). The map $w \mapsto y$ is called the Itô map. Recall that we assume the vector fields are smooth. When there is only one driving Wiener process ($d = 1$) the Itô map is continuous in the topology of uniform convergence (Theorem 1.1.1. in Lyons and Qian [33]). When there are two or more driving processes ($d \geq 2$) the Universal Limit Theorem (Theorem 6.2.2. in Lyons and Qian [33]) tells us that the Itô map is continuous in the p-variation topology, in particular for $2 \leq p < 3$. A Wiener path with $d \geq 2$ has p-variation with $p > 2$, and the p-variation metric in this case includes information about the Lévy chordal areas of the path (Lyons [32]). Hence we must choose suitable piecewise smooth approximations to the driving path process w. The following result follows from the corresponding result for ordinary differential equations in Hairer, Lubich and Wanner [22] (p. 112) as well as directly from Chapter VIII in Malliavin [36] on the Transfer Principle (see also Emery [15]).

LEMMA 8.1. *A necessary and sufficient condition for the solution to the stochastic differential equation* (1.1) *to evolve on a smooth n-dimensional submanifold $\mathcal{M}$ of $\mathbb{R}^N$ ($n \leq N$) up to a stopping time T_* is that $V_i(y, t) \in T_y\mathcal{M}$ for all $y \in \mathcal{M}$, $i = 0, 1, \ldots, d$.*

Hence the stochastic Taylor expansion for the flow φ_t is a diffeomorphism on $\mathcal{M}$. However a truncated version of the stochastic Taylor expansion for the flow $\hat{\varphi}_t$ will not in general keep you on the manifold, i.e. if $y_0 \in \mathcal{M}$ then $\hat{\varphi}_t \circ y_0$ need not necessarily lie in $\mathcal{M}$. On the other hand, the exponential Lie series ψ_t, or any truncation $\hat{\psi}_t$ of it, lies in $\mathfrak{X}(\mathcal{M})$. By Lemma 8.1 this is a necessary and sufficient condition for the corresponding flow-map $\exp \hat{\psi}_t$ to be a diffeomorphism on $\mathcal{M}$. Hence if $u(0) = y_{t_n} \in \mathcal{M}$, then $y_{t_{n+1}} \approx u(1) \in \mathcal{M}$. When solving the ordinary differential equation (8.1), classical geometric integration methods, for example Lie group integrators such as Runge–Kutta Munthe-Kaas methods, over the interval $\tau \in [0, 1]$ will numerically ensure $y_{t_{n+1}}$ stays in $\mathcal{M}$. Additionally, as the following result reveals, numerical methods constructed using the Castell–Gaines Lie series approach can also be more accurate (a proof is provided in Appendix B). We define the *strong global error* at time T associated with an approximate solution $\hat{y}_T$ as $\mathcal{E} \equiv \|y_T - \hat{y}_T\|_{L^2}$.

THEOREM 8.2. *In the case of two independent Wiener processes and under the assumptions of Theorem 7.1, for any initial condition $y_0 \in \mathcal{M}$ and a sufficiently small fixed stepsize $h = t_{n+1} - t_n$, the order $1/2$ Lie series integrator is globally more accurate in L^2 than the order $1/2$ stochastic Taylor integrator. In addition, in the case of one Wiener process, the order 1 and $3/2$ uniformly accurate exponential Lie series integrators generated by $\hat{\psi}^{(1)}_{t_n,t_{n+1}} = J_0 V_0 + J_1 V_1 + \frac{h^2}{12}\big([V_1, [V_1, V_0]]\big)$ and*

$$\hat{\psi}^{(3/2)}_{t_n,t_{n+1}} = J_0 V_0 + J_1 V_1 + \tfrac{1}{2}(J_{01} - J_{10})[V_0, V_1] + \tfrac{h^2}{12}\big([V_1, [V_1, V_0]]\big),$$

respectively, are globally more accurate in L^2 than their corresponding stochastic Taylor integrators. In other words, if $\mathcal{E}^{ls}_m$ denotes the global error of the exponential Lie series integrators of order m above, and $\mathcal{E}^{st}_m$ is the global error of the stochastic Taylor integrators of the corresponding order, then $\mathcal{E}^{ls}_m \leq \mathcal{E}^{st}_m$ for $m = 1/2, 1, 3/2$.

Remarks.

1. The result for $\hat{\psi}^{(3/2)}$ is new. That the order-$1/2$ Lie series integrator (for two Wiener processes) and the order 1 integrator generated by $\hat{\psi}^{(1)}$ are uniformly more accurate confirms the asymptotically efficient properties of these schemes proved by

Castell and Gaines [8]. The proof follows along the lines of an analogous result for linear stochastic systems considered in Lord, Malham and Wiese [31].

2. Consider the order $1/2$ exponential Lie series with no vector field commutations. Solving the ordinary differential equation (8.1) using an (ordinary) Euler Munthe-Kaas method and approximating $\mathrm{dexp}_\sigma^{-1} \approx \mathsf{id}$ is equivalent to the order $1/2$ stochastic Taylor Munthe-Kaas method (for the same Lie group and action).

9. Numerical examples.

9.1. Rigid body. We consider the dynamics of a rigid body such as a satellite (see Marsden and Ratiu [37]). We will suppose that the rigid body is perturbed by two independent multiplicative stochastic processes W^1 and W^2 with the corresponding vector fields $V_i(y) \equiv \xi_i(y)\, y$, for $i = 0, 1, 2$, with $\xi_i \in \mathfrak{so}(3)$. If we normalize the initial data y_0 so that $|y_0| = 1$ then the dynamics evolves on $\mathcal{M} = \mathbb{S}^2$. We naturally suppose $\mathcal{G} = SO(3)$, and $\Lambda_{y_0}(S) \equiv Sy_0$ so that $\lambda_{\xi_i}(y) = \xi_i(y)\, y$, and we can pull back the flow generated by V on $\mathcal{M}$ to the flow on $\mathcal{G}$ generated by $X_{\xi_i}(S, t) = \xi_i\big(\Lambda_{y_0}(S)\big)\, S$, $i = 0, 1, 2$. We use the following matrix representation for the $\xi_i(y) \in \mathfrak{so}(3)$:

$$\xi_i(y) = \begin{pmatrix} 0 & -y_3/\alpha_{i,3} & y_2/\alpha_{i,2} \\ y_3/\alpha_{i,3} & 0 & -y_1/\alpha_{i,1} \\ -y_2/\alpha_{i,2} & y_1/\alpha_{i,1} & 0 \end{pmatrix},$$

where the constants $\alpha_{i,j}$ for $j = 1, 2, 3$ are chosen so that the vector fields V_i and matrices ξ_i do not commute for $i = 0, 1, 2$: $\alpha_{0,1} = 3$, $\alpha_{0,2} = 1$, $\alpha_{0,3} = 2$, $\alpha_{1,1} = 1$, $\alpha_{1,2} = 1/2$, $\alpha_{1,3} = 3/2$, $\alpha_{2,1} = 1/4$, $\alpha_{2,2} = 1$, $\alpha_{2,3} = 1/2$. The vector fields V_i satisfy the conditions of Theorem 7.1 since the manifold is compact in this case.

We will numerically solve (1.1) using three different order 1 methods: stochastic Taylor, stochastic Taylor Munthe-Kaas based on (6.1) and Castell–Gaines (a standard non-geometric Runge–Kutta method is used to solve the ordinary differential equation (8.1)). The vector field compositions $V_i V_j$ needed for the stochastic Taylor and Castell–Gaines methods are readily computed. For the Munthe-Kaas method we note that we have $v_{\xi_i} \circ \mathsf{o} = \xi_i(y_0)$ and

$$v_{\xi_i} v_{\xi_j} \circ \mathsf{o} = \hat{A}(y_0, y_0; \alpha_i, \alpha_j) - \tfrac{1}{2}[\xi_i(y_0), \xi_j(y_0)].$$

Here $\mathsf{o} \in \mathfrak{so}(3)$ is the zero element on the Lie algebra, and for all $y, z \in \mathbb{R}^3$ we define

$$A(y, z; \alpha, \beta) \equiv \begin{pmatrix} \left(\frac{y_2 z_3}{\alpha_2} - \frac{y_3 z_2}{\alpha_3}\right)\frac{1}{\beta_1} \\ \left(\frac{y_3 z_1}{\alpha_3} - \frac{y_1 z_3}{\alpha_1}\right)\frac{1}{\beta_2} \\ \left(\frac{y_1 z_2}{\alpha_1} - \frac{y_2 z_1}{\alpha_2}\right)\frac{1}{\beta_3} \end{pmatrix},$$

and $\hat{\ }: \mathbb{R}^3 \to \mathfrak{so}(3)$ denotes the vector space isomorphism $\sigma \mapsto \hat{\sigma}$ where

$$\hat{\sigma} \equiv \begin{pmatrix} 0 & -\sigma_3 & \sigma_2 \\ \sigma_3 & 0 & -\sigma_1 \\ -\sigma_2 & \sigma_1 & 0 \end{pmatrix}.$$

Note that $\hat{y}\, z \equiv y \wedge z$ (see Marsden and Ratiu [37]). Note also since $\sigma \in \mathfrak{so}(3)$, $\exp \sigma \in SO(3)$ can be conveniently and cheaply computed using Rodrigues' formula (see Marsden and Ratiu [37] or Iserles *et al.* [25]).

In Figure 9.1 we show the distance from the manifold $\mathbb{S}^2$ of each the three approximations; we start with initial data $y_0 = (\sqrt{2}, \sqrt{2}, 0)^\mathrm{T}$. The stochastic Taylor Munthe-Kaas method can be seen to preserve the solution in the unit sphere to within machine
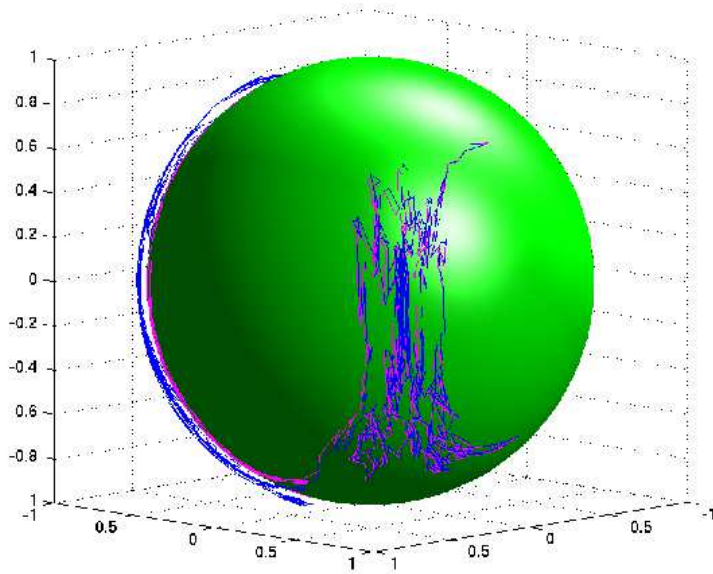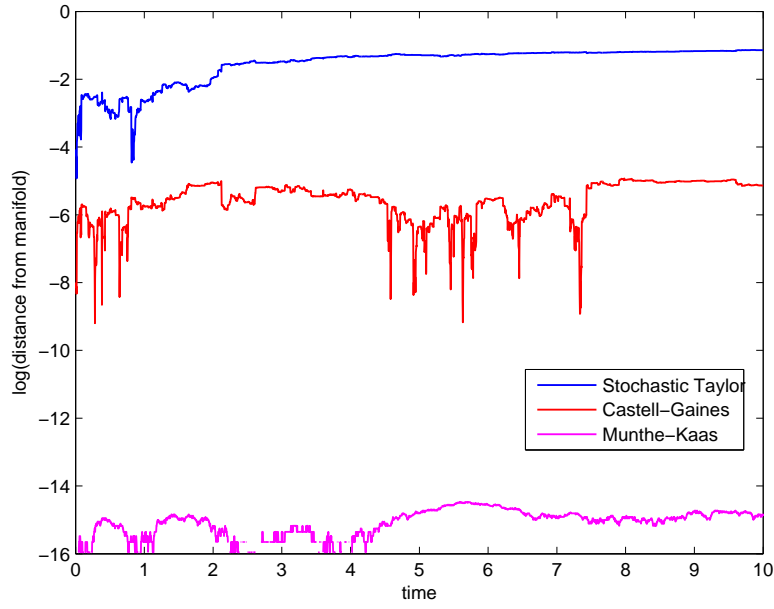
Fig. 9.1. *Rigid body: We show the log-distance of the approximate solution to the unit sphere as a function of time for each of the methods. Below we show the approximate solutions as a function of time for the stochastic Taylor (blue) and Munthe-Kaas methods (magenta). The trajectory starts at the top right and eventually drifts over the left horizon.*

error. We also see that the stochastic Taylor method clearly drifts off the sphere as the integration time progresses, as does the non-geometric Castell-Gaines method—which does however remain markedly closer to the manifold than the stochastic Taylor scheme.
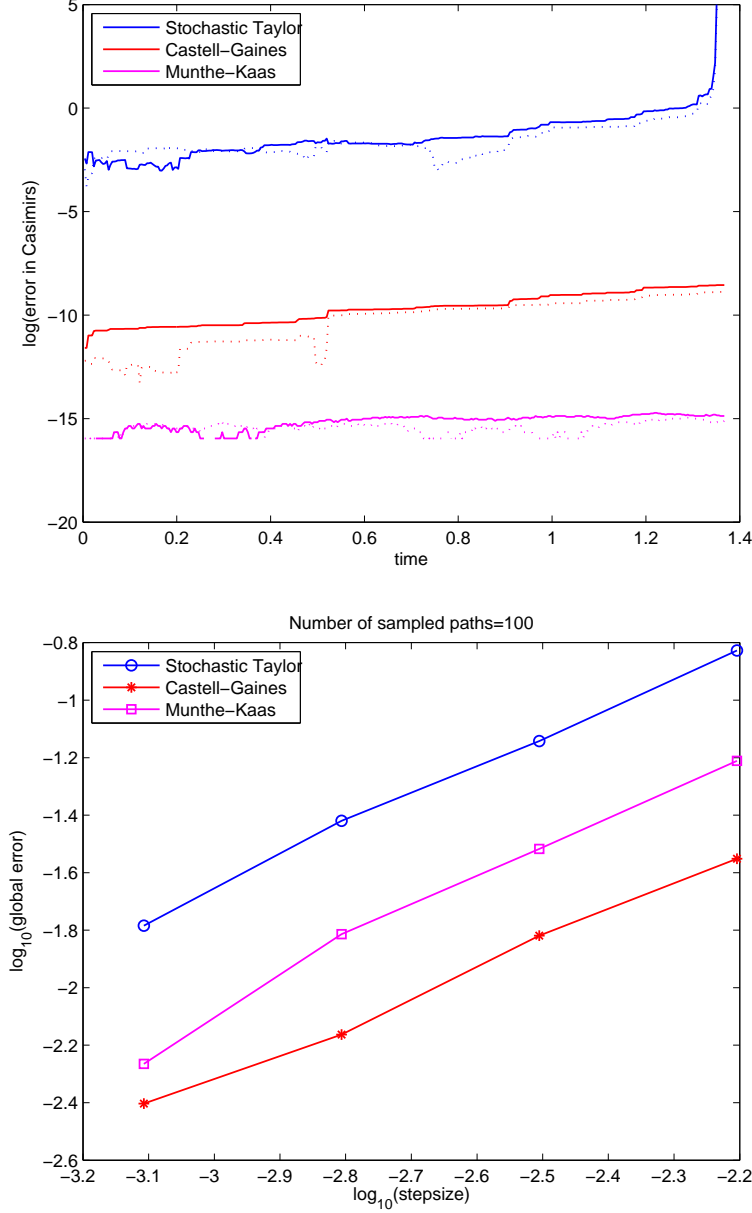

Fig. 9.2. *Autonomous underwater vehicle:* We show the log-distance of the approximate solution to the two Casimirs $C_1 = \pi \cdot p$ (dotted line) and $C_2 = |p|^2$ (solid line) as a function of time for each of the methods. Below, we also show the global error as a function of stepsize.

9.2. Autonomous underwater vehicle. The dynamics of an ellipsoidal autonomous underwater vehicle is prescribed by the state $y = (\pi, p) \in \mathfrak{se}(3)^*$ where $\pi \in \mathfrak{so}(3)^*$ is its angular momentum and $p \in (\mathbb{R}^3)^*$ its linear momentum (see Holmes, Jenkins and Leonard [24], Egeland, Dalsmo and Sørdalen [12] and Marsden and Ratiu [37]). We suppose that the vehicle is perturbed by two independent multiplicative stochastic processes. The governing vector fields are for $i = 0, 1, 2$:

$$V_i(y) = \mathrm{ad}^*_{\xi_i} \circ y \,.$$

Here $\xi_i(y) = \big(\omega_i(y), u_i(y)\big) \in \mathfrak{se}(3)$ where $\omega_i(y) = I_i^{-1}\pi$ and $u_i(y) = M_i^{-1}p$ are the angular and linear velocity, and $I_i = \mathrm{diag}(\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3})$ and $M_i = \mathrm{diag}(\beta_{i,1}, \beta_{i,2}, \beta_{i,3})$ are the constant moment of inertia and mass matrices, respectively. Explicitly for $\xi \in \mathfrak{se}(3)$ we have

$$\mathrm{ad}^*_{\xi} \circ y \equiv (\pi \wedge \omega + p \wedge u, p \wedge \omega) \,.$$

The system of vector fields V_i, $i = 0, 1, 2$ represents the Lie–Poisson dynamics on $\mathcal{M} = \mathfrak{se}(3)^*$ (Marsden and Ratiu [37]). There are two independent Casimir functions $C_k \colon \mathfrak{se}(3)^* \to \mathbb{R}$, $k = 1, 2$, namely $C_1 = \pi \cdot p$ and $C_2 = |p|^2$; these are conserved by the flow on $\mathfrak{se}(3)^*$. Note that the Hamiltonian, i.e. total kinetic energy $\frac{1}{2}(\pi \cdot \omega + p \cdot u)$, is also exactly conserved (and helpful for establishing the sufficiency conditions in Theorem 7.1), but that is not our focus here.

If $\mathcal{G} = SE(3) \cong SO(3) \times \mathbb{R}^3$, then the coadjoint action of $SE(3)$ on $\mathfrak{se}(3)^*$, $\mathrm{Ad}^* \colon SE(3) \times \mathfrak{se}(3)^* \to \mathfrak{se}(3)^*$ is defined for all $S = (s, \rho) \in SE(3)$, where $s \in SO(3)$ and $\rho \in \mathbb{R}^3$, and $y \in \mathfrak{se}(3)^*$ by: $\Lambda_y \circ S = \mathrm{Ad}^*_{S^{-1}} \circ y \equiv \big(s\pi + \rho \wedge (sp), sp\big)$. The corresponding infinitesimal action $\lambda \colon \mathfrak{se}(3) \times \mathfrak{se}(3)^* \to \mathfrak{se}(3)^*$ for all $\xi \in \mathfrak{se}(3)$ and $y \in \mathfrak{se}(3)^*$ is given by (see Marsden and Ratiu [37], p. 477)

$$\lambda_\xi \circ y = -\mathrm{ad}^*_\xi \circ y \,.$$

Since $\mathrm{ad}^*_\xi(y) = -\lambda_\xi(y) = \lambda_{-\xi}(y)$ the governing set of vector fields on $\mathfrak{se}(3)^*$ are

$$V_i(y) = \lambda_{-\xi_i} \circ y \,.$$

We can now pull back this flow on $\mathfrak{se}(3)^*$ to a flow on $SE(3)$ via Λ_{y_0}. The corresponding flow on $SE(3)$ is generated by the governing set of vector fields for $i = 0, 1, 2$:

$$X_{-\xi_i} \circ S = -\big(\omega_i(y) \wedge s, \omega_i(y) \wedge \rho + u_i(y)\big) \,,$$

with $y = \Lambda_{y_0}(S)$.

To aid implementation note that $SE(3) = \big\{(s, \rho) \in SE(3) \colon s \in SO(3), \ \rho \in \mathbb{R}^3\big\}$ embeds into $SL(4; \mathbb{R})$ via the map

$$S = (s, \rho) \mapsto \begin{pmatrix} s & \rho \\ O^{\mathrm{T}} & 1 \end{pmatrix} \,,$$

where O is the three-vector of zeros. Also $\mathfrak{se}(3)$ is isomorphic to a Lie subalgebra of $\mathfrak{sl}(4; \mathbb{R})$ with elements of the form

$$\sigma = (\theta, \zeta) \mapsto \begin{pmatrix} \hat{\theta} & \zeta \\ O^{\mathrm{T}} & 0 \end{pmatrix} \,.$$

Hence the governing vector fields on $SE(3)$ are of the form $X_{\xi_i} = -\xi_i(y)\,S$, where

$$\xi_i(y) = \begin{pmatrix} \hat{\omega}_i(\pi) & u_i(p) \\ O^{\mathrm{T}} & 0 \end{pmatrix}.$$

The governing vector fields on $\mathfrak{se}(3)$ are $v_{\xi_i}(\sigma) = -\mathrm{dexp}_\sigma \circ \xi_i\big(\Lambda_{y_0}(\exp\sigma)\big)$. Again the vector field compositions $V_i V_j$ needed for the stochastic Taylor and Castell–Gaines methods can be computed straightforwardly. Direct calculation also reveals that in block matrix form

$$v_{\xi_i} v_{\xi_j} \circ \mathbf{o} = \begin{pmatrix} \hat{A}(\pi_0, \pi_0; \alpha_i, \alpha_j) + \hat{A}(p_0, p_0; \beta_i, \alpha_j) & A(\pi_0, p_0; \alpha_i, \beta_j) \\ O^{\mathrm{T}} & 0 \end{pmatrix} - \tfrac{1}{2}\left[\xi_i(y_0), \xi_j(y_0)\right].$$

Here $A(y, z; \alpha, \beta)$ is defined as for the rigid body example. Note that the exponential map $\exp_{\mathfrak{se}(3)} \colon \mathfrak{se}(3) \to SE(3)$ is defined for all $\sigma = (\theta, \zeta) \in \mathfrak{se}(3)$ by

$$\exp_{\mathfrak{se}(3)} \sigma = \begin{pmatrix} \exp_{\mathfrak{so}(3)} \hat{\theta} & f(\theta)\zeta \\ O^{\mathrm{T}} & 1 \end{pmatrix},$$

where $\exp_{\mathfrak{so}(3)}$ is the exponential map from $\mathfrak{so}(3)$ to $SO(3)$ which can be computed using Rodrigues' formula and (see Bullo and Murray [4], p. 5)

$$f(\theta) = I_{3\times 3} + \big(1 - \cos\|\theta\|\big)\hat{\theta}/\|\theta\|^2 + \big(1 - (\sin\|\theta\|)/\|\theta\|\big)\hat{\theta}^2/\|\theta\|^2.$$

In Figure 9.2 we show the distance from the manifold $\mathfrak{se}(3)^*$ of each the three approximations; in particular how far the individual trajectories stray from the Casimirs $C_1 = \pi \cdot p$ and $C_2 = |p|^2$. We start with the initial data $y_0 = (\sqrt{2}, \sqrt{2}, 0, 0, \sqrt{2}, \sqrt{2})^{\mathrm{T}}$. As before the stochastic Taylor Munthe-Kaas method can be seen to preserve the Casimirs to within machine error. We also see that the stochastic Taylor method clearly drifts off the manifold as the integration time progresses and at a particular time depending on the Wiener path shoots off very rapidly away from the manifold. Note also that for large stepsizes the stochastic Taylor method is unstable. However the non-geometric Castell–Gaines and stochastic Munthe-Kaas methods still give reliable results in that regime. Lastly, although the the stochastic Munthe-Kaas method adheres to the manifold to within machine error, the error of the non-geometric Castell–Gaines method is actually smaller.

10. Conclusions. We have established and implemented stochastic Lie group integrators based on stochastic Munthe-Kaas methods and also derived geometric Castell–Gaines methods. We have also revealed several aspects of these integrators that require further investigation.

1. We could construct a stochastic nonlinear Magnus method by approximating the solution to the stochastic differential equation (5.1) on the Lie algebra using Picard iterations (see Casas and Iserles [6]).

2. We would like to develop a practical procedure for implementing ordinary Munthe-Kaas methods for higher order Castell–Gaines integrators. We need to determine the element $\xi \colon \mathcal{M} \to \mathfrak{g}$ so that in (8.1) we have $\hat{\psi} = \lambda_\xi$.

3. We need to determine the properties of the local and global errors for the stochastic Munthe-Kaas methods. Also a thorough investigation of the stability properties of the stochastic Munthe-Kaas and Castell–Gaines methods is required. For the autonomous underwater vehicle simulations they were both superior to the direct stochastic Taylor method, especially for larger stepsizes. We also need to compare the relative efficiency of the methods concerned, in particular to compare an optimally efficient geometric Castell–Gaines method with the stochastic Munthe-Kaas method.

4. Although we have chiefly confined ourselves to driving paths that are Wiener processes, we can extend Munthe-Kaas and Castell–Gaines methods to rougher driving paths (Lyons and Qian [33], Friz [18], Friz and Victoir [19]). Further, what happens when we consider processes involving jumps? For example Srivastava, Miller and Grenander [44] consider jump diffusion processes on matrix Lie groups for Bayesian inference. Or what if we consider fractional Brownian driving paths; Baudoin and Coutin [3] investigate fractional Brownian motions on Lie groups?

5. Schiff and Shnider [43] have used Lie group methods to derive Möbius schemes for numerically integrating deterministic Riccati systems beyond finite time removable singularities and numerical instabilities. They integrate a linear system of equations on the general linear group $\mathrm{GL}(n)$ which corresponds to a Riccati flow on the Grassmannian manifold $\mathrm{Gr}(k, n)$ via the Möbius action map. Lord, Malham and Wiese [31] implemented stochastic Möbius schemes and show that they can be more accurate and cost effective than directly solving stochastic Riccati systems using stochastic Taylor methods. We would like to investigate further their effectiveness for stochastic Riccati equations arising in Kalman filtering (Kloeden and Platen [27]) and to backward stochastic Riccati equations arising in optimal stochastic linear-quadratic control (see for example Kohlmann and Tang [28] and Estrade and Pontier [16]).

6. Other areas of potential application of the methods we have presented in this paper are for example: term-structure interest rate models evolving on finite dimensional invariant manifolds (see Filipovic and Teichmann [17]); stochastic dynamics triggered by DNA damage (Chickarmane, Ray, Sauro and Nadim [10]) and stochastic symplectic integrators for which the gradient of the solution evolves on the symplectic Lie group (see Milstein and Tretyakov [38]).

Appendix A. Proof of Theorem 7.1. We follow the proof for linear stochastic differential equations in Lord, Malham and Wiese [31] (where further technical details on estimates for multiple Stratonovich integrals can be found). Suppose $\hat{\psi}_t \equiv \hat{\psi}_t(m)$ is the truncated Lie series (7.1). First we show that $\exp \hat{\psi}_t \circ y_0 \in L^2$. We see that for any number k, $\left(\hat{\psi}_t\right)^k \circ y_0$ is a sum of $|\mathbb{Q}_m|^k$ terms, each of which is a k-multiple product of terms $J_\alpha\, c_\alpha \circ y_0$. It follows that

$$\left\| \left(\hat{\psi}_t\right)^k \circ y_0 \right\|_{L^2} \leq \left(\max_{\alpha \in \mathbb{Q}_m} \| c_\alpha \circ y_0 \| \right)^k \cdot \sum_{\substack{\alpha_i \in \mathbb{Q}_m \\ i=1,\ldots,k}} \| J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_k} \|_{L^2}. \tag{A.1}$$

Note that the maximum of the norm of the compositions of vector fields $c_\alpha \circ y_0$ is taken over a finite set. Repeated application of the product rule reveals that for $i = 1, \ldots, k$, each term '$J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_k}$' in (A.1) is the sum of at most 2^{2mk-1} Stratonovich integrals J_β, where for $t \leq 1$, $\| J_\beta \|_{L^2} \leq 2^{4mk-1}\, t^{k/2}$. Since the right hand side of equation (A.1) consists of $|\mathbb{Q}_m|^k\, 2^{2mk-1}$ Stratonovich integrals J_β, we conclude that,

$$\left\| \left(\hat{\psi}_t\right)^k \circ y_0 \right\|_{L^2} \leq \left(\max_{\alpha \in \mathbb{Q}_m} \| c_\alpha \circ y_0 \| \cdot |\mathbb{Q}_m| \cdot 2^{6m} \cdot t^{1/2} \right)^k.$$

Hence $\exp\hat{\psi}_t \circ y_0$ is square-integrable.

Second we prove (7.2). Let $\hat{y}_t$ denote the stochastic Taylor series solution, truncated to included terms of order up to and including t^m. We have

$$\left\|y_t - \exp\hat{\psi}_t \circ y_0\right\|_{L^2} \le \left\|y_t - \hat{y}_t\right\|_{L^2} + \left\|\hat{y}_t - \exp\hat{\psi}_t \circ y_0\right\|_{L^2}.$$

We know $y_t \in L^2$—see Lemma III.2.1 in Gihman and Skorohod [21]. Note that the assumptions there are fulfilled, since the uniform boundedness of the derivatives implies uniform Lipschitz continuity of the vector fields by the mean value theorem, and uniform Lipschitz continuity in turn implies a linear growth condition for the vector fields since they are autonomous. Note that $\hat{y}_t$ is a strong approximation to y_t up to and including terms of order t^m, with the remainder consisting of $\mathcal{O}(t^{m+1/2})$ terms (see Proposition 5.9.1 in Kloeden and Platen [27]). It follows from the definition of the exponential Lie series as the logarithm of the stochastic Taylor series, that the terms of order up to and including t^m in $\exp\hat{\psi}_t \circ y_0$ correspond with $\hat{y}_t$; the error consists of $\mathcal{O}(t^{m+1/2})$ terms.

Appendix B. Proof of Theorem 8.2. Our proof follows along the lines of that for uniformly accurate Magnus integrators for linear constant coefficient systems (see Lord, Malham & Wiese [31] and Malham and Wiese [35]). Let $\varphi_{t_n,t_{n+1}}$ and $\hat{\varphi}_{t_n,t_{n+1}}$ denote the exact and approximate flow-maps constructed on the interval $[t_n, t_{n+1}]$ of length h. We define the local flow remainder as

$$R_{t_n,t_{n+1}} \equiv \varphi_{t_n,t_{n+1}} - \hat{\varphi}_{t_n,t_{n+1}},$$

and so the local remainder is $R_{t_n,t_{n+1}} \circ y_{t_n}$. Let R^{ls} and R^{st} denote the local flow remainders corresponding to the exponential Lie series and stochastic Taylor approximations, respectively.

B.1. Order $1/2$ integrator: two Wiener processes. For the global order $1/2$ integrators we have to leading order $R^{\text{ls}} = \frac{1}{2}(J_{12} - J_{21})[V_1, V_2]$ and $R^{\text{st}} = J_{12}V_1V_2 + J_{21}V_2V_1$. Note that we have included the terms $J_{11}V_1^2$ and $J_{22}V_2^2$ in the integrators. A direct calculation reveals that

$$\mathbb{E}\left((R^{\text{st}} \circ y_0)^{\text{T}} R^{\text{st}} \circ y_0\right) = \mathbb{E}\left((R^{\text{ls}} \circ y_0)^{\text{T}} R^{\text{ls}} \circ y_0\right) + h^{2m} U^{\text{T}} B U + \mathcal{O}\left(h^{2m+\frac{1}{2}}\right). \quad \text{(B.1)}$$

Here $m = 1/2$ (for the order $1/2$ integrators), $U = (V_1V_2 \circ y_0, V_2V_1 \circ y_0)^{\text{T}} \in \mathbb{R}^{2n}$, and $B \in \mathbb{R}^{2n\times 2n}$ consists of $n \times n$ diagonal blocks of the form $b_{ij}I_{n\times n}$ where

$$b = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

and $I_{n\times n}$ is the $n \times n$ identity matrix. Since b is positive semi-definite, the matrix $B = b \otimes I_{n\times n}$ is positive semi-definite. Hence the order $1/2$ exponential Lie series integrator is locally more accurate than the corresponding stochastic Taylor integrator.

B.2. Order 1 integrator: one Wiener process. For the global order 1 integrators we have to leading order $R^{\text{ls}} = \frac{1}{2}(J_{01} - J_{10})[V_0, V_1]$ and $R^{\text{st}} = J_{01}V_0V_1 + J_{10}V_1V_0 + J_{111}V_1^3 + \frac{1}{4}h^2(V_0V_1^2 + V_1^2V_0)$. The terms of order h^2 shown are significant when we consider the global error in Section B.4 below. The estimate (B.1) also applies in this case with $m = 1$ and $U = (V_0V_1 \circ y_0, V_1V_0 \circ y_0, V_1^3 \circ y_0)^{\text{T}} \in \mathbb{R}^{3n}$; and $B \in \mathbb{R}^{3n\times 3n}$ consists of $n \times n$ diagonal blocks of the form $b_{ij}I_{n\times n}$ where

$$b = \frac{1}{12}\begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 5 \end{pmatrix}.$$

Since b is positive semi-definite, the matrix $B = b \otimes I_{n \times n}$ is positive semi-definite. Hence the order 1 exponential Lie series integrator is locally more accurate than the corresponding stochastic Taylor integrator.

B.3. Order $3/2$ integrator: one Wiener process. The local flow remainders are $R^{\mathrm{ls}} = \frac{1}{6}\big(J_{110} - 2J_{101} + J_{011} - \frac{1}{2}h^2\big)[V_1,[V_1,V_0]]$ and $R^{\mathrm{st}} = J_{011}V_0V_1^2 + J_{101}V_1V_0V_1 + J_{110}V_1^2V_0 + J_{1111}V_1^4 - \frac{1}{4}h^2(V_0V_1^2 + V_1^2V_0 + \frac{1}{2}V_1^4)$. The terms of order h^2 shown are significant when we consider the global error—but for a different reason this time—see Section B.4 below. Again, the estimate (B.1) applies in this case with $m = 3/2$ and $U = (V_0V_1^2 \circ y_0, V_1V_0V_1 \circ y_0, V_1^2V_0 \circ y_0, V_1^4 \circ y_0)^{\mathrm{T}} \in \mathbb{R}^{4n}$; and $B \in \mathbb{R}^{4n \times 4n}$ consists of $n \times n$ diagonal blocks of the form $b_{ij}I_{n \times n}$ where

$$b = \frac{1}{144}\begin{pmatrix} 11 & 8 & 5 & 12 \\ 8 & 8 & 8 & 12 \\ 5 & 8 & 11 & 12 \\ 12 & 12 & 12 & 24 \end{pmatrix}.$$

Again, B is positive semi-definite and the order $3/2$ exponential Lie series integrator is locally more accurate than the corresponding stochastic Taylor integrator.

B.4. Global error. Recall that we define the *strong global error* at time T associated with an approximate solution $\hat{y}_T$ as $\mathcal{E} \equiv \|y_T - \hat{y}_T\|_{L^2}$. The exact and approximate solutions can be constructed by successively applying the exact and approximate flow maps $\varphi_{t_n,t_{n+1}}$ and $\hat{\varphi}_{t_n,t_{n+1}}$ on the successive intervals $[t_n, t_{n+1}]$ to the initial data y_0. A straightforward calculation shows for a small fixed stepsize h,

$$\mathcal{E}^2 = \mathbb{E}\,(\mathcal{R} \circ y_0)^{\mathrm{T}}\,\mathcal{R} \circ y_0\,, \tag{B.2}$$

up to higher order terms, where $\mathcal{R} \equiv \sum_{n=0}^{N-1} \varphi_{t_{n+1},t_N} \circ R_{t_n,t_{n+1}} \circ \varphi_{t_0,t_n}$ is the standard accumulated local error contribution to the global error. The important conclusion is that when we construct the global error (B.2), the terms of leading order in the local flow remainders R^{ls} or R^{st} with zero expectation lose only a half order of convergence in this accumulation effect. Hence in the local flow remainders shown above, for the terms of zero expectation, the local superior accuracy for the Lie series integrators transfers to the corresponding global errors (see Lord, Malham and Wiese [31] for more details). Terms of non-zero expectation however behave like deterministic error terms losing a whole order (in the local to global convergence); they contribute to the global error through their expectations. Hence we include such terms of order h^2 in the order $3/2$ integrators above and they appear as the terms subtracted from the remainders shown. For the order 1 integrators we do not need to include the order h^2 terms in the integrator to obtain the correct mean-square convergence. However to guarantee that the global error for the exponential Lie series integrator is always smaller than that for the stochastic Taylor scheme, we include this term in the integrator.

REFERENCES

[1] G. Ben Arous, *Flots et series de Taylor stochastiques*, Probab. Theory Related Fields, 81 (1989), pp. 29–77.
[2] F. Baudoin, *An introduction to the geometry of stochastic flows*, Imperial College Press, 2004.
[3] F. Baudoin and L. Coutin, *Self-similarity and fractional Brownian motions on Lie groups*, arXiv:math.PR/0603199 v1, 2006.

[4]   F. Bullo and R. M. Murray, *Proportional derivative (PD) control on the Euclidean group*, CDS Technical Report 95-010, 1995.

[5]   K. Burrage and P. M. Burrage, *High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula*, Phys. D, 133 (1999), pp. 34–48.

[6]   F. Casas and A. Iserles, *Explicit Magnus expansions for nonlinear equations*, Cambridge NA reports, 2005.

[7]   F. Castell, *Asymptotic expansion of stochastic flows*, Probab. Theory Related Fields, 96 (1993), pp. 225–239.

[8]   F. Castell and J. Gaines, *An efficient approximation method for stochastic differential equations by means of the exponential Lie series*, Math. Comp. Simulation, 38 (1995), pp. 13–19.

[9]   K. T. Chen, *Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula*, Annals of Mathematics, 65(1) (1957), pp. 163–178.

[10] V. Chickarmane, A. Ray, H. M. Sauro and A. Nadim, *A model for p53 dynamics triggered by DNA damage*, SIAM J. Applied Dynamical Systems, 6(1) (2007), pp.61–78.

[11] P. E. Crouch and R. Grossman, *Numerical integration of ordinary differential equations on manifolds*, J. Nonlinear Sci., 3 (1993), pp. 1–33.

[12] O. Egeland, M. Dalsmo and O. J. Sørdalen, *Feedback control of a nonholonomic underwater vehicle with a constant desired configuration*, The International Journal of Robotics Research, 15(1) (1996), pp. 24–35.

[13] K. D. Elworthy, *Stochastic differential equations on manifolds*, London Mathematical Society Lecture Note Series 70, Cambridge University Press, 1982.

[14] M. Emery, *Stochastic Calculus on manifolds*, Universitext, Springer–Verlag, 1989.

[15] ———, *On two transfer principles in stochastic differential geometry*, Séminaire de probabilités (Strasbourg), 24 (1990), pp. 407–441.

[16] A. Estrade and M. Pontier, *Backward stochastic differential equations in a Lie group*, Séminaire de probabilités (Strasbourg), 35 (2001), pp. 241–259.

[17] D. Filipović and J. Teichmann, *On the geometry of the term structure of interest rates*, Proc. R. Soc. Lond. A, 460 (2004), pp. 129–167.

[18] P. Friz, *Continuity of the Itô-map for Hölder rough paths with applications to the support theorem in Hölder norm*, arXiv:math.PR/0304501 v2, 2003.

[19] P. Friz and N. Victoir, *Euler estimates for rough differential equations*, Preprint, 2007.

[20] J. G. Gaines and T. J. Lyons, *Variable step size control in the numerical solution of stochastic differential equations*, SIAM J. Appl. Math., 57(5) (1997), pp. 1455–1484.

[21] I. I. Gihman, and A. V. Skorohod, *The theory of stochastic processes III*, Springer, 1979.

[22] E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration*, Springer Series in Computational Mathematics, 2002.

[23] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Academic Press, 1978.

[24] P. Holmes, J. Jenkins and N. E. Leonard, *Dynamics of the Kirchoff Equations I: coincident centers of gravity and bouyancy*, Phys. D, 118 (1998), pp. 311–342.

[25] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, *Lie-group methods*, Acta Numer., (2000), pp. 215–365.

[26] A. Iserles and A. Zanna, *Efficient computation of the matrix exponential by generalized polar decompositions*, SIAM J. Numer. Anal., 42(5) (2005), pp. 2218–2256.

[27] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Springer, 1999.

[28] M. Kohlmann and S. Tang, *Multidimensional backward stochastic Riccati equations and applications*, SIAM J. Control Optim., 41(6) (2003), pp. 1696–1721.

[29] H. Kunita, *On the representation of solutions of stochastic differential equations*, LNM 784, Springer–Verlag, 1980, pp. 282–304.

[30] ———, *Stochastic flows and stochastic differential equations*, Cambridge University Press, 1990.

[31] G. Lord, S. J.A. Malham and A. Wiese, *Efficient strong integrators for linear stochastic systems*, 2006, Submitted.

[32] T. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana, 14(2) (1998), pp. 215–310.

[33] T. Lyons and Z. Qian, *System control and rough paths*, Oxford University Press, 2002.

[34] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Comm. Pure Appl. Math., 7 (1954), pp. 649–673.

[35] S. J.A. Malham and A. Wiese, *Universal optimal stochastic expansions*, 2007, Preprint.

[36] P. Malliavin, *Stochastic analysis*, Grundlehren der mathematischen Wissenschaften 313, Springer, 1997.

[37] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry*, Second edition,

Springer, 1999.

[38] G. N. MILSTEIN AND M. V. TRETYAKOV, *Stochastic numerics for mathematical physics*, Springer, 2004.

[39] T. MISAWA, *A Lie algebraic approach to numerical integration of stochastic differential equations*, SIAM J. Sci. Comput., 23(3) (2001), pp. 866–890.

[40] H. MUNTHE-KAAS, *High order Runge–Kutta methods on manifolds*, Appl. Numer. Math., 29 (1999), pp. 115–127.

[41] N. J. NEWTON, *Asymptotically efficient Runge–Kutta methods for a class of Itô and Stratonovich equations*, SIAM J. Appl. Math., 51 (1991), pp. 542–567.

[42] P. J. OLVER, *Equivalence, invariants, and symmetry*, Cambridge University Press, 1995.

[43] J. SCHIFF AND S. SHNIDER, *A natural approach to the numerical integration of Riccati differential equations*, SIAM J. Numer. Anal., 36(5) (1999), pp. 1392–1413.

[44] A. SRIVASTAVA, M. I. MILLER AND U. GRENANDER, *Jump-diffusion processes on matrix Lie groups for Bayesian inference*, preprint, 2000.

[45] R. S. STRICHARTZ, *The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations*, J. Funct. Anal., 72 (1987), pp. 320–345.

[46] H. J. SUSSMANN, *Product expansions of exponential Lie series and the discretization of stochastic differential equations, in Stochastic Differential Systems*, Stochastic Control Theory, and Applications, W. Fleming and J. Lions, eds., Springer IMA Series, Vol. 10 (1988), pp. 563–582.

[47] V. S. VARADARAJAN, *Lie groups, Lie algebras, and their representations*, Springer, 1984.

[48] F. W. WARNER, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, Springer–Verlag, 1983.

[49] Y. YAMATO, *Stochastic differential equations and nilpotent Lie algebras*, Z. Wahrsch. Verw. Gebiete, 47(2) (1979), pp 213–229.