



**HAL**  
open science

# Towards a BCH Formula on the Diffeomorphism Group with a Right-Invariant Metric

James Benn, Stephen Marsland

► **To cite this version:**

James Benn, Stephen Marsland. Towards a BCH Formula on the Diffeomorphism Group with a Right-Invariant Metric. 2023. hal-04360903v1

**HAL Id: hal-04360903**

**<https://hal.science/hal-04360903v1>**

Preprint submitted on 22 Dec 2023 (v1), last revised 24 May 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Towards a BCH Formula on the Diffeomorphism Group with a Right-Invariant Metric

James Benn · Stephen Marsland

Received: date / Accepted: date

**Abstract** The Baker-Campbell-Hausdorff (BCH) formula links the Lie group and Lie algebra by giving a formula for group multiplication of elements close to the identity in terms of the Lie bracket. This has applications in geometry, algebra, and partial differential equations, and provides simple proofs of further developments of Lie theory. To date, there is no BCH formula for the diffeomorphism groups, although they are infinite dimensional analogues of Lie groups, because the intrinsic group exponential map is not even locally surjective. In this paper we provide the relevant background in an expository manner, starting with the finite dimensional case, and then take a first step towards developing a BCH formula for the exponential map of a right-invariant metric on a Lie Group  $G$  by deriving an explicit formula for its derivative, in the spirit of Poincaré. We then identify features that allow for the finite dimensional results to lift to the infinite dimensional setting of the Diffeomorphism groups and re-prove our finite dimensional results on the diffeomorphism groups. Along the way we discuss additional applications of our formulas and conclude with a program to complete the BCH formula.

## Introduction

In a normal coordinate neighbourhood of the identity in a (finite dimensional) Lie group  $G$  with Lie algebra  $\mathfrak{g}$  the Baker-Campbell-Hausdorff (BCH) formula expresses group multiplication in terms of Lie algebraic data in the form of a power series whose terms are given by iterated Lie brackets. This provides an

---

JB has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 786854). SM is supported by Te Pūnaha Matatini, a New Zealand Centre of Research Excellence in Complex Systems.

---

J. Benn  
Université Côte d'Azur and INRIA, Epione team, France  
E-mail: james.benn@inria.fr

S. Marsland  
School of Mathematics and Statistics, Victoria University of Wellington, New Zealand  
E-mail: stephen.marsland@vuw.ac.nz

extremely useful general purpose tool with wide-reaching applications in physics, group theory, the analysis of linear PDEs, the structure theory of Lie groups and algebras, and even numerical analysis.

Achilles and Bonfiglioli [1] construct a fascinating timeline (summarised here) documenting the earliest contributions to the formula, which go back to Friedrich Heinrich Schur, Jules Henri Poincaré, and Ernesto Pascal. Around 1890, Schur developed an alternative approach to the foundations of Lie's theory of transformation groups; he explicitly parameterised a neighbourhood of the identity in  $G$  by a neighbourhood of the origin in  $\mathfrak{g}$  in such a way that group multiplication could be expressed as a convergent power series depending only on the algebra's structure constants and the Bernoulli numbers.

Pascal (1902) would later show how Schur's explicit parameterisation could be obtained from the BCH formula itself. The methods employed by Pascal, beginning in 1901, involve developing and deconstructing power series expansions for the group operation, and then reconstructing the result using a certain symmetrisation of polynomials from which an iterative formula in terms of Lie brackets and the Bernoulli numbers emerges. Campbell had earlier (1897) introduced a series similar to the one derived by Pascal, and used it to recursively construct a sequence consisting of Lie algebra polynomials approximating the group multiplication of two elements.

Poincaré (1900) effectively framed the problem as an ODE by determining an explicit formula for the derivative of the exponential map as a power series of an analytic function whose inverse is also a power series with the Bernoulli numbers as coefficients. He was able to obtain an explicit solution to the ODE via residue calculus, which expressed group multiplication as the exponential of an infinite series whose terms consist of polynomials in an indeterminate Lie algebra. In the process, Poincaré invented nothing less than the universal enveloping algebra of a Lie algebra, but what is striking is that the symmetrisation methods employed by Poincaré and Pascal are remarkably similar, despite the fact that Pascal was probably unaware of Poincaré's earlier work and development of the universal algebra.

Unfortunately, problems of universality and convergence lurked within these contributions, as pointed out by numerous critics including Engel, Hausdorff, and Bourbaki. These two issues were addressed by Baker (1905) and Hausdorff (1906), although only Hausdorff was credited by Bourbaki as a reliable source. However, in almost all of these early works a formula is actually difficult to find, and what *was* proved was that group multiplication *could* be expressed as the exponential of a power series involving Lie algebraic data; precisely, *in the associative algebra of the formal power series in two non-commuting indeterminates  $x$  and  $y$ , the series related to  $\log e^x e^y$  is a series of Lie polynomials.*

Forty years after Baker and Hausdorff, Dynkin provided the first explicit closed-form expression for  $\log e^x e^y$  in terms of iterated Lie brackets (note that the only other contributor to obtain a formula in terms of Lie brackets was Pascal).

Moving to the infinite dimensional (pseudo-) Lie groups, which is our primary focus here, things are more complicated. Infinite dimensional Lie groups that do admit a BCH formula are known as BCH-Lie groups or exponential-Lie groups. In the survey by Neeb [36] one can find the fundamental results on these Lie groups; in particular, every Banach Lie group is a BCH-Lie group and being a BCH-Lie group is more-or-less a topological property.

There is, however, a class of Lie groups that are not BCH: the diffeomorphism groups. There exist diffeomorphisms arbitrarily close to the identity diffeomorphism that do not embed in any flow, thus the intrinsic group exponential map is not even locally surjective – see [24],[38], [10], [30], [40] for the earliest and original examples, and [50], [25], or [37] for extensions and collected results. The situation is actually worse: there exist arc-wise connected free subgroups on infinitely many generators which, except for the identity, do not lie in the image of the exponential map [11]. In view of the simplicity of the identity component of the diffeomorphism group (see [47] for the original theorem or [3] for a proof) every diffeomorphism can be expressed as the product of a finite number of diffeomorphisms that *do* embed in a flow, but the surjectivity failure still forbids writing down any mathematical expression of the form:

$$e^{Z(t)} = e^{tX_1} \circ \dots \circ e^{tX_n}$$

which is also the starting point of Campbell’s, Poincaré’s, Baker’s and Hausdorff’s derivation of the BCH statement. In summary, there seems to be a fundamental algebraic and topological disconnect between the diffeomorphism group and its Lie algebra of vector fields.

On the other hand, Riemannian exponential maps of right-invariant metrics on diffeomorphism groups have been of interest ever since the work of Arnold [2], where he showed that the  $L^2$  metric on vector fields becomes right-invariant when restricted to the group of volume-preserving diffeomorphisms  $\mathcal{D}_\mu$  and its tangent space of divergence-free vector fields  $T_e\mathcal{D}_\mu$ . Moreover, geodesics of this metric correspond to solutions of the Euler equations of hydrodynamics on the underlying manifold  $M$  on which the diffeomorphism group acts. Ebin and Marsden [7] exploited Arnold’s observation to great effect and showed that the  $L^2$  metric is indeed a smooth Riemannian metric on  $\mathcal{D}_\mu$  that admits a smooth right-invariant Levi-Civita connection and, consequently, a smooth right-invariant exponential map that *is* a diffeomorphism in a neighbourhood of the identity.

Since then, much work has been done on the geometry and analysis of diffeomorphism groups with right-invariant metrics and their relationship to equations from continuum mechanics. Misiolek and Preston [34] have shown that many right-invariant metrics of Sobolev type admit “well behaved” exponential maps that are, themselves, diffeomorphisms in a neighbourhood of the identity element of the group on which they are defined, and are nonlinear Fredholm maps of index zero in the sense of Smale. Recall that a bounded linear operator between Banach spaces is said to be Fredholm if it has closed range and its kernel and co-kernel are finite dimensional; the index is defined as the difference between the dimension of the kernel and the dimension of the co-kernel. A smooth map  $f$  is said to be a Fredholm map if its Frechet derivative  $df(p)$  is a Fredholm operator at each  $p$ . In particular, when equipped with a certain right-invariant metric the diffeomorphism group and its algebra are reconnected, and a neighbourhood of the identity may be parameterised by infinitesimal transformations via an exponential map whose topological and analytic properties are similar to those of a finite dimensional exponential map.

This seems a promising goal and line of attack: to produce a BCH-type formula for the diffeomorphism group by using the exponential map of a right-invariant Sobolev type metric whose index is sufficiently high so that the Fredholm property is assured. Obviously, there is a price to pay when using these exponential maps:

most straight lines in the algebra are *not* mapped to one-parameter subgroups, which prevents any direct adaptation of the proof of the classical BCH formula and greatly complicates the analysis.

In this paper we take a first step towards a BCH formula on the diffeomorphism group by deriving an explicit formula for the derivative of the exponential map defined by a right-invariant metric on a Lie group  $G$  (Theorem 1), in the spirit of Poincaré. This is done first for a finite dimensional Lie group  $G$  with Lie algebra  $\mathfrak{g}$  and arbitrary right-invariant metric  $\langle \cdot, \cdot \rangle$ . The formula we obtain reduces to the known formula for the group exponential map when  $\langle \cdot, \cdot \rangle$  is bi-invariant (assuming  $G$  allows for one). Our derivation is different from the original and operator theoretic in nature.

As an immediate application, of independent interest, we give a dynamical characterisation of the conjugate locus of a right-invariant metric (Theorem 4), which provides a new geometric description of conjugate points. In addition, we use the formula to show how in/stability of the geodesic system measured through linearisation of the geodesic equation (Eulerian stability) is related to in/stability measured through linearisation of the corresponding flow. When  $G$  is compact and the metric is bi-invariant, for example in a symmetric space, these two notions of stability coincide; our two examples (Theorems 4 and 6) illustrate the topological and algebraic influence on the disconnect between these two in/stability measures when the symmetry is broken by a one-sided invariant metric. The conclusion is that if one has control over the algebra and topology then it is enough to study the first order stability problem (Eulerian stability) to understand the second order stability problem (Lagrangian stability) which describes the geometry of  $G$ .

Following the finite dimensional analysis we reformulate the results in the context of right-invariant Sobolev metrics on the infinite-dimensional diffeomorphism groups using the theory of [34], and show that our finite dimensional results persist whenever the exponential maps have the Fredholm property. These results become immediately relevant to the stability of the Euler equations of hydrodynamics and show that  $L^2$  Eulerian and Lagrangian stabilities, as well as instabilities, coincide; thus the  $L^2$  metric on the volume-preserving diffeomorphism groups mimics the behaviour of a bi-invariant metric on a compact Lie group. Both the conjugate point and stability results seem directly relevant to a problem (problem 11) described in [20].

The purpose of presenting our finite dimensional results first is so that the reader can see the derivation and application clearly without having to worry about function spaces and regularity assumptions from the outset. Identifying which Fredholm properties allow for the finite dimensional proofs to lift to the infinite dimensional setting then greatly simplifies the second part.

It is worth explaining the potential interest in a "right-invariant" BCH formula. While we know that the intrinsic group exponential map is not locally surjective, we do not have a good understanding of what the image of the group exponential map is. The results of Palis [38] show that the set of diffeomorphisms generated by vector fields form a set of first Baire category; consequently, cataloging vector fields and the diffeomorphisms they integrate to, although interesting, won't necessarily identify those diffeomorphisms that do not embed any flow. Given that the exponential map of a sufficiently regular right-invariant metric on the diffeomorphism group is a local diffeomorphism, an accompanying BCH formula could assist in characterising those diffeomorphisms that lie in the image of the group

exponential map and those that do not, while also quantifying the extent to which the classical BCH formula fails to hold on the diffeomorphism group. For example, the results of Zajtz [50] show that Anosov diffeomorphisms do not embed in any flow: on the one hand these diffeomorphisms exhibit exponential stretching and compressing, while on the other hand they are structurally stable. Understanding the qualitative stability properties of those diffeomorphisms which do not embed in any flow could have interesting consequences for autonomous dynamical systems, as-well as time-dependent systems where Anosov diffeomorphisms could embed. In view of the above remarks on the simplicity of the diffeomorphism groups, a BCH formula may also aid in determining the minimum number of embeddable diffeomorphisms that need to be composed together in order to reach a diffeomorphism that does not embed. Additionally, a qualitative and quantitative understanding of the failure of the classical BCH formula in the diffeomorphism group setting is important for the Stationary Velocity Fields method in medical imaging and registration since it is a key tool for the development of efficient gradient-based optimization methods (see Chapter 5 of Pennec et al. [39]).

We close the article with a discussion on the difficulties in completing the BCH formula from the results of this work, and outline a program of future development. The central problem involves dealing with spectral representations of differences and products of non-commuting operators, thus we require an appropriate functional calculus for systems of non-commuting operators.

The paper is structured as follows: we first provide a very accessible introduction to the algebra and right-invariant geometry on finite dimensional Lie groups. This will obviously be of use for the finite dimensional results, but will also highlight the difficulties of working on diffeomorphism groups, which will be discussed in the second half of the paper. We will discuss Arnold's framework for studying geodesics of one-sided metrics on Lie groups and derive the basic results, along with a new form of the geodesic equations (Theorem 2). This will set us up for analysing Jacobi fields and the derivative of the right-invariant exponential map. Theorem 3 contains our main results on the derivative of the right-invariant exponential map, while Theorems 4, 5, and 6 contain applications to conjugate points and geodesic stability. We then proceed to discuss diffeomorphism groups. A self-contained presentation of the analysis and algebra of diffeomorphism groups is impossible here and so we have done our best to summarise the aspects relevant for our purposes while referring the reader to several comprehensive sources. The goal is to identify which features of the finite dimensional geometry can be reformulated in the infinite dimensional setting with comparable consequences. Theorem 9 contains the diffeomorphism analogue of the derivative of the right-invariant exponential map while Theorems 10 and 11 contain applications to conjugate points and geodesic stability. Finally, potential BCH formulas and their difficulties are discussed.

## Algebra and Geometry

Let  $G$  be a finite dimensional Lie group with group operation  $\star$  and identity element  $e$ . The classical operations of right and left translation are

$$\begin{aligned} R : G \times G &\rightarrow G & L : G \times G &\rightarrow G \\ R(g, h) &:= R_g(h) = h \star g & L(g, h) &:= L_g(h) = g \star h \end{aligned}$$

and for a fixed  $g \in G$  both  $R_g$  and  $L_g$  are smooth diffeomorphisms of  $G$ . A vector field  $V$  on  $G$  is said to be right-invariant if the derivative  $dR_g$  satisfies

$$dR_g \cdot V(h) = V(R_g(h)).$$

Analogously, a vector field  $U$  on  $G$  is said to be left-invariant if

$$dL_g \cdot U(h) = U(L_g(h)).$$

In what follows we will denote a vector field by a capital letter, for example  $V$ , and it's value at  $e$  by the corresponding lower case letter, for example  $V(e) = v$ .

From here we can deduce several facts about right/left-invariant vector fields: (1) the set of right/left-invariant vector fields on  $G$  is isomorphic to the tangent space at the identity  $T_e G =: \mathfrak{g}$  via the evaluation at  $e$  map; (2) right/left-invariant vector fields are smooth which can be seen by writing the map  $g \mapsto v(f \circ R_g)$  as  $\frac{d}{dt}|_{t=0} f \circ R(g, \eta(t))$  for any curve  $\eta(t)$  with  $\eta(0) = e$  and  $\frac{d}{dt}|_{t=0} \eta(t) = v$ , and any  $f \in C^\infty(G)$ ; (3) the Lie bracket of two right/left-invariant vector fields is itself right/left-invariant which follows from general principles of diffeomorphisms and the Lie bracket; and (4) the vector space  $\mathfrak{g}$  forms a Lie algebra under the Lie bracket of right/left-invariant vector fields. Specifically, the bracket operation  $[\cdot, \cdot]_{\mathfrak{g}}$  in  $\mathfrak{g}$  is determined by:

$$[V, W] = [dR_g \cdot v, dR_g \cdot w] = dR_g \cdot [v, w]_{\mathfrak{g}} \quad (1)$$

$$[V, W] = [dL_g \cdot v, dL_g \cdot w] = dL_g \cdot [v, w]_{\mathfrak{g}} \quad (2)$$

depending on whether one is considering right-, or left-, invariant vector fields, respectively.

An operator that will feature significantly in later sections is the group adjoint operator  $\text{Ad} : G \rightarrow \text{Aut}(\mathfrak{g})$  which gives a representation of  $G$  through an action on its Lie algebra  $\mathfrak{g}$

$$\begin{aligned} \text{Ad}_g &: \mathfrak{g} \rightarrow \mathfrak{g} \\ \text{Ad}_g v &= dL_g dR_{g^{-1}} \cdot v. \end{aligned} \quad (3)$$

Because the left and right translation maps commute their differentials commute as-well and so the ordering of the differentials in (3) will not always be respected but rather chosen according to what is convenient in the context (see the proof of Proposition 1, for example).

The differential of  $\text{Ad}$  is the map  $\text{ad} : \mathfrak{g} \rightarrow \text{Aut}(\mathfrak{g})$  and gives a complementary representation of  $\mathfrak{g}$  on itself

$$\begin{aligned} \text{ad}_v &: \mathfrak{g} \rightarrow \mathfrak{g} \\ \text{ad}_v w &= \frac{d}{dt}|_{t=0} \text{Ad}_{\eta(t)} w, \end{aligned} \quad (4)$$

where  $\eta(t)$  is any smooth curve in  $G$  passing through the identity  $e$  at  $t = 0$  with velocity  $v$ .

To relate these operators to the Lie bracket in  $\mathfrak{g}$  we need to understand the flows of right/left-invariant vectors and it is convenient to introduce the group exponential map  $\exp_e^G$  for this. Following Warner [48], the map

$$\lambda \cdot \frac{d}{dt} \mapsto \lambda \cdot v \in \mathfrak{g}$$

is a  $C^\infty$  Lie algebra homomorphism from  $\mathbb{R}$  into  $\mathfrak{g}$ . Since  $\mathbb{R}$  is simply connected there exists a *unique*  $C^\infty$  homomorphism from  $\mathbb{R}$  into  $G$  which we call

$$\exp_v : \mathbb{R} \rightarrow G$$

and has the property

$$d\exp_v \left( \lambda \cdot \frac{d}{dt} \right) = \lambda \cdot v.$$

- see Theorem 3.27 [48]. In other words, the map  $t \mapsto \exp_v(t)$  is the unique one-parameter subgroup of  $G$  whose tangent vector at the identity  $e$  is  $v$ . Define the group exponential map

$$\begin{aligned} \exp_e^G : \mathbb{R} \times \mathfrak{g} &\rightarrow G \\ \exp_e^G(tv) &= \exp_v(t). \end{aligned} \quad (5)$$

To be clear, the group homomorphism property implies two things about  $\exp_e^G(tv)$ :

$$\exp_e^G(sv) \star \exp_e^G(tv) = \exp_e^G((s+t)v), \quad \exp_e^G(tv)^{-1} = \exp_e^G(-tv)$$

which is what it means for  $\exp_e^G(tv)$  to be a one-parameter subgroup of  $G$ .

If  $V$  is a right-invariant vector field on  $G$  then the fundamental theorem on existence and uniqueness of ODE's says there exists a one-parameter subgroup of diffeomorphisms  $\varphi_t^V$  acting on  $G$  and satisfying

$$\partial_t \varphi_t^V(g) = V \left( \varphi_t^V(g) \right).$$

Since  $\exp_e^G(tv)$  is the unique one parameter subgroup (5) of  $G$  passing through the identity in the direction  $v = V(e)$  the flow  $\varphi_t^V(e)$  through the identity is given by  $\exp_e^G(tv)$ . Right translating  $\exp_e^G(tv)$  by an element  $g$  we see that  $R_g \exp_e^G(tv)$  satisfies the ODE

$$\partial_t R_g \exp_e^G(tv) = dR_g V(\exp_e^G(tv)) = V(R_g \exp_e^G(tv)),$$

where we have used the right-invariance of  $V$ . So  $R_g \exp_e^G(tv)$  satisfies the same ODE as  $\varphi_t^V(g)$  and has the same initial conditions and therefore the two curves must coincide by uniqueness of solutions to ODE's. Consequently, the flow  $\varphi_t^V(g)$  is given by

$$\varphi_t^V(g) = R_g \left( \exp_e^G(tv) \right) = \exp_e^G(tv) \star g = L_{\exp_e^G(tv)}(g). \quad (6)$$

An identical argument shows that the flow  $\psi_t^U$  of a left-invariant vector field  $U$  is given by

$$\psi_t^U(g) = R_{\exp_e^G(tv)}(g). \quad (7)$$

*Remark 1* Normally the group exponential map is defined by

$$\begin{aligned} \exp_e^G : \mathfrak{g} &\rightarrow G \\ \exp_e^G(v) &= \exp_v(1) \end{aligned}$$

and then it is proved, using only the homomorphism property of  $\exp_v$ , that  $\exp_v(t)$  is the unique integral curve of the right/left-invariant vector field  $V$  whose value at the identity is  $v$ , from which it can be deduced that  $\exp_e^G(tv) = \exp_v(t)$  (see Theorem 3.31 of [48]). Knowing this we have chosen to take  $\exp_e^G(tv) = \exp_v(t)$  as the definition and work backwards.



**Proposition 1** *Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ .*

1. *If  $X$  and  $Y$  are two left-invariant vector fields on  $G$  then the algebra adjoint operator (4) is given by the Lie bracket in  $\mathfrak{g}$*

$$\mathrm{ad}_x y = dL_{g^{-1}} \cdot [X, Y] = [x, y]_{\mathfrak{g}}$$

2. *If  $V$  and  $W$  are two right-invariant vector fields on  $G$  then the algebra adjoint operator (48) is given by minus the Lie bracket in  $\mathfrak{g}$*

$$\mathrm{ad}_v w = -dR_{g^{-1}} \cdot [V, W] = -[v, w]_{\mathfrak{g}}$$

*Proof.* (1) Proofs are better without surprises: what we're going to show is

$$\mathrm{ad}_{X_e} Y_e = \frac{d}{dt} \Big|_{t=0} d\psi_{-t}^X Y \left( \psi_t^X(e) \right) =: [X, Y](e) = dR_{g^{-1}} \cdot [X, Y](g) = [X_e, Y_e]_{\mathfrak{g}}$$

where  $\psi_t^X$  is the flow of the left-invariant vector field  $X$  given in (7). Using the definition of the algebra adjoint (4), the definition of the group adjoint (3), left-invariance, the flow (7), and the definition of the Lie bracket of vector fields we compute

$$\begin{aligned} \mathrm{ad}_{X_e} Y_e &= \frac{d}{dt} \Big|_{t=0} \mathrm{Ad}_{\exp_e^G(tX_e)} Y_e = \frac{d}{dt} \Big|_{t=0} dR_{\exp_e^G(-tX_e)} dL_{\exp_e^G(tX_e)} Y_e \\ &= \frac{d}{dt} \Big|_{t=0} dR_{\exp_e^G(-tX_e)} Y \left( L_{\exp_e^G(tX_e)}(e) \right) = \frac{d}{dt} \Big|_{t=0} d\psi_{-t}^X \cdot Y \left( R_{\exp_e^G(tX_e)}(e) \right) \\ &= \frac{d}{dt} \Big|_{t=0} d\psi_{-t}^X \cdot Y \left( \psi_t^X(e) \right) =: [X, Y](e) \\ &= dL_{g^{-1}} \cdot [X, Y](g) = [X_e, Y_e]_{\mathfrak{g}}. \end{aligned}$$

The statement is completed by using left-invariance of the vector fields and their Lie bracket.

- (2) Details will not be skipped: as above we will show that

$$\mathrm{ad}_v w = \frac{d}{dt} \Big|_{t=0} d\psi_t^V W \left( \psi_{-t}^V(e) \right) =: [-V, W](e) = -dR_{g^{-1}} \cdot [V, W](g) = -[v, w]_{\mathfrak{g}}$$

but notice that the time parameter has the opposite sign. If  $\varphi_t^V$  is the flow of the right-invariant vector field  $V$  then  $\varphi_{-t}^V$  is the flow of the right-invariant vector field  $-V$ . The difference between the algebra adjoint operator for left-invariant vector fields and the algebra adjoint operator for right-invariant vector fields is due to the way in which left/right-invariance absorbs the respective factors appearing in the group adjoint operator  $\mathrm{Ad}$ . Observe: using the definition of the algebra adjoint (4), the definition of the group adjoint (3), right-invariance, the flow (6) and the definition of the Lie bracket

$$\begin{aligned} \mathrm{ad}_v w &= \frac{d}{dt} \Big|_{t=0} \mathrm{Ad}_{\exp_e^G(tv)} w = \frac{d}{dt} \Big|_{t=0} dL_{\exp_e^G(tv)} dR_{\exp_e^G(-tv)} w \\ &= \frac{d}{dt} \Big|_{t=0} dL_{\exp_e^G(tv)} W \left( R_{\exp_e^G(-tv)}(e) \right) = \frac{d}{dt} \Big|_{t=0} d\varphi_t^V \cdot W \left( L_{\exp_e^G(-tv)}(e) \right) \\ &= \frac{d}{dt} \Big|_{t=0} d\varphi_t^V \cdot W \left( \varphi_{-t}^V(e) \right) =: [-V, W](e) \\ &= -dR_{g^{-1}} \cdot [V, W](g) = -[v, w]_{\mathfrak{g}} \end{aligned}$$

As in (1) the statement is completed by using right-invariance of the vector fields and their Lie bracket.

□

□

*Remark 2* The relationship between the contents of Proposition 1 and the differential geometry of a right/left-invariant metric on a Lie group has been opaque for quite some time in the literature. Proposition 1 is purely algebraic and has absolutely nothing to do with a choice of Riemannian metric on  $G$ ; rather, it is analogous to a choice of orientation on the group. We will refer to the “algebraic orientation” in (1) as the left-handed algebra and the “algebraic orientation” in (2) the right-handed algebra. When choosing a right/left-invariant metric on a Lie group it makes sense to pair the right/left-invariant geometry with the corresponding right/left-handed algebra but only because it simplifies the relationship between the geometry and the algebra - there is no mathematical implication in either direction.

But speaking of geometry, we haven’t seen any yet. So let’s give  $G$  a right-invariant Riemannian metric  $\langle \cdot, \cdot \rangle$  - there is no obstruction in doing so since we can give  $\mathfrak{g}$  any inner product and smoothly extend it to the rest of the group via

$$\langle V_g, W_g \rangle_g = \langle dR_{g^{-1}} \cdot v_g, dR_{g^{-1}} \cdot W_g \rangle_e \quad \forall V_g, W_g \in T_g G.$$

From now on we will restrict to right-invariant objects but the reader can make the necessary changes for left-invariance using the above algebra.

We have the usual geometric objects associated with  $\langle \cdot, \cdot \rangle$ ; namely, a smooth Levi-Civita connection  $\nabla$  and smooth Riemann curvature tensor  $R$ . Since right translation  $R_g : G \rightarrow G$  is a global isometry of the metric  $\langle \cdot, \cdot \rangle$  both  $\nabla$  and  $R$  are invariant under the pullback action of  $R_g$  which follows from general properties of Levi-Civita connections (see Lee [26], for example):  $R_g^* \nabla = \nabla$  and  $R_g^* R = R$ . In symbols, for any  $g \in G$  and any vector fields  $U, V$  and  $W$  on  $G$

$$dR_g \cdot \nabla_V W = \nabla_{dR_g \cdot V} dR_g \cdot W \quad (8)$$

$$dR_g \cdot R(U, V)W = R(dR_g \cdot U, dR_g \cdot V)dR_g \cdot W. \quad (9)$$

So just like the metric  $\langle \cdot, \cdot \rangle$ , both  $\nabla$  and  $R$  are themselves right-invariant. With this in mind let’s relate the geometry back to the algebra.

The group coadjoint operator  $\text{Ad}_g^* : \mathfrak{g} \rightarrow \mathfrak{g}$  is the metric adjoint of the group adjoint operator

$$\langle \text{Ad}_g^* v, w \rangle = \langle v, \text{Ad}_g w \rangle, \quad \forall v, w \in \mathfrak{g} \quad (10)$$

while the algebra coadjoint operator  $\text{ad}_u^* : \mathfrak{g} \rightarrow \mathfrak{g}$  is the metric adjoint of the algebra adjoint operator

$$\langle \text{ad}_u^* v, w \rangle = \langle v, \text{ad}_u w \rangle, \quad \forall v, w \in \mathfrak{g}. \quad (11)$$

**Proposition 2** *Let  $G$  be a Lie group with a right-handed Lie algebra  $\mathfrak{g}$ , and right-invariant metric  $\langle \cdot, \cdot \rangle$ . Suppose  $V$  and  $W$  are any two right-invariant vector fields on  $G$ . Then*

$$\nabla_V W = -\frac{1}{2} dR_g \cdot (\text{ad}_v w - \text{ad}_v^* w - \text{ad}_w^* v)$$

*Proof.* Let  $Z$  be any other right-invariant vector field on  $G$ . Since the metric  $\langle \cdot, \cdot \rangle$  is right invariant the functions

$$\langle V, W \rangle, \quad \langle V, Z \rangle, \quad \langle W, Z \rangle$$

are all constant and Koszul's formula for the connection gives

$$2\langle \nabla_V W, Z \rangle = \langle [V, W], z \rangle - \langle [W, Z], V \rangle - \langle [V, Z], W \rangle.$$

Since  $V$ ,  $W$ , and  $Z$  are right invariant their Lie brackets are right-invariant and since the connection is right-invariant (8) Koszul's formula collapses to the Lie algebra as

$$2\langle \nabla_v w, z \rangle = \langle [v, w]_{\mathfrak{g}}, z \rangle - \langle [w, z]_{\mathfrak{g}}, v \rangle - \langle [v, z]_{\mathfrak{g}}, w \rangle.$$

Using Proposition 1 (2) to write the Lie bracket in the algebra in terms of the adjoint operator  $\text{ad}_v w = -[v, w]_{\mathfrak{g}}$  and taking metric adjoints we arrive at

$$2\langle \nabla_v w, z \rangle = -\langle \text{ad}_v w, z \rangle + \langle z, \text{ad}_w^* v \rangle + \langle z, \text{ad}_v^* w \rangle.$$

Since this holds for any vector  $z \in \mathfrak{g}$  we obtain

$$\nabla_v w = -\frac{1}{2} (\text{ad}_v w - \text{ad}_w^* v - \text{ad}_v^* w).$$

Multiplying both sides by  $dR_g$  and using (8) once more gives the Proposition.  $\square$

*Remark 3* Proposition 2 does *not* hold for arbitrary smooth vector fields on  $G$ . If  $\{e_1, \dots, e_n\}$  is an orthonormal basis of  $\mathfrak{g}$  then we can right translate these vectors around the group  $G$  to obtain a global right-invariant orthonormal frame  $\{E_1, \dots, E_n\}$ . Writing the vector fields  $V$  and  $W$  in terms of the global frame ( $E_i$ ):

$$\begin{aligned} V &= v^1(g)E_1 + \dots + v^n(g)E_n \\ W &= w^1(g)E_1 + \dots + w^n(g)E_n \end{aligned}$$

we see that the covariant derivative  $(\nabla_V W)(e)$  depends on the values of the functions  $w^i(g)$  in a neighbourhood of the identity due to the Leibniz rule. But compare this with Marsden, Ratiu, and Raugel ("Symplectic Connections and the Linearisation of Hamiltonian Systems" - Proposition 3.1 and Corollary 3.2) or Bao, Lafontaine, and Ratiu ("On a Non-Linear Equation Related to the Geometry of the Diffeomorphism Group" - Proposition 1) where it is shown that any connection is uniquely determined by its values on right-invariant vector fields, albeit through a different relationship.

The tangent vector of a curve  $\eta : I \rightarrow G$  defined on an interval  $I \subset \mathbb{R}$  containing zero such that  $\eta(0) = e$  can be right-translated back to the identity  $e$  to give a curve  $v(t) \in \mathfrak{g}$  and we may write

$$\partial_t \eta(t) = dR_{\eta(t)} \cdot v(t).$$

If  $U : I \rightarrow TG$  is any vector field along  $\eta(t)$ , i.e.  $U(t) \in T_{\eta(t)}G$ , then it too can be right-translated back to the identity to give a curve  $u(t) \in \mathfrak{g}$ :

$$U(t) = dR_{\eta(t)} \cdot u(t).$$

The  $t$ -derivative (covariant derivative) of  $U(t)$  along  $\eta(t)$  is given as

$$\begin{aligned}\partial_t U(t) &= dR_{\eta(t)} \cdot \partial_t u(t) + \nabla_{\partial_t \eta} dR_{\eta(t)} \cdot u(t) \\ &= dR_{\eta(t)} \cdot (\partial_t u(t) + \nabla_{v(t)} u(t))\end{aligned}\quad (12)$$

where right-invariance (8) of the connection has been used and the covariant derivative in the Lie algebra has the interpretation given by Proposition 2.

Proposition 1 and 2 also give us a way of computing the Lie bracket of vector fields along curves in  $\mathfrak{g}$ . If  $U : I \rightarrow G$  and  $V : I \rightarrow G$  are two vector fields along a curve  $\eta : I \rightarrow G$  with  $dR_{\eta(t)} \cdot U(t) = u(t) \in \mathfrak{g}$  and  $dR_{\eta(t)} \cdot V(t) = v(t) \in \mathfrak{g}$  then subtracting the formula for  $\nabla_{v(t)} u(t)$  in Proposition 2 from the formula for  $\nabla_{u(t)} v(t)$  in Proposition 2

$$\nabla_{u(t)} v(t) - \nabla_{v(t)} u(t) = [u(t), v(t)]_{\mathfrak{g}} = -\text{ad}_{u(t)} v(t) \quad (13)$$

which is consistent with the symmetry of the Levi-Civita connection  $\nabla$ .

### The Geodesic Equations of a Right-Invariant Metric

Arnold developed a general framework in which to study geodesic equations of both left- as well as right-invariant metrics on arbitrary Lie groups as Euler equations on the associated Lie algebras. In everything that follows we will assume a right-handed algebra.

**Theorem 1** *If  $G$  is a Lie group with a right-invariant metric  $\langle \cdot, \cdot \rangle$ , then a curve  $\eta(t)$  is a geodesic if and only if the curve  $v(t)$  in  $\mathfrak{g}$ , defined by the flow equation*

$$\dot{\eta}(t) = dR_{\eta(t)} v(t) \quad (14)$$

*satisfies the Euler equations*

$$\dot{v}(t) = -\text{ad}_{v(t)}^* v(t). \quad (15)$$

*Proof.* To derive the Euler equations on  $\mathfrak{g}$  we will write the energy of a one-parameter family of curves  $(-\epsilon, \epsilon) \ni s \mapsto \eta(s, t)$  with fixed endpoints  $t = a$  and  $t = b$  in the form

$$E(s) = \frac{1}{2} \int_a^b \|\eta(s, t)\|^2 dt = \frac{1}{2} \int_a^b \langle dR_{\eta(s, t)^{-1}} \partial_t \eta(s, t), dR_{\eta(s, t)^{-1}} \partial_t \eta(s, t) \rangle dt$$

using right-invariance of the metric. To calculate  $\partial_s|_{s=0} (dR_{\eta(s, t)^{-1}} \partial_t \eta(s, t))$  set

$$\begin{aligned}z(t) &= \partial_s|_{s=0} v(s, t) = \partial_s|_{s=0} (dR_{\eta(s, t)^{-1}} \partial_t \eta(s, t)) \\ y(t) &= dR_{\eta(t)^{-1}} \partial_s \eta(s, t)|_{s=0}\end{aligned}$$

- the right translate of the variation vector field induced by  $\eta(s, t)$ . Then using (12)

$$\partial_t \partial_s|_{s=0} \eta(s, t) = \partial_t dR_{\eta(t)} y(t) = dR_{\eta(t)} (\partial_t y(t) + \nabla_{v(t)} y(t)).$$

On the other hand, using (12) again

$$\partial_s|_{s=0} \partial_t \eta(s, t) = \partial_s|_{s=0} dR_{\eta(s, t)} \cdot v(s, t) = dR_{\eta(t)} \cdot z(t) + dR_{\eta(t)} \cdot (\nabla_{y(t)} v(t)).$$

Since the left hand sides of each of the above equations are equal, putting the two together and rearranging using (13) we have

$$\partial_s|_{s=0} (dR_{\eta(s,t)^{-1}} \partial_t \eta(s,t)) = z(t) = \partial_t y(t) - \text{ad}_{v(t)} y(t).$$

Differentiating the energy  $E$  with respect to  $s$ , then integrating by parts and using the fact that  $y(a) = y(b) = 0$  we obtain

$$\partial_s|_{s=0} E(s) = \int_a^b \langle v, \partial_t y - \text{ad}_v y \rangle dt = - \int_a^b \langle \partial_t v + \text{ad}_v^* v, y \rangle dt$$

which is zero for any variation field  $y$  if and only if (15) holds.  $\square$   $\square$

**Corollary 1** *If  $\eta(t)$  is a curve in  $G$  with velocity field  $v(t) \in \mathfrak{g}$  defined through (14) and satisfying (15) with initial conditions  $\eta(0) = e$  and  $v(0) = v_o$  then we have the conservation law*

$$v(t) = \text{Ad}_{\eta^{-1}(t)}^* v_o. \quad (16)$$

*Proof.* From the definition of the group and algebra adjoint operators we have

$$\partial_t \text{Ad}_{\eta^{-1}(t)} = -\text{Ad}_{\eta^{-1}(t)} \text{ad}_{v(t)}$$

and by general properties of adjoints

$$\partial_t \text{Ad}_{\eta^{-1}(t)}^* = -\text{ad}_{v(t)}^* \text{Ad}_{\eta^{-1}(t)}^*.$$

Multiplying both sides of the Euler equations (15) by  $\text{Ad}_{\eta^{-1}(t)}^*$ , reversing the product rule, and integrating in time gives the result.  $\square$   $\square$

Introduce one more operator  $K_v : \mathfrak{g} \rightarrow \mathfrak{g}$

$$K_v(w) = \text{ad}_w^* v. \quad (17)$$

for later use.

**Corollary 2** *If  $\eta(t)$  is a curve in  $G$  with velocity field  $v(t) \in \mathfrak{g}$  defined through (14) and satisfying (15) with initial conditions  $\eta(0) = e$  and  $v(0) = v_o$  then*

$$K_{v(t)} = \text{Ad}_{\eta^{-1}(t)}^* K_{v_o} \text{Ad}_{\eta^{-1}(t)}. \quad (18)$$

*Proof.* For any  $u, w \in \mathfrak{g}$  we compute using the relationship between the algebra adjoint and the Lie bracket, and the latter's behaviour under pushforward by diffeomorphisms:

$$\begin{aligned} \langle \text{Ad}_{\eta(t)}^* K_{v(t)} u, w \rangle &= \langle v(t), \text{ad}_u \text{Ad}_{\eta(t)} w \rangle = \langle v(t), \text{Ad}_{\eta(t)} \text{ad}_{\text{Ad}_{\eta^{-1}(t)} u} w \rangle \\ &= \langle \text{Ad}_{\eta(t)}^* v(t), \text{ad}_{\text{Ad}_{\eta^{-1}(t)} u} w \rangle \\ &= \langle v_o, \text{ad}_{\text{Ad}_{\eta^{-1}(t)} u} w \rangle \\ &= \langle K_{v_o} \text{Ad}_{\eta^{-1}(t)} u, w \rangle \end{aligned}$$

where we have used Corollary 1 in the penultimate equality.  $\square$   $\square$

## The Exponential Map

We have not discussed the existence and uniqueness of solutions to the geodesic equations, nor their smooth dependence on initial conditions, but the standard analysis of geodesic systems on smooth finite dimensional Riemannian manifolds applies (see [26] or [22], for example) and it can be shown that geodesics of a right-invariant metric on  $G$  exist, are unique, and do depend smoothly on initial conditions. Given this, we now introduce the exponential of the right-invariant metric.

In parallel with the group exponential map  $\exp_e^G$ , the (Riemannian) exponential map of the right invariant metric  $\langle \cdot, \cdot \rangle$

$$\exp_e : \mathfrak{g} \rightarrow G$$

maps a vector  $v \in \mathfrak{g}$  to the value of the (Riemannian) geodesic  $\eta(t)$  with initial conditions  $\eta(0) = e$ ,  $\eta'(0) = v$  at time  $t = 1$

$$\exp_e(v) = \eta(1).$$

The group exponential map  $\exp_e^G$  is *never* identically equal to the Riemannian exponential map  $\exp_e$  unless  $\langle \cdot, \cdot \rangle$  is both left and right invariant, although it is a Theorem of Kaiser [19] that there always exists a one-parameter subgroup of  $G$  which is a solution to the right-invariant geodesic system (15) so that the images of the two exponential maps do have a non-empty intersection. A standard rescaling procedure show that the exponential map is the data-to-solution map of the geodesic equations (15) assigning to each  $(e, v) \in \{e\} \times \mathfrak{g}$  the unique geodesic  $\eta(t)$  with those initial conditions:

$$\exp_e(tv) = \eta(t).$$

Since solutions of the geodesic system depend smoothly on the initial conditions, the exponential map is smooth and whose derivative at  $t = 0$  can be shown to equal the identity operator - by the inverse function theorem  $\exp_e$  is a diffeomorphism between a neighbourhood of  $e \in G$  and a neighbourhood of the origin in  $\mathfrak{g}$ .

As a map from the Lie algebra  $\mathfrak{g}$  - the model space - to the Lie group  $G$  - the manifold - the exponential map  $\exp_e$  contains the geometry of  $G$  with  $\langle \cdot, \cdot \rangle$  and it is therefore of interest to study its properties and singularities. Alternatively, one could ask in what way the singularities of  $\exp_e$  could restrict the algebraic, topological, and differentiable structure of  $G$ . See [23], [12], [21], or [49] for early examples of the latter line of questioning; we'll proceed more in the direction of the former (with a view towards a BCH type formula).

So we turn to the Jacobi equation whose solutions give us precise information about the Riemannian exponential map  $\exp_e : \mathfrak{g} \rightarrow G$  and its derivative. Consider a geodesic  $\eta(t) = \exp_e(tv_o)$  starting from the identity  $e$  in the direction  $v_o \in \mathfrak{g}$ ; perturb the initial condition  $v_o$  by a vector  $z_o$ , and consider the one-parameter family of geodesics  $\eta(s, t) = \exp_e(t(v_o + sz_o))$ . The Jacobi field  $J(t)$  along  $\eta(t)$  satisfying  $J(0) = 0$  and  $\dot{J}(0) = z_o$  is the variation field of  $\eta(s, t)$

$$J(t) = \partial_s|_{s=0}\eta(s, t) = D \exp_e(tv_o)tz_o.$$

**Proposition 3** *Suppose  $G$  is any Lie group with a right-invariant metric  $\langle \cdot, \cdot \rangle$ . Let  $\eta(t)$  be a geodesic with  $\eta(0) = e$  and  $\dot{\eta}(0) = v_o$ . Then every Jacobi field  $J(t)$  along  $\eta(t)$ , with initial conditions  $J(0) = 0$  and  $\dot{J}(0) = z_o$ , satisfies the following system of equations on  $\mathfrak{g}$*

$$\partial_t y - \text{ad}_v y = z \quad (19)$$

$$\partial_t z + \text{ad}_v^* z + K_v z = 0 \quad (20)$$

where  $J(t) = dR_{\eta(t)}y(t)$  and  $\dot{\eta}(t) = dR_{\eta(t)}v(t)$  as in (15).

*Proof.* Equation (19) is the linearisation of the flow equation (14) and was derived in the proof of Theorem 1. Indeed, if  $\eta(s, t)$  is a variation of a geodesic  $\eta(t)$  which continues to be a geodesic for each  $s$ , with  $v(s, t)$  the corresponding solution to the Euler equations (15), then after setting

$$\begin{aligned} z(t) &= \partial_s|_{s=0}v(s, t) = \partial_s|_{s=0} (dR_{\eta(s, t)^{-1}}\partial_t\eta(s, t)) \\ y(t) &= dR_{\eta(t)^{-1}}\partial_s\eta(s, t)|_{s=0} \end{aligned}$$

we saw that

$$z(t) = \partial_s|_{s=0} (dR_{\eta(s, t)^{-1}}\partial_t\eta(s, t)) = \partial_t y - \text{ad}_v y.$$

Linearising the Euler equations (15) gives (20).  $\square$   $\square$

Equation (20) is the linearised Euler equation and we denote its solution operator by  $S(t) : \mathfrak{g} \rightarrow \mathfrak{g}$

$$S(t)z_o = z(t). \quad (21)$$

Denote the solution operator to the Jacobi system (20) and (19) by

$$\Phi(t)z_o := J(t) \quad (22)$$

which is a bounded linear operator from  $\mathfrak{g} = T_e G$  to  $T_{\eta(t)}G$ .

**Proposition 4** *Let  $G$ ,  $\eta(t)$ , and  $v(t)$  be as in Proposition 3. Then the solution operator (21) to the linearised Euler equation (20) is related to the solution operator (22) to the Jacobi system (19) and (20) by*

$$\text{Ad}_{\eta^{-1}(t)}^* = S(t) + K_{v(t)} \circ dR_{\eta^{-1}(t)} \circ \Phi(t) \quad (23)$$

*Proof.* If  $\eta(s, t)$  is a smooth variation of  $\eta(t)$  with corresponding variation of the velocity  $v(s, t)$  then  $z(t) = \partial_s|_{s=0}v(s, t)$ ,  $z_o = \partial_s|_{s=0}v(s, 0)$  and  $y(t) = dR_{\eta^{-1}(t)}\partial_s|_{s=0}\eta(s, t)$ . By Corollary 1

$$v(s, t) = \text{Ad}_{\eta^{-1}(s, t)}^* v_o(s)$$

for each  $s$  and differentiating both sides with respect to  $s$  and setting  $s = 0$  we obtain

$$z(t) = -\text{ad}_{y(t)}^* \text{Ad}_{\eta^{-1}(t)}^* v_o + \text{Ad}_{\eta^{-1}(t)}^* z_o.$$

Using Corollary 1 again and rewriting the expression in terms of the operator  $K_{v(t)}$  and the solution operators (21) and (22) we get

$$\text{Ad}_{\eta^{-1}(t)}^* z_o = S(t)z_o + K_{v(t)} \circ dR_{\eta^{-1}(t)} \circ \Phi(t)z_o$$

$\square$

$\square$

Proposition 4 together with Corollary 1 give another form of the Euler-Arnold equations. A simple of read of this new form for the equations is that the geodesic evolution is the time average of the linearised geodesic evolution.

**Theorem 2** *The Euler-Arnold equations of Theorem 1 can be re-written as a feedback system*

$$\begin{aligned}\partial_t \tilde{v}(t) &= -\text{ad}_{v(t)}^* \tilde{v}(t) - K_{v(t)} \tilde{v}(t) \\ v(t) &= \frac{1}{t} \int_0^t \tilde{v}(s) ds \\ v(0) &= \tilde{v}(0) = v_o\end{aligned}$$

*Proof.* If  $v(t)$  is a solution to the Euler-Arnold equations with initial condition  $v_o$  then setting  $z_o = v_o$  and applying Corollary 1 to the left hand side of equation (23) in Proposition 4, while making use of the fact that  $\Phi(t)v_o = t dR_{\eta(t)} \cdot v(t)$  for the right hand side of equation (23), we get

$$v(t) = S(t)v_o + t \cdot K_{v(t)}v(t).$$

The curve  $v(t)$  does *not* solve the linearised Euler equation (20) *unless*  $v(t) = v_o$  is stationary (in which case the above equation reads  $v_o = v_o$ ). So if  $v(t)$  is non-stationary this equation says  $v(t)$  decomposes as the solution to the linearised Euler equation (20) with the same initial condition as  $v(t)$  plus an extra term  $K_v v$ . Making use of the Euler-Arnold equations once more we can substitute  $-\partial_t v(t)$  for  $K_v v$  and obtain a new form of the geodesic equations:

$$t \cdot \partial_t v(t) = -v(t) + S(t)v_o = -v(t) + \tilde{v}(t).$$

which folds, using the product rule, as

$$\partial_t (t \cdot v(t)) = \tilde{v}(t).$$

Integrating both sides in  $t$ , we see that the geodesic equation of a right-invariant metric on a Lie group can be written as a feedback system

$$\begin{aligned}\partial_t \tilde{v}(t) &= -\text{ad}_{v(t)}^* \tilde{v}(t) - K_{v(t)} \tilde{v}(t) \\ v(t) &= \frac{1}{t} \int_0^t \tilde{v}(s) ds \\ v(0) &= \tilde{v}(0) = v_o\end{aligned}$$

where  $v(t)$  is the output - the geodesic dynamics - which, at every moment, is fed back into the first equation generating the dynamics.  $\square$   $\square$

## The Derivative of the Exponential Map

Using Propositions 3 and 4 we can derive a formula for the derivative of the exponential map associated to a right-invariant metric  $\langle \cdot, \cdot \rangle$  on a Lie group  $G$ .



**Theorem 3** *Suppose  $G$  is any Lie group with a right-invariant metric  $\langle \cdot, \cdot \rangle$ . Let  $\eta(t)$  be a geodesic with  $\eta(0) = e$  and  $\dot{\eta}(0) = v_o \in \mathfrak{g}$ , let  $J(t) = dR_{\eta(t)}y(t)$  be the Jacobi field along  $\eta(t)$  with initial conditions  $J(0) = 0$ ,  $\dot{J}(0) = z_o \in \mathfrak{g}$ , and let  $z(t)$  be the solution to (20) with initial condition  $z(0) = z_o \in \mathfrak{g}$ . Then*

1. *If  $z_o \in (\ker K_{v_o})^\perp$  then  $y(t) \in (\ker K_{v(t)})^\perp$  for all  $t$  and the derivative of the exponential map is given by*

$$D \exp_e(tv_o)tz_o = dR_{\eta(t)}\tilde{K}_{v(t)}^{-1} (\text{Ad}_{\eta^{-1}(t)}^* - S(t)) z_o \quad (24)$$

where  $\tilde{K}_{v(t)}$  denotes the restriction of  $K_{v(t)}$  to  $(\ker K_{v(t)})^\perp$  on which it is an isomorphism.

2. *If  $z_o \in \ker K_{v_o}$  then  $y(t) \in \ker K_{v(t)}$  for all  $t$  and is never zero for  $t > 0$ . The derivative of the exponential map is given by*

$$D \exp_e(tv_o)tz_o = dL_{\eta(t)} \int_0^t \text{Ad}_{\eta^{-1}(t)} \text{Ad}_{\eta^{-1}(s)}^* z_o ds \quad (25)$$

Observe that when  $\langle \cdot, \cdot \rangle$  is bi-invariant (assuming  $G$  allows for it) formula (24) reduces to the known formula for the derivative of the exponential map. Indeed, in this setting

$$\text{Ad}_{\eta(t)^{-1}}^* = \text{Ad}_{\eta(t)}, \quad \text{ad}_{v_o}^* = -\text{ad}_{v_o}, \quad K_{v_o} = \text{ad}_{v_o}$$

and together these imply that all solutions to the Euler-Arnold equations are stationary ( $v(t) = v_o$ ), the Riemannian exponential map and group exponential map coincide ( $\exp_e = \exp_e^G$ ) and

$$S(t) = I, \quad \forall t.$$

Substituting these relationships into (24) for appropriate  $z_o$  gives

$$\begin{aligned} D \exp_e^G(tv_o)tz_o &= dR_{\eta(t)} \text{ad}_{v_o}^{-1} (\text{Ad}_{\eta(t)} - I) z_o \\ &= dL_{\eta(t)} \text{ad}_{v_o}^{-1} (I - \text{Ad}_{\eta(t)^{-1}}) z_o \\ &= dL_{\exp_e^G(tv_o)} \text{ad}_{v_o}^{-1} (I - e^{-t \text{ad}_{v_o}}) z_o. \end{aligned}$$

Substituting the same relationships into (25) for appropriate  $z_o$  gives

$$D \exp_e^G(tv_o)tz_o = dL_{\exp_e^G tv_o} \cdot tz_o.$$

Comparing the Taylor series of the former with the latter we see that the two expressions coincide when  $z_o \in \ker \text{ad}_{v_o}$  and can therefore be reduced to a single formula written in terms of the functional calculus

$$D \exp_e(tv_o)tz_o = dL_{\exp_e^G(tv_o)} \cdot \frac{I - e^{-t \text{ad}_{v_o}}}{\text{ad}_{v_o}} z_o = dR_{\exp_e^G(tv_o)} \cdot \frac{e^{t \text{ad}_{v_o}} - I}{\text{ad}_{v_o}} z_o.$$

This is exactly the derivative of the (group) exponential map of a bi-invariant metric, see Duistermaat and Kolk [6] or Hall [14].

The proof of Theorem 2 will be broken down into a series of Lemmas which collectively show that  $\Phi(t) \ker K_{v_o} = \ker K_{v_o}$  and  $\Phi(t) (\ker K_{v_o})^\perp \subseteq (\ker K_{v_o})^\perp$  for each  $t$ . This allows us to invert  $K_{v(t)}$  in (23) and obtain (24). The bulk of the argument amounts to keeping track of how the operators  $\text{Ad}_{\eta(t)}$  and  $\text{Ad}_{\eta(t)}^*$  act on the subspace  $\ker K_{v_o}$ .

**Lemma 1** For each  $t$  we have

$$\text{Ad}_{\eta^{-1}(t)} \ker K_{v(t)} = \ker K_{v_o}, \quad \text{Ad}_{\eta^{-1}(t)} (\ker K_{v(t)})^\perp = (\ker K_{v_o})^\perp \quad (26)$$

$$\text{Ad}_{\eta^{-1}(t)}^* \ker K_{v_o} = \ker K_{v(t)}, \quad \text{Ad}_{\eta^{-1}(t)}^* (\ker K_{v_o})^\perp = (\ker K_{v(t)})^\perp \quad (27)$$

*Proof.* Corollary 2 gave a decomposition of the operator  $K_{v(t)}$  as

$$K_{v(t)} = \text{Ad}_{\eta^{-1}(t)}^* K_{v_o} \text{Ad}_{\eta^{-1}(t)}$$

from which it becomes clear that  $w \in \ker K_{v(t)}$  if and only if  $\text{Ad}_{\eta^{-1}(t)} w \in \ker K_{v_o}$ . Since  $\text{Ad}_{\eta^{-1}(t)}$  is an automorphism of  $\mathfrak{g}$  for each  $t$  we must have that  $w \in (\ker K_{v(t)})^\perp$  if and only if  $\text{Ad}_{\eta^{-1}(t)} w \in (\ker K_{v_o})^\perp$  as well; this gives (26).

Suppose  $w \in \ker K_{v(t)}$  so that  $\text{Ad}_{\eta(t)^{-1}} w \in \ker K_{v_o}$  by (26), and let  $u \in (\ker K_{v_o})^\perp$ . Then

$$0 = \langle u, \text{Ad}_{\eta(t)^{-1}} w \rangle = \langle \text{Ad}_{\eta(t)^{-1}}^* u, w \rangle.$$

Since this identity holds for all  $w \in \ker K_{v(t)}$  we must have  $\text{Ad}_{\eta(t)^{-1}}^* u \in (\ker K_{v(t)})^\perp$ ; that is,

$$\text{Ad}_{\eta(t)^{-1}}^* (\ker K_{v_o})^\perp \subseteq (\ker K_{v(t)})^\perp.$$

Similarly, if  $w \in \ker K_{v(t)}^\perp$  so that  $\text{Ad}_{\eta(t)^{-1}} w \in \ker K_{v_o}^\perp$  by (26), and  $u \in \ker K_{v_o}$  then

$$0 = \langle u, \text{Ad}_{\eta(t)^{-1}} w \rangle = \langle \text{Ad}_{\eta(t)^{-1}}^* u, w \rangle$$

so that  $\text{Ad}_{\eta(t)^{-1}}^* u$  is orthogonal to  $\ker K_{v(t)}^\perp$ ; that is,

$$\text{Ad}_{\eta(t)^{-1}}^* (\ker K_{v_o}) \subseteq (\ker K_{v(t)}).$$

Since  $\text{Ad}_{\eta(t)}^*$  is also an automorphism of  $\mathfrak{g}$  the inclusions must also be equalities - this gives (27).  $\square$   $\square$

**Lemma 2** The operator

$$\Omega(t) : \mathfrak{g} \rightarrow \mathfrak{g} \quad (28)$$

$$\Omega(t)w = \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* w ds$$

is an isomorphism with

$$\Omega(t) \ker K_{v_o} = \ker K_{v_o} \quad \Omega(t) (\ker K_{v_o})^\perp = (\ker K_{v_o})^\perp \quad (29)$$

for each  $t$ .

*Proof.* We have

$$\begin{aligned} \langle w, \Omega(t)w \rangle &= \int_0^t \langle w, \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* w \rangle ds = \int_0^t \|\text{Ad}_{\eta^{-1}(s)}^* w\|^2 ds \\ &\geq \left( \int_0^t \frac{1}{\|\text{Ad}_{\eta^{-1}(s)}^*\|^2} ds \right) \|w\|^2. \end{aligned}$$

Now

$$\|\text{Ad}_{\eta(s)}^*\|^2 = \|\text{Ad}_{\eta(s)}\|^2 \leq \left\| dL_{\eta(s)}^T dL_{\eta(s)} \right\|_{L^\infty}$$

where the  $L^\infty$  norm denotes the absolute value of the largest eigenvalue of the symmetric matrix  $dL_{\eta(s)}^T dL_{\eta(s)}$ . By the Schwartz inequality we have

$$C(t) \|w\| \leq \|\Omega(t)w\|,$$

$$C(t) = \int_0^t \frac{1}{\left\| dL_{\eta(s)}^T dL_{\eta(s)} \right\|_{L^\infty}} ds.$$

Since  $C(t)$  is the integral of a positive function it is itself positive and therefore  $\Omega(t)$  has empty kernel. Since  $\Omega(t)$  is self-adjoint it also has empty co-kernel and is therefore an isomorphism on  $\mathfrak{g}$ .

From (26) and (27) we have

$$\text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* \ker K_{v_o} = \ker K_{v_o}, \quad \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* (\ker K_{v_o})^\perp = (\ker K_{v_o})^\perp$$

for each  $s$ . If  $\{E_1^{\ker}, \dots, E_m^{\ker}\}$  span  $\ker K_{v_o}$  and  $\{E_{m+1}^\perp, \dots, E_n^\perp\}$  span  $(\ker K_{v_o})^\perp$  then for any  $w \in \ker K_{v_o}$  and any  $u \in (\ker K_{v_o})^\perp$  we can express  $\text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* w$  and  $\text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* u$  as

$$\text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* w = w_1(s) \cdot E_1^{\ker} + \dots + w_m(s) \cdot E_m^{\ker}$$

$$\text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* u = u_{m+1}(s) \cdot E_{m+1}^\perp + \dots + u_n(s) \cdot E_n^\perp.$$

Linearity of the integral then gives

$$\Omega(t)w = \left( \int_0^t w_1(s) ds \right) \cdot E_1^{\ker} + \dots + \left( \int_0^t w_m(s) ds \right) \cdot E_m^{\ker} \in \ker K_{v_o}$$

$$\Omega(t)u = \left( \int_0^t u_{m+1}(s) ds \right) \cdot E_{m+1}^\perp + \dots + \left( \int_0^t u_n(s) ds \right) \cdot E_n^\perp \in (\ker K_{v_o})^\perp$$

so that

$$\Omega(t) \ker K_{v_o} \subset \ker K_{v_o} \quad \Omega(t) (\ker K_{v_o})^\perp \subset (\ker K_{v_o})^\perp.$$

But since  $\Omega(t)$  is an isomorphism we must have

$$\Omega(t) \ker K_{v_o} = \ker K_{v_o} \quad \Omega(t) (\ker K_{v_o})^\perp = (\ker K_{v_o})^\perp.$$

□

□

**Lemma 3** *The operator*

$$\Gamma(t) : \mathfrak{g} \rightarrow \mathfrak{g} \tag{30}$$

$$w \mapsto \int_0^t \text{Ad}_{\eta^{-1}(s)} K_{v(s)} dR_{\eta^{-1}(s)} \Phi(s) w ds$$

*satisfies*

$$\Gamma(t)\mathfrak{g} \subseteq (\ker K_{v_o})^\perp \tag{31}$$

*for each  $t$ .*

*Proof.* The map  $\Phi(t)$  sends  $w$  to the Jacobi field  $J(t)$  along  $\eta(t)$  with initial conditions  $J(0) = 0$  and  $\dot{J}(0) = w$ . Let  $y(s) = dR_{\eta^{-1}(s)}\Phi(s)w$  be the right-translated Jacobi field which is a curve in  $\mathfrak{g}$ . For each  $s$  we have  $K_{v(s)}y(s) \in (\ker K_{v(s)})^\perp$ . To see this, suppose that  $y^\perp(s) \in (\ker K_{v(s)})^\perp$  and  $y^{\ker}(s) \in \ker K_{v(s)}$  then using Corollary 2 and (26)

$$K_{v(s)}y(s) = \text{Ad}_{\eta^{-1}(s)}^* K_{v_o} \text{Ad}_{\eta^{-1}(s)} y(s) = \text{Ad}_{\eta^{-1}(s)}^* K_{v_o} \text{Ad}_{\eta^{-1}(s)} y^\perp(s)$$

since  $K_{v_o} \text{Ad}_{\eta^{-1}(s)} y^{\ker}(s) = 0$ . What remains belongs to  $(\ker K_{v(s)})^\perp$  by (27). Using (26) of Lemma 1 once more

$$\text{Ad}_{\eta^{-1}(s)} K_{v(s)} y(s) \in (\ker K_{v_o})^\perp$$

for each  $s$ . Expanding  $\text{Ad}_{\eta^{-1}(s)} K_{v(s)} y(s)$  in a basis of eigenvectors of  $K_{v_o}$  spanning  $(\ker K_{v_o})^\perp$

$$\text{Ad}_{\eta^{-1}(s)} K_{v(s)} y(s) = w_1(s) \cdot E_1^\perp + \cdots + w_n(s) \cdot E_k^\perp.$$

Then

$$\Gamma(t)w = \left( \int_0^t w_1(s) ds \right) \cdot E_1^\perp + \cdots + \left( \int_0^t w_n(s) ds \right) \cdot E_k^\perp \in (\ker K_{v_o})^\perp.$$

□

□

**Lemma 4** *Let  $\eta(t)$  be a geodesic in  $G$  with velocity  $v(t)$  defined through (14) and let  $J(t) = dR_{\eta(t)}y(t)$  be the Jacobi field along  $\eta$  with initial conditions  $J(0) = 0$  and  $\dot{J}(0) = z_o$ . Then  $y(t) \in \mathfrak{g}$  is given by*

$$y(t) = \text{Ad}_{\eta(t)} [\Omega(t) - \Gamma(t)] z_o \quad (32)$$

*Proof.* Since

$$\frac{d}{dt} \text{Ad}_{\eta(t)} w = \text{ad}_v \text{Ad}_{\eta(t)} w$$

differentiating both sides of the identity

$$\text{Ad}_{\eta(t)} \text{Ad}_{\eta(t)^{-1}} w = w$$

shows that

$$\frac{d}{dt} \text{Ad}_{\eta(t)^{-1}} w = -\text{Ad}_{\eta(t)^{-1}} \text{ad}_v w.$$

Multiplying both sides of (19) by  $\text{Ad}_{\eta(t)^{-1}}$  and using the product rule we get

$$\frac{d}{dt} (\text{Ad}_{\eta(t)^{-1}} y(t)) = \text{Ad}_{\eta(t)^{-1}} z(t). \quad (33)$$

Re-arranging the identity in Proposition 4 for  $z(t)$  and substituting the result into (33) gives

$$\partial_t (\text{Ad}_{\eta^{-1}(t)} y(t)) = \text{Ad}_{\eta^{-1}(t)} \text{Ad}_{\eta^{-1}(t)}^* z_o - \text{Ad}_{\eta^{-1}(t)} K_{v(t)} y(t)$$

Integrating this equation from 0 to  $t$ , multiplying both sides by  $\text{Ad}_{\eta(t)}$  and rewriting the result in terms of  $\Omega(t)$  and  $\Gamma(t)$  gives (32). □ □

*Proof of Theorem 2.* Suppose  $z_o \in (\ker K_{v_o})^\perp$  and let  $J(t) = \Phi(t)z_o$  be the Jacobi field along  $\eta(t)$  with initial conditions  $J(0) = 0$  and  $\dot{J}(0) = z_o$ . From Lemma 4 the right-translated Jacobi field  $y(t) = dR_{\eta^{-1}(t)}J(t)$  is given by

$$y(t) = \text{Ad}_{\eta(t)} [\Omega(t) - \Gamma(t)] z_o. \quad (34)$$

By Lemmas 2 and 3

$$\Omega(t)z_o - \Gamma(t)z_o \in (\ker K_{v_o})^\perp$$

and by (26) of Lemma 1 we have

$$y(t) = \text{Ad}_{\eta(t)} [\Omega(t) - \Gamma(t)] z_o \in (\ker K_{v(t)})^\perp$$

for all  $t$ . By Proposition 4

$$\text{Ad}_{\eta^{-1}(t)}^* z_o = z(t) + K_{v(t)} y(t)$$

and since  $y(t) \in (\ker K_{v(t)})^\perp$  we can invert  $K_{v(t)}$  to obtain

$$y(t) = \tilde{K}_{v(t)}^{-1} (\text{Ad}_{\eta^{-1}(t)}^* z_o - z(t)) \quad (35)$$

where  $\tilde{K}_{v(t)}$  is the restriction of  $K_{v(t)}$  to  $(\ker K_{v(t)})^\perp$ . Right-translating by  $\eta(t)$  and rewriting the right-hand side of (35) in terms of the solution operator  $S(t)$  to the linearised Euler equation (20) we obtain

$$D \exp_e(tv_o)tz_o = J(t) = dR_{\eta(t)} \tilde{K}_{v(t)}^{-1} (\text{Ad}_{\eta^{-1}(t)}^* z_o - S(t)) z_o$$

which is part 1. of Theorem 2.

For part 2. we take a detour: Part 1. showed that given  $z_o \in (\ker K_{v_o})^\perp$  the curve  $y(t) = dR_{\eta^{-1}(t)}J(t) \in \mathfrak{g}$  defined through the Jacobi field  $J(t)$  along  $\eta(t)$  with initial conditions  $J(0) = 0$  and  $\dot{J}(0) = z_o$  belongs to  $(\ker K_{v(t)})^\perp$  for all  $t$  - we'll first show that  $z(t) = S(t)z_o \in (\ker K_{v(t)})^\perp$  for all  $t$  whenever  $z_o \in (\ker K_{v_o})^\perp$  and  $z(t) = S(t)z_o \in \ker K_{v(t)}$  for all  $t$  whenever  $z_o \in \ker K_{v_o}$ . Let  $z_o \in (\ker K_{v_o})^\perp$ ; as noted, the associated curve  $y(t) \in (\ker K_{v(t)})^\perp$  for all  $t$  and according to (26) of Lemma 1,  $\text{Ad}_{\eta^{-1}(t)} y(t) \in (\ker K_{v_o})^\perp$  for all  $t$ . Write this in the basis of eigenvectors provided by  $K_{v_o}$

$$\text{Ad}_{\eta^{-1}(t)} y(t) = y_1(t)E_1^\perp + \dots + y_k(t)E_k^\perp$$

and note that  $\partial_t (\text{Ad}_{\eta^{-1}(t)} y(t)) \in (\ker K_{v_o})^\perp$  aswell. Equation (33) relates  $\partial_t (\text{Ad}_{\eta^{-1}(t)} y(t))$  to the solution  $z(t)$  of the linearised Euler equations (20) with initial condition  $z(0) = z_o$  and shows that

$$\text{Ad}_{\eta^{-1}(t)} z(t) \in (\ker K_{v_o})^\perp.$$

Lemma 1, (26) then gives

$$z(t) \in (\ker K_{v(t)})^\perp \quad \forall t$$

and shows that

$$S(t) (\ker K_{v_o})^\perp \subset (\ker K_{v(t)})^\perp. \quad (36)$$

Since  $S(t)$  is the solution operator to a first order linear differential equation,  $S(t^*)w = 0$  at some time  $t^*$  if and only if  $w = 0$  - for otherwise this contradicts uniqueness of solutions. Therefore  $S(t)$  has empty kernel for as long as it is defined and consequently is an automorphism of  $\mathfrak{g}$  for as long as it is defined - this, in conjunction with the fact that  $\dim(\ker K_{v_o})^\perp = \dim(\ker K_{v(t)})^\perp$  by Lemma 1, means the inclusion (36) is an *equality* and

$$S(t)(\ker K_{v_o})^\perp = (\ker K_{v(t)})^\perp. \quad (37)$$

Now isomorphism of  $S(t)$  also implies

$$S(t)\ker K_{v_o} = \ker K_{v(t)}. \quad (38)$$

Now let  $z_o \in \ker K_{v_o}$ ,  $z(t) = S(t)z_o$ , and  $y(t) = dR_{\eta^{-1}(t)}\Phi(t)z_o$ . Then by Proposition 2, Lemma 1 (27), and (38)

$$K_{v(t)}y(t) = \text{Ad}_{\eta^{-1}}^* z_o - z(t) \in \ker K_{v(t)}$$

for each  $t$ . But this implies  $K_{v(t)}y(t) = 0$  for each  $t$ : indeed, for a fixed  $t$  let  $u \in (\ker K_{v(t)})^\perp$ ; then, since  $K_{v(t)}$  is a skew self-adjoint operator on  $\mathfrak{g}$

$$0 = \langle K_{v(t)}y(t), u \rangle = -\langle y(t), K_{v(t)}u \rangle$$

and since  $K_{v(t)}$  is also an isomorphism on  $(\ker K_{v(t)})^\perp$  and the identity holds for any  $u \in (\ker K_{v(t)})^\perp$  we must have that  $y(t) \in \ker K_{v(t)}$ . Since this holds for any  $t$

$$y(t) \in \ker K_{v(t)} \quad \forall t$$

and

$$\text{Ad}_{\eta^{-1}}^* z_o = z(t) \quad \forall t. \quad (39)$$

Finally, to prove (65) use (39) and (33) to write

$$\partial_t (\text{Ad}_{\eta^{-1}(t)} y(t)) = \text{Ad}_{\eta^{-1}(t)} \text{Ad}_{\eta^{-1}(t)}^* z_o.$$

Integrating this equation from 0 to  $t$  and multiplying both sides by  $dL_{\eta(t)}$  we obtain

$$D \exp_e(tv_o)tz_o = dR_{\eta(t)}y(t) = dL_{\eta(t)} \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* z_o ds = dL_{\eta(t)} \Omega(t)z_o$$

which is never zero for all  $t$  since  $\Omega(t)$  is an automorphism by Lemma 2 and  $dL_{\eta(t)}$  is an isomorphism.  $\square$

## Conjugate Points

An immediate application of Theorem 2 is to conjugate points of a right-invariant metric on  $G$  - these are singularities of  $\exp_e$ . Let  $\eta(t)$  be a geodesic of the right invariant metric  $\langle \cdot, \cdot \rangle$  in  $G$  issuing from the identity  $e$  in the direction  $v_o$ . A point  $\eta(t^* > 0)$  is said to be conjugate to  $e$  if the linear operator

$$D \exp_e(t^* v_o)t^* : \mathfrak{g} \rightarrow T_{\eta(t^*)}G$$

fails to be an isomorphism. If  $\dim \ker D \exp_e(t^* v_o)t^* = k$  then  $k$  is called the multiplicity of the conjugate point  $\eta(t^*)$ . After the first conjugate point the infinitesimal geometry along  $\eta(t)$  ceases to resemble the infinitesimal geometry along the straight line segment  $tv_o$  in the model space. They contain information about the topological and differentiable structure of  $G$  and relate to covering properties of the exponential map. In the infinite dimensional situation conjugate points contain information on the stability of physical systems described by geodesics; for example, Hoffman and Manning showed how conjugate points are related to the buckling of elastic planar rods, while Sokolowski [44] studied conjugate points and their relationship to the twin paradox in general relativity (although the metrics are not Riemannian here).

An immediate consequence of Theorem 2 is the following

**Theorem 4** *Let  $G$ ,  $\langle \cdot, \cdot \rangle$ ,  $\eta(t)$ ,  $v(t)$ ,  $J(t)$ , and  $z(t)$  be as in Theorem 2. Then  $\eta(t^*)$  is conjugate to  $e$  if and only if there exists a  $z_o \in (\ker K_{v_o})^\perp$  and orthogonal to  $v_o$  such that*

$$\text{Ad}_{\eta^{-1}(t^*)}^* z_o = S(t^*) z_o \quad (40)$$

*Proof.* If  $\eta(t^*)$  is conjugate to the identity  $e$  then  $D \exp_e(t^* v_o)t^*$  fails to be an isomorphism between  $\mathfrak{g}$  and  $T_{\eta(t^*)}G$ . Since  $D \exp_e(t^* v_o)t^*|_{\ker K_{v_o}}$  is never zero by part 2. of Theorem 2, the injectivity failure must occur on  $(\ker K_{v_o})^\perp$ . Formula (24) shows that  $D \exp_e(t^* v_o)t^*$  fails to be injective if and only if there exists a  $z_o \in (\ker K_{v_o})^\perp$  such that (40) holds.

Conversely if there exists a  $z_o \in (\ker K_{v_o})^\perp$  such that (40) holds then  $D \exp_e(t^* v_o)t^*$  fails to be an isomorphism by Theorem 2 and therefore  $\eta(t^*)$  is conjugate to the identity  $e$ . □

By Corollary 1, a solution to the geodesic equations (15) evolves on the coadjoint orbit  $\{\text{Ad}_g^* v_o : g \in G\}$  of its initial condition. The solution operator  $S(t)$  to the Linearised Euler equations (20) therefore measures the rate at which coadjoint orbits separate. Theorem 3 says a conjugate point along  $\eta(t)$  occurs whenever a solution to the linearised Euler equations passes through the geodesic orbit  $\text{Ad}_{\eta^{-1}(t^*)}^* z_o$  of its initial condition. It could be interesting to understand what this actually means.

## Stability

Another application of Proposition 4 and Theorem 2 is to the stability of the Euler-Arnold equations and the stability of geodesics. In continuum mechanics it is common to study the stability of stationary or equilibrium solutions to a

set of equations. In our setting, a vector  $v_o \in \mathfrak{g}$  is a stationary or equilibrium point of the Euler-Arnold equations (15) if and only if  $\text{ad}_{v_o}^* v_o = 0$ . The linearised Euler-Arnold equations (20) describe the effect small perturbations of  $v_o$  have on the corresponding solutions to the Euler-Arnold equations and analysing the growth of the norm of solutions to (20) is called Eulerian stability. The geodesic  $\eta(t) = \exp_e(tv_o)$  generated by the stationary solution  $v_o$  is called a stationary geodesic. The Jacobi equation (??) describes the effect small perturbations of  $v_o$  have on the corresponding geodesics and analysing the growth of the norm of Jacobi fields is called Lagrangian stability.

Since the Jacobi equation (??) decouples as the linearised Euler equation and the linearised flow equation it should come as no surprise that Eulerian stability is related to Lagrangian stability. Proposition 4 tells us that this relationship is determined by the coadjoint operator and  $K_{v_o}$ . In some simple but general situations we can determine the precise relationship between Eulerian and Lagrangian stability and the conclusion is that it is enough to study the first order stability problem (Eulerian stability) to understand the second order stability problem (Lagrangian stability).

In both of the following theorems, and their proofs, we will assume:  $G$  is a Lie group with Lie algebra  $\mathfrak{g}$  and right-invariant metric  $\langle \cdot, \cdot \rangle$  with induced norm  $\|\cdot\|_{\mathcal{R}}$ ;  $v_o$  is a stationary solution to the Euler-Arnold equations (15) with corresponding geodesic  $\eta(t)$ ;  $z(t)$  is a solution to the linearised (about  $v_o$ ) Euler-Arnold equations (20) with initial condition  $z_o$ ; and  $J(t)$  is a solution to the Jacobi equation along  $\eta(t)$  with initial conditions  $J(0) = 0$ ,  $\frac{d}{dt}J(0) = z_o$ .

**Theorem 5** *Suppose  $G$  is compact. Then:*

1. *If  $z_o \in \ker K_{v_o}$  then the norm of  $z(t)$  is uniformly bounded in  $t$  while the corresponding Jacobi field grows at most linearly in  $t$*

$$\|z(t)\|_{\mathcal{R}} \leq C_1 \|z_o\|_{\mathcal{R}}, \quad \|J(t)\|_{\mathcal{R}} \leq C_2 t \|z_o\|_{\mathcal{R}}$$

*for finite constants  $C_1$  and  $C_2$ ;*

2. *If  $z_o \in (\ker K_{v_o})^\perp$  then Eulerian and Lagrangian stabilities and instabilities are equivalent*

$$C_1 \|J(t)\|_{\mathcal{R}} \leq C_2 \cdot \|z_o\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} \leq 2C_2 \cdot \|z_o\|_{\mathcal{R}} + C_3 \|J(t)\|_{\mathcal{R}}.$$

*for finite constants  $C_1$ ,  $C_2$ , and  $C_3$ ; and*

3. *In general, Eulerian and Lagrangian stabilities and instabilities are equivalent modulo a linear factor of  $t$ :*

$$\begin{aligned} C_1 \|J(t)\|_{\mathcal{R}} &\leq C_2 \cdot \|z_o\|_{\mathcal{R}} + t \cdot C_3 \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} \\ &\leq 2C_2 \cdot \|z_o\|_{\mathcal{R}} + t \cdot C_3 \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}} + C_4 \|J(t)\|_{\mathcal{R}} \end{aligned}$$

*for finite constants  $C_1$ ,  $C_2$ ,  $C_3$ , and  $C_4$ .*

*Proof.* Suppose  $G$  is compact. Then, as is well known,  $G$  admits a bi-invariant metric  $\langle \cdot, \cdot \rangle_{\mathcal{B}}$  for which  $\text{Ad}_g$  is unitary for any  $g \in G$ ; in particular, its spectrum lies on the unit circle and its operator norm is 1:

$$\|\text{Ad}_g\|_{\text{Op}(\mathcal{B})} = 1$$



for any  $g \in G$ . Since  $\text{Ad}_g$  is an automorphism of  $\mathfrak{g}$  and the automorphism group  $\text{Aut}(\mathfrak{g})$  is finite dimensional all norms are equivalent:

$$c_1 = c_1 \|\text{Ad}_g\|_{\text{Op}(\mathcal{B})} \leq \|\text{Ad}_g\|_{\text{Op}(\mathcal{R})} \leq c_2 \|\text{Ad}_g\|_{\text{Op}(\mathcal{B})} = c_2 \quad (41)$$

for positive constants  $c_1$  and  $c_2$ . From the general properties of operator norms the equality  $\|\text{Ad}_g\|_{\text{Op}(\mathcal{R})} = \|\text{Ad}_g^*\|_{\text{Op}(\mathcal{R})}$  holds and we obtain

$$c_1 \leq \|\text{Ad}_g^*\|_{\text{Op}(\mathcal{R})} \leq c_2 \quad (42)$$

for any  $g \in G$ .

For part 1. of the Theorem, assume that  $z_o \in \ker K_{v_o}$ ; then  $y(t) = dR_{\eta(t)^{-1}} \cdot J(t) \in \ker K_{v_o}$  for all  $t$  by part 2 of Theorem 2, and by Proposition 4 and (42) we have

$$\|z(t)\|_{\mathcal{R}} \leq \|\text{Ad}_{\eta^{-1}(t)}^* z_o\|_{\mathcal{R}} \leq c_2 \|z_o\|_{\mathcal{R}}.$$

The formula in part 2 of Theorem 2, along with right-invariance, (41) and (42) gives

$$\begin{aligned} \|J(t)\|_{\mathcal{R}} &= \left\| \text{Ad}_{\eta(t)} \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* z_o ds \right\|_{\mathcal{R}} \\ &\leq \|\text{Ad}_{\eta(t)}\|_{\text{Op}(\mathcal{R})} \cdot \int_0^t \|\text{Ad}_{\eta^{-1}(s)}\|_{\text{Op}(\mathcal{R})} \cdot \|\text{Ad}_{\eta^{-1}(s)}^*\|_{\text{Op}(\mathcal{R})} \cdot \|z_o\|_{\mathcal{R}} ds \\ &\leq tc_2^3 \cdot \|z_o\|_{\mathcal{R}}. \end{aligned}$$

For part (3) suppose  $\ker K_{v_o} \subsetneq \mathfrak{g}$  and let  $z(t)$  be the solution to the linearised Euler equations (20) with initial condition  $z_o$  and let  $J(t) = dR_{\eta(t)} y(t)$  be the Jacobi field with initial conditions  $J(0) = 0$ ,  $\frac{d}{dt} J(0) = z_o$ . Let  $z_o^{\text{ker}}$  denote the part of  $z_o$  which lies in  $\ker K_{v_o}$  and  $J^{\text{ker}}$  the corresponding part of  $J(t)$  which lies in  $\ker K_{v_o}$ ; similarly, let  $z_o^\perp$  denote the part of  $z_o$  which lies in  $(\ker K_{v_o})^\perp$  and  $J^\perp$  the corresponding part of  $J(t)$  which lies in  $(\ker K_{v_o})^\perp$ . Using Proposition 4, skew self-adjointness of  $K_{v_o}$ , and the bounds (42) we can estimate

$$\begin{aligned} \|z(t)\|_{\mathcal{R}} &\leq \|\text{Ad}_{\eta^{-1}(t)}^* z_o\|_{\mathcal{R}} + \|K_{v_o} y(t)\|_{\mathcal{R}} \\ &\leq \|\text{Ad}_g^*\|_{\text{Op}(\mathcal{R})} \|z_o\|_{\mathcal{R}} + \max_{\lambda \in \sigma(K_{v_o})} |\lambda| \|J(t)\|_{\mathcal{R}} \\ &\leq c_2 \|z_o\|_{\mathcal{R}} + \max_{\lambda \in \sigma(K_{v_o})} |\lambda| \|J(t)\|_{\mathcal{R}} \end{aligned} \quad (43)$$

Now estimate in the other direction: Proposition 4, right-invariance and (42) immediately give

$$\min_{\lambda \in \sigma(K_{v_o})/\{0\}} |\lambda| \left\| J^\perp(t) \right\|_{\mathcal{R}} \leq \|K_{v_o} y(t)\|_{\mathcal{R}} \leq c_2 \cdot \|z_o\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}}$$

Adding  $\min_{\lambda \in \sigma(K_{v_o})/\{0\}} |\lambda| \left\| J^{\text{ker}}(t) \right\|_{\mathcal{R}}$  to both sides of the inequality gives

$$\min_{\lambda \in \sigma(K_{v_o})/\{0\}} |\lambda| \|J(t)\|_{\mathcal{R}} \leq c_2 \cdot \|z_o\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} + \min_{\lambda \in \sigma(K_{v_o})/\{0\}} |\lambda| \left\| J^{\text{ker}}(t) \right\|_{\mathcal{R}}$$

Using part (2) of Theorem 2, right-invariance of the metric, and the bounds (41) and (42) we can bound the last term of the inequality:

$$\begin{aligned} \left\| J^{\ker}(t) \right\|_{\mathcal{R}} &= \left\| \text{Ad}_{\eta(t)} \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* z_o^{\ker} ds \right\|_{\mathcal{R}} \\ &\leq \left\| \text{Ad}_{\eta(t)} \right\|_{\text{Op}(\mathcal{R})} \cdot \int_0^t \left\| \text{Ad}_{\eta^{-1}(t)} \right\|_{\text{Op}(\mathcal{R})} \cdot \left\| \text{Ad}_{\eta^{-1}(t)}^* \right\|_{\text{Op}(\mathcal{R})} \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}} ds \\ &\leq t c_2^3 \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}} \end{aligned} \quad (44)$$

and arrive at

$$\min_{\lambda \in \sigma(K_{v_o}/\{0\})} |\lambda| \|J(t)\|_{\mathcal{R}} \leq c_2 \cdot \|z_o\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} + t \cdot c_2^3 \cdot \min_{\lambda \in \sigma(K_{v_o}/\{0\})} |\lambda| \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}}. \quad (45)$$

Putting the inequalities (43) and (45) together we get

$$\begin{aligned} \min_{\lambda \in \sigma(K_{v_o}/\{0\})} |\lambda| \|J(t)\|_{\mathcal{R}} &\leq c_2 \cdot \|z_o\|_{\mathcal{R}} + t \cdot c_2^3 \cdot \min_{\lambda \in \sigma(K_{v_o}/\{0\})} |\lambda| \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} \\ &\leq 2c_2 \cdot \|z_o\|_{\mathcal{R}} + t \cdot c_2^3 \cdot \min_{\lambda \in \sigma(K_{v_o}/\{0\})} |\lambda| \cdot \left\| z_o^{\ker} \right\|_{\mathcal{R}} \\ &\quad + \max_{\lambda \in \sigma(K_{v_o})} |\lambda| \|J(t)\|_{\mathcal{R}}, \end{aligned}$$

which completes part (3).

If  $z_o \in (\ker K_{v_o})^\perp$  then  $z_o^{\ker} = 0$  and  $y(t) = dR_{\eta(t)^{-1}} \cdot J(t) \in (\ker K_{v_o})^\perp$  for all  $t$  by part 1 of Theorem 2 so that part (2) follows from part (3).  $\square$   $\square$

**Theorem 6** *Suppose  $G$  is a nilpotent Lie group with nilpotent Lie algebra  $\mathfrak{g}$  of step  $k$ . Then*

1. *If  $z_o \in \ker K_{v_o}$  then the norm of  $z(t)$  grows at most polynomially of order  $k-1$  while the norm of the corresponding Jacobi field grows at most polynomially of order  $3k-2$*

$$\begin{aligned} \|z(t)\|_{\mathcal{R}} &\leq \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_{v_o}^*)^{k-i} z_o \right\|_{\mathcal{R}} \\ \|J(t)\|_{\mathcal{R}} &\leq \sum_{i,j,l=1}^k t^{3k+1-i-j-l} \left\| (\text{ad}_{v_o})^{2k-i-j} (\text{ad}_{v_o}^*)^{k-l} z_o \right\|_{\mathcal{R}} \end{aligned}$$

2. *If  $z_o \in (\ker K_{v_o})^\perp$  then Eulerian and Lagrangian stabilities and instabilities are equivalent modulo a polynomial of order  $k-1$ :*

$$C_1 \|J(t)\|_{\mathcal{R}} \leq \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_v^*)^{k-i} z_o \right\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} \leq 2 \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_v^*)^{k-i} z_o \right\|_{\mathcal{R}} + C_2 \cdot \|J(t)\|_{\mathcal{R}}$$

for finite constants  $C_1$  and  $C_2$ ; and

3. In general, Eulerian and Lagrangian stabilities and instabilities are equivalent modulo a polynomial of order  $3k - 2$ :

$$\begin{aligned}
C_1 \|J(t)\|_{\mathcal{R}} &\leq \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_v^*)^{k-i} z_o \right\|_{\mathcal{R}} + \|z(t)\|_{\mathcal{R}} \\
&\quad + C_1 \cdot \sum_{i=1}^k \sum_{j=1}^k \sum_{l=1}^k t^{3k+1-i-j-l} \left\| (\text{ad}_{v_o})^{2k-i-j} (\text{ad}_{v_o}^*)^{k-l} z_o^{\text{ker}} \right\|_{\mathcal{R}} \\
&\leq 2 \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_v^*)^{k-i} z_o \right\|_{\mathcal{R}} + C_2 \cdot \|J(t)\|_{\mathcal{R}} \\
&\quad + C_1 \cdot \sum_{i=1}^k \sum_{j=1}^k \sum_{l=1}^k t^{3k+1-i-j-l} \left\| (\text{ad}_{v_o})^{2k-i-j} (\text{ad}_{v_o}^*)^{k-l} z_o^{\text{ker}} \right\|_{\mathcal{R}}.
\end{aligned}$$

for finite constants  $C_1$  and  $C_2$ .

*Proof.* If  $G$  is nilpotent of step  $k$  then  $(\text{ad}_v)^k z_o = 0$ , where “to the power of  $k$ ” means we have applied the adjoint operator  $k$  times. This also means  $(\text{ad}_v^*)^k z_o = 0$ . Using the definition of the algebra adjoint and coadjoint operators we find that

$$\frac{d^k}{dt^k} \text{Ad}_{\eta(t)} z_o = 0, \quad \frac{d^k}{dt^k} \text{Ad}_{\eta^{-1}(t)} z_o = 0, \quad \frac{d^k}{dt^k} \text{Ad}_{\eta^{-1}(t)}^* z_o = 0$$

which give

$$\begin{aligned}
\text{Ad}_{\eta(t)} z_o &= t^{k-1} \cdot (\text{ad}_{v_o})^{k-1} z_o + \cdots + t \cdot \text{ad}_{v_o} z_o + z_o \\
\text{Ad}_{\eta^{-1}(t)} z_o &= (-1)^{k-1} t^{k-1} \cdot (\text{ad}_{v_o})^{k-1} z_o + \cdots + (-1)t \cdot \text{ad}_{v_o} z_o + z_o \\
\text{Ad}_{\eta^{-1}(t)}^* z_o &= (-1)^{k-1} t^{k-1} \cdot (\text{ad}_{v_o}^*)^{k-1} z_o + \cdots + (-1)t \cdot \text{ad}_{v_o}^* z_o + z_o
\end{aligned}$$

and consequently

$$\begin{aligned}
\left\| \text{Ad}_{\eta(t)} z_o \right\|_{\mathcal{R}} &\leq \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_v)^{k-i} z_o \right\|_{\mathcal{R}} \\
\left\| \text{Ad}_{\eta^{-1}(t)} z_o \right\|_{\mathcal{R}} &\leq \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_{v_o})^{k-i} z_o \right\|_{\mathcal{R}} \\
\left\| \text{Ad}_{\eta^{-1}(t)}^* z_o \right\|_{\mathcal{R}} &\leq \sum_{i=1}^k t^{k-i} \left\| (\text{ad}_{v_o}^*)^{k-i} z_o \right\|_{\mathcal{R}} \\
\left\| \text{Ad}_{\eta(t)} \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* z_o ds \right\|_{\mathcal{R}} &\leq \sum_{i=1}^k \sum_{j=1}^k \sum_{l=1}^k t^{3k+1-i-j-l} \left\| (\text{ad}_{v_o})^{2k-i-j} (\text{ad}_{v_o}^*)^{k-l} z_o \right\|_{\mathcal{R}}.
\end{aligned} \tag{46}$$

Using these formulas the proof follows that of Theorem 5, exactly.  $\square$   $\square$

## Diffeomorphism Groups

The question is: “what additional properties of the adjoint and coadjoint operators are required for Theorem 3 and its proof to lift to diffeomorphism groups?” There are two characteristics of the operators that track throughout the proof of Lemmas 1 - 3, and their amalgamation in the final proof of Theorem 3; these are

1. Boundedness and invertibility of the group adjoint and coadjoint operators, aswell as their product and the integral of their product on  $\mathfrak{g}$ ;
2. the admission of a complete orthonormal basis of eigenvectors by  $K_v$  which span the Lie algebra  $\mathfrak{g}$ .

The operator  $K_v$ , regardless of the dimension of  $G$ , is always a skew self-adjoint operator. In an infinite dimensional Hilbert space  $H$  the Spectral Theorem tells us that compact (skew) self-adjoint operators admit an orthonormal basis of eigenvectors spanning  $H$  and whose eigenvalues accumulate on zero only. So we can replace (1) and (2) by

- (A) Boundedness and invertibility of the group adjoint and coadjoint operators, aswell as their product and the integral of their product on  $\mathfrak{g}$ ;
- (B)  $K_v$  is a compact operator on  $\mathfrak{g}$ .

If one could find a suitable functional analytic topology for the Diffeomorphism group so that it becomes a smooth manifold in which left and right translation are at least continuous, and the above two characteristics hold then one immediately obtains Theorem 3 with the same proof, aswell as the corollaries on conjugate points, and Eulerian/Lagrangian stability.

The purpose of this section is to show that such a topology in which (A) and (B) hold exists and has been comprehensively analyzed by Misiolek and Preston in their study of Fredholm properties of exponential maps on diffeomorphism groups [34]. Properties (A) and (B) are precisely what makes an exponential map a non-linear Fredholm map of index zero, and these properties are not satisfied in just any topology or for any metric defining the exponential map - there are conditions under which these holds. That Fredholm exponential maps do admit a formula like that in Theorem 3 is in keeping with topological nature of the problem and of a possible BCH formula in this setting.

We will not give an account of Fredholm maps here and instead refer the reader to [34] for details and applications. We will also not recreate the analysis and results given in [34] but supply the reader with the parts of the theory that are relevant for Theorem 3 in this setting.

## Background and Preparation

The main reference for this section is [33] and in the sections that follow we will quote their fundamental results which are relevant for lifting Theorem 3 to this setting. Additional references on Sobolev topologies and the manifold structure for the diffeomorphism groups are Ebin and Marsden [7], and the monographs of Hebey [15] and Inci, Kappeler, and Topalov [17].

For a closed  $n$ -dimensional smooth manifold  $M$ , the class  $H^s(M, M)$  is the set of maps  $f : M \rightarrow M$  whose first  $s$  derivatives are square-integrable in every coordinate chart of  $M$ . It is a classical theorem of Sobolev that if  $s$  has sufficiently many

square integrable derivatives then it is also classically differentiable: if  $s > \frac{n}{2} + 1$  then every map  $f$  is also  $C^1$  and we can define the set of Sobolev diffeomorphisms of  $M$  by

$$\mathcal{D}^s(M) = \{\eta \in H^s(M, M) : \eta^{-1} \text{ exists and } \eta^{-1} \in H^s(M, M)\}.$$

The set  $\mathcal{D}^s(M)$  is open, inherits a manifold structure from  $H^s(M, M)$ , and is a topological group under composition whose tangent space at the identity (its Lie algebra) is the set of Sobolev  $H^s$  vector fields on  $M$ .

The operation of right translation is smooth but left translation is not even continuous

$$\begin{aligned} R_\eta : G &\rightarrow G & L_\eta : G &\rightarrow G \\ R_\eta \xi &= \xi \circ \eta & L_\eta \xi &= \eta \circ \xi. \end{aligned}$$

To see this let  $\xi(t)$  be a curve in  $\mathcal{D}^s(M)$  with  $\xi(0) = e$  and  $\xi'(0) = v \in T_e \mathcal{D}^s(M)$ ; the differential of right translation is again right translation while the chain rule shows that the differential of left translation is multiplication by the Jacobian matrix of  $\eta$ :

$$\begin{aligned} dR_\eta \cdot v &= \frac{d}{dt} \Big|_{t=0} R_\eta \xi(t) = \frac{d}{dt} \Big|_{t=0} \xi(t) \circ \eta = v \circ \eta \\ dL_\eta \cdot v &= \frac{d}{dt} \Big|_{t=0} L_\eta \xi(t) = \frac{d}{dt} \Big|_{t=0} \eta \circ \xi(t) = D\eta \cdot v. \end{aligned}$$

As a consequence, any operation involving the derivative of left translation doesn't make sense on  $T_e \mathcal{D}^s(M)$  and is only defined on a space of lower order regularity -  $T_e \mathcal{D}^r(M)$  for  $r < s$ . In the rest of this section we will make use of two Sobolev indices:  $s$  for the strong Sobolev topology which defines the manifold topology on  $\mathcal{D}^s(M)$ , and  $r$  for the weaker Sobolev topology on which  $dL_\eta$  is a *bounded* linear operator. In actuality, the weaker  $H^r$  topology will be the completion of the space  $T_e \mathcal{D}^s(M)$  in an  $H^r$  norm introduced in the next section.

Just as in the finite dimensional case we can define the group adjoint operator which can be written explicitly using the latter calculations as

$$\begin{aligned} \text{Ad}_\eta : T_e \mathcal{D}^r(M) &\rightarrow T_e \mathcal{D}^r(M) \\ \text{Ad}_\eta v &= dL_\eta dR_{\eta^{-1}} \cdot v = D\eta \cdot v \circ \eta^{-1} = \eta_* v. \end{aligned} \tag{47}$$

for any  $\eta \in \mathcal{D}^s(M)$  and  $r < s$ . That is, the group adjoint operator on the diffeomorphism group is the usual action of diffeomorphisms on vector fields via push-forward. As mentioned above, the reason that  $\text{Ad}_\eta$  only makes sense on  $T_e \mathcal{D}^r(M)$  with  $r < s$  is that if  $\eta$  is of class  $H^s$  then the calculation for the derivative of left translation above shows that  $dL_\eta$  is only defined on  $T_e \mathcal{D}^r(M)$  for  $r < s$ .

The algebra coadjoint operator can also be defined as in finite dimensions by

$$\begin{aligned} \text{ad}_v : T_e \mathcal{D}^r(M) &\rightarrow T_e \mathcal{D}^r(M), \quad r < s \\ \text{ad}_v w &= \frac{d}{dt} \Big|_{t=0} \text{Ad}_{\eta(t)} w \end{aligned} \tag{48}$$

where  $\eta(t)$  is a curve in  $\mathcal{D}^s(M)$  with  $\eta(0) = e$ ,  $\eta'(0) = v \in T_e\mathcal{D}^s(M)$ . To calculate the algebra adjoint operator explicitly let  $\eta(t)$  be the curve in  $\mathcal{D}^s(M)$  corresponding to the flow of  $v$  on  $M$  (i.e.  $\eta(0) = e$ ,  $\eta'(0) = v \in T_e\mathcal{D}^s(M)$ ). Then using the definition of the Lie derivative and the explicit form of the group adjoint operator

$$[v, w] = \frac{d}{dt}\Big|_{t=0}\eta(t)_*^{-1}w = \frac{d}{dt}\Big|_{t=0}\text{Ad}_{\eta(t)^{-1}}w = -\text{ad}_v w$$

where the last equality follows from differentiating the expression  $\text{Ad}_{\eta(t)^{-1}}\text{Ad}_{\eta(t)} = I$ . In particular, the minus sign is a consequence of the fact that the Lie derivative is associated with  $\text{Ad}_{\eta(t)^{-1}}$ . So in some sense the group structure on  $\mathcal{D}^s(M)$  forces the sign in the algebra adjoint operator through the explicit form of the group adjoint operator as pushforward - compare with Proposition 1 in the finite dimensional setting.

As for the case of the group adjoint operator, if  $u \in T_e\mathcal{D}^s(M)$  then  $\text{ad}_u$  is only defined on  $T_e\mathcal{D}^r(M)$  for  $r < s$  due to  $u$ 's derivative loss within  $\text{ad}$ .

If we assume that  $M$  is a Riemannian manifold with metric  $g$  and Riemannian volume form  $\mu$  then the Sobolev  $H^s$  volumorphism group  $\mathcal{D}_\mu^s(M)$  consists of those diffeomorphisms  $\eta \in \mathcal{D}^s(M)$  such that  $\eta^*\mu = \mu$ . Continuing with the assumption  $s > \frac{n}{2} + 1$  it is possible to show that  $\mathcal{D}_\mu^s(M)$  is a smooth submanifold of  $\mathcal{D}^s(M)$  whose tangent space at the identity consists of  $H^s$  divergence-free vector fields [7]. The group and algebra adjoint operators defined above restrict in the normal way and the vector space of divergence-free vector fields is closed under the Lie bracket, thereby forming a sub-algebra of the full algebra of vector fields.

In the next section we will put a right invariant metric on the full diffeomorphism group  $\mathcal{D}^s(M)$  and the volume preserving diffeomorphism group  $\mathcal{D}_\mu^s(M)$  and give conditions under which the metric admits a smooth right-invariant Levi-Civita connection and smooth exponential map which is a local diffeomorphism in a neighbourhood of the identity element.

### Right-Invariant Sobolev Metrics

On  $\mathcal{D}^s(M)$  we put a right-invariant  $H^r$  metric defined at the identity by

$$(v, w)_{H^r} = \int_M g(v, A^r w) d\mu \quad (49)$$

for any  $v, w \in T_e\mathcal{D}_\mu^s$ , where  $A^r$  is a self-adjoint, elliptic, invertible differential operator of order  $2r$ . Right-invariance comes from translating the metric around the group in the same we did for an inner product on a finite dimensional Lie algebra  $\mathfrak{g}$  with Lie group  $G$ . Typical examples of  $A^r$  include  $I + \Delta^r$ ,  $(I + \Delta)^r$ , or for  $r$  an integer  $\sum_{k=0}^r \Delta^k$ . Here  $\Delta$  is the Laplace-Beltrami operator of the Riemannian metric  $g$  on  $M$ . For this metric to actually be usable it is often required that  $s > \frac{n}{2} + 1 + 2r$  so that  $A^r w$  is at least  $C^1$ .

#### Theorem 7 [34]

1. Suppose  $\mathcal{D}^s(M)$  is the diffeomorphism group equipped with the right-invariant  $H^r$  metric (49). If  $r \geq 1$  is an integer and  $s > \frac{n}{2} + 2r$  then the metric and its Levi-Civita connection are both  $C^\infty$ . Therefore the Riemannian exponential map is  $C^\infty$  and gives a local diffeomorphism at the identity.

2. Suppose that  $\mathcal{D}_\mu^s(M)$  is the volumorphism group equipped with the right-invariant  $H^r$  metric (49). If  $r \geq 0$  is an integer and  $s > \frac{n}{2} + 1 + 2r$  then the metric and its Levi-Civita connection are both  $C^\infty$ . Therefore the Riemannian exponential map is  $C^\infty$  and gives a local diffeomorphism at the identity.
3. Suppose  $G$  is either  $\mathcal{D}^s(M)$  or  $\mathcal{D}_\mu^s(M)$  and  $r$  and  $s$  satisfy the conditions of either (1) or (2), respectively. Then for any  $v_o \in T_e G$  the map

$$D \exp_e(tv_o) : T_{tv_o} T_e G \rightarrow T_{\exp_e(tv_o)} G$$

is a bounded linear operator satisfying  $D \exp_e(0) = I$ .

Consequently the metrics (49) admit smooth right-invariant Levi-civita connections  $\nabla^r$ , right-invariant curvature tensors  $R^r$ , and carry geodesics defined for at least a short time interval. So as in the first half of this study we can use finite dimensional techniques to analyze the geodesic equation, the derivative of the exponential map, and study conjugate points in terms of the Jacobi equation.

Let  $G^s$  be either  $\mathcal{D}^s(M)$  or  $\mathcal{D}_\mu^s(M)$ ,  $\mathfrak{g}^s$  their tangent spaces at the identity, and assume from now on that  $r$  and  $s$  satisfy the conditions of either (1) or (2) in Theorem 7. A superscript  $r$  will denote the completion of  $\mathfrak{g}^s$  in the norm induced by (49). The group coadjoint operator  $\text{Ad}_g^* : \mathfrak{g}^r \rightarrow \mathfrak{g}^r$  is the metric adjoint of the group adjoint operator

$$(\text{Ad}_g^* v, w)_{H^r} = (v, \text{Ad}_g w)_{H^r}, \quad \forall v, w \in \mathfrak{g} \quad (50)$$

while the algebra coadjoint operator  $\text{ad}_u^* : \mathfrak{g}^r \rightarrow \mathfrak{g}^r$  is the metric adjoint of the algebra adjoint operator

$$(\text{ad}_u^* v, w)_{H^r} = (v, \text{ad}_u w)_{H^r}, \quad \forall v, w \in \mathfrak{g}. \quad (51)$$

In complete parallel with Theorem 1, Corollary 1, and Corollary 2, without change in the proofs, we obtain

**Theorem 8** *If  $G^s$  is the full diffeomorphism group  $\mathcal{D}^s(M)$  or the volumorphism group  $\mathcal{D}_\mu^s(M)$  with a right-invariant metric Sobolev metric (49) such that  $r$  and  $s$  satisfy the conditions of either (1) or (2), respectively, in Theorem 6 then a curve  $\eta(t) \in G^s$  is a geodesic if and only if the curve  $v(t)$  in  $\mathfrak{g}^s$ , defined by the flow equation*

$$\dot{\eta}(t) = dR_{\eta(t)} v(t) \quad (52)$$

satisfies the Euler equations

$$\dot{v}(t) = -\text{ad}_{v(t)}^* v(t). \quad (53)$$

**Corollary 3** *If  $\eta(t)$  is a curve in  $G^s$  with velocity field  $v(t) \in \mathfrak{g}^s$  defined through (52) and satisfying (53) with initial conditions  $\eta(0) = e$  and  $v(0) = v_o$  then we have the conservation law*

$$v(t) = \text{Ad}_{\eta^{-1}(t)}^* v_o. \quad (54)$$

As before, define  $K_v : \mathfrak{g}^r \rightarrow \mathfrak{g}^r$  by

$$K_v(w) = \text{ad}_w^* v. \quad (55)$$

**Corollary 4** *If  $\eta(t)$  is a curve in  $G^s$  with velocity field  $v(t) \in \mathfrak{g}^s$  defined through (52) and satisfying (53) with initial conditions  $\eta(0) = e$  and  $v(0) = v_o$  then*

$$K_{v(t)} = \text{Ad}_{\eta^{-1}(t)}^* K_{v_o} \text{Ad}_{\eta^{-1}(t)}. \quad (56)$$

$H^r$  Compactness of  $K_v$  and  $H^r$  Boundedness of the Adjoint Operators

Properties (A) and (B) are not automatic in infinite dimensions and require proof on a case-by-case basis. We now summarise the results of [34] which demonstrate boundedness and invertibility of the adjoint and coadjoint operators, as well as their product on diffeomorphism groups, and which demonstrate the compactness of the operator  $K_{v(t)}$  on diffeomorphism groups and their Lie algebras.

**Proposition 5** (Lemma 6.3 and Remark 6.4 [34]) *For each of the groups considered in Theorem 6 there is a critical value  $r_o$  such that when  $r > r_o$  and  $s$  is sufficiently large the operator  $K_v : \mathfrak{g}^r \rightarrow \mathfrak{g}^r$  is compact for any  $v \in \mathfrak{g}^s$ . The critical values  $r_o$  and required values of  $r$  and  $s$  so that Theorem 6 is also satisfied are:*

1. For  $\mathcal{D}^s(M)$ ,  $r_o = \frac{1}{2}$ ,  $r \geq 1$  and  $s > 3r + \frac{n}{2}$ ;
2. For  $\mathcal{D}_\mu^s(M^2)$ ,  $r_o = -\frac{1}{2}$ ,  $r \geq 0$  and  $s > 3r + 3$
3. For  $\mathcal{D}_\mu^s(M^n)$ ,  $n \geq 3$ ,  $r_o = 0$ ,  $r \geq 1$  and  $s > 3r + 1 + \frac{n}{2}$

*Remark 4* Proposition 5 holds independently of the well-posedness results given in Theorem 6 and is valid for even smaller values of  $r$  and  $s$ . But the application of these results in our situation require the well-posedness results of Theorem 6 to hold which is why we have stated Proposition in this way. This is explained in [34].

For boundedness and invertibility of the adjoint and coadjoint operators, and their product, we have

**Proposition 6** (Proposition [34]) *Assume that  $r > r_o$  and  $s > \frac{n}{2} + |r| + 1$  for  $G^s = \mathcal{D}^s(M)$  or  $G^s = \mathcal{D}_\mu^s(M)$ . Then for any  $\eta$  in either of the considered groups the maps  $\text{Ad}_\eta$ ,  $\text{Ad}_\eta^*$ , and  $\Lambda_\eta = \text{Ad}_\eta^* \text{Ad}_\eta$  are bounded invertible linear operators on  $\mathfrak{g}^r$ .*

Finally,

**Proposition 7** (Theorem 6.5 [34]) *Assume that  $r$  and  $s$  satisfy the conditions of Proposition 6. Then for a geodesic  $\eta(t)$  of an  $H^r$  metric on one of the groups  $G^s$ , the operator*

$$\begin{aligned} \Omega(t) : \mathfrak{g}^r &\rightarrow \mathfrak{g}^r & (57) \\ \Omega(t)w &= \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(s)}^* w \, ds \end{aligned}$$

*is invertible on  $\mathfrak{g}^r$ .*

*Proof.* The proof of this statement is the first half of the proof of Theorem 6.5 in [34].  $\square$

## Jacobi Fields

From now on we will assume that the Sobolev indices  $s$  and  $r$ , relative to the group  $G^s$  we are talking about are chosen so that the conclusions of Propositions 5 and 6 hold. Consider a geodesic  $\eta(t) = \exp_e(tv_o)$  in  $G^s$  starting from the identity  $e$  in



the direction  $v_o \in \mathfrak{g}^s$  and let  $sz_o$  be a perturbation of  $v_o$  for some small parameter  $s \in (-\epsilon, \epsilon)$  so that  $\eta(s, t) = \exp_e(t(v_o + sz_o))$  is a variation of geodesics; the Jacobi field  $J(t)$  along  $\eta(t)$  satisfying  $J(0) = 0$  and  $\dot{J}(0) = z_o$  is defined by

$$J(t) = \partial_s|_{s=0}\eta(s, t) = D \exp_e(tv_o)tz_o.$$

Denote the solution operator to the Jacobi equation (??) by  $\Phi(t)$

$$\Phi(t)z_o := J(t) \tag{58}$$

which is a bounded linear operator from  $T_eG^s$  to  $T_{\eta(t)}G^s$  by Theorem 6.

**Proposition 8** *Suppose  $G^s$  is one of the groups given in Theorem 6 with right-invariant  $H^r$  metric. Let  $\eta(t)$  be an  $H^s$  geodesic with  $\eta(0) = e$  and  $\dot{\eta}(0) = v_o \in T_eG^s$ . Then every Jacobi field  $J(t)$  along  $\eta(t)$ , with initial conditions  $J(0) = 0$  and  $\dot{J}(0) = z_o$ , satisfies the following system of equations on  $\mathfrak{g}^s$*

$$\partial_t y - \text{ad}_v y = z \tag{59}$$

$$\partial_t z + \text{ad}_v^* z + K_v z = 0 \tag{60}$$

where  $J(t) = dR_{\eta(t)}y(t)$  and  $\dot{\eta}(t) = dR_{\eta(t)}v(t)$  as in (53).

When the exponential map is  $C^1$  the Jacobi system (59) and (60) is well-posed on  $\mathfrak{g}^s$  - this is part (3) of Theorem 6 - despite the fact that the adjoint and coadjoint operators involve derivative loss. The derivation of this system is formally the same as the one given in the proof of Proposition 3.

Equation (20) is the linearised Euler equation and we denote its solution operator by

$$S(t)z_o = z(t). \tag{61}$$

Following the remarks above,  $S(t)$  is a bounded linear operator between  $\mathfrak{g}^r$  and  $\mathfrak{g}^r$ . Denote the solution operator to the Jacobi system (59) and (60) by  $\Phi(t)$

$$\Phi(t)z_o := J(t) \tag{62}$$

which is a bounded linear operator from  $T_eG^s$  to  $T_{\eta(t)}G^s$  by Theorem 7.

**Proposition 9** *Let  $G^s$ ,  $\eta(t)$ , and  $v(t)$  be as in Proposition 8 with  $s$  and  $r$  chosen so that the conclusions of Propositions 5 and 6 hold. Then the solution operator (61) to the linearised Euler equation (60) is related to the solution operator (62) to the Jacobi equation (??) by*

$$\text{Ad}_{\eta^{-1}(t)}^* = S(t) + K_{v(t)} \circ dR_{\eta^{-1}(t)} \circ \Phi(t) \tag{63}$$

which is defined on  $\mathfrak{g}^r$

*Proof.* Identical to the proof of Proposition 4. □ □

Analogous to Theorem 2 we see that the Euler-Arnold equations can be written as a feedback system and the geodesic evolution is the time average of the linearised geodesic evolution with the same initial conditions.

**Theorem 9** *The Euler-Arnold equations of Theorem 8 can be re-written as a feedback system*

$$\begin{aligned} \partial_t \tilde{v}(t) &= -\text{ad}_{v(t)}^* \tilde{v}(t) - K_{v(t)} \tilde{v}(t) \\ v(t) &= \frac{1}{t} \int_0^t \tilde{v}(s) ds \\ v(0) &= \tilde{v}(0) = v_o \end{aligned}$$

## The Exponential Map

Before stating the infinite dimensional analogue of Theorem 3, let us summarise what has been done so far. Given a geodesic  $\eta(t)$  of a right-invariant  $H^r$  metric, belonging to one of the groups  $G^s$ , with  $H^s$  velocity vector  $dR_{\eta(t)^{-1}} \cdot \dot{\eta}(t) = v(t) \in \mathfrak{g}^s$ , the solution operator to the linearized Euler equations is related to the coadjoint operator and the solution operator to the Jacobi equation via

$$\text{Ad}_{\eta^{-1}(t)}^* = S(t) + K_{v(t)} \circ dR_{\eta^{-1}(t)} \circ \Phi(t)$$

which is valid on  $\mathfrak{g}^r$  for  $r < s$ . If, in addition,  $r$  and  $s$  satisfy the hypotheses of Propositions 5 and 6 then the operator  $K_v : \mathfrak{g}^r \rightarrow \mathfrak{g}^r$  is compact for all  $v \in \mathfrak{g}^s$  and the operators  $\text{Ad}_\eta$ ,  $\text{Ad}_\eta^*$ ,  $\text{Ad}_\eta^* \text{Ad}_\eta$ , and  $\Omega(t)$  are invertible bounded linear operators on  $\mathfrak{g}^r$ .

**Theorem 10** *Suppose  $G^s$  is any one of the above Lie groups with right-invariant  $H^r$  metric  $(\cdot, \cdot)_{H^r}$  and that  $r$  and  $s$  satisfy the hypotheses of Propositions 5 and 6. Let  $\eta(t)$  be a geodesic with  $\eta(0) = e$  and  $\dot{\eta}(0) = v_o \in \mathfrak{g}^s$ , let  $J(t) = dR_{\eta(t)}y(t)$  be the Jacobi field along  $\eta(t)$  with initial conditions  $J(0) = 0$ ,  $\dot{J}(0) = z_o \in \mathfrak{g}^r$ , and let  $z(t)$  be the solution to (60) with initial condition  $z(0) = z_o \in \mathfrak{g}^r$ . Then*

1. *If  $z_o \in (\ker K_{v_o})^\perp \subset T_e G^r$  then  $y(t) \in (\ker K_{v(t)})^\perp \subset \mathfrak{g}^r$  for all  $t$  and the derivative of the exponential map can be represented on  $\mathfrak{g}^r$  by*

$$D \exp_e(tv_o)tz_o = dR_{\eta(t)} \tilde{K}_{v(t)}^{-1} (\text{Ad}_{\eta^{-1}(t)}^* - S(t)) z_o, \quad (64)$$

where  $\tilde{K}_{v(t)}$  denotes the restriction of  $K_{v(t)}$  to  $(\ker K_{v(t)})^\perp$  on which it is an isomorphism.

2. *If  $z_o \in \ker K_{v_o} \subset \mathfrak{g}^r$  then  $y(t) \in \ker K_{v(t)} \subset \mathfrak{g}^r$  for all  $t$  and is never zero for  $t > 0$ . The derivative of the exponential map is given by*

$$D \exp_e(tv_o)tz_o = dL_{\eta(t)} \int_0^t \text{Ad}_{\eta^{-1}(s)} \text{Ad}_{\eta^{-1}(t)}^* z_o ds. \quad (65)$$

*Proof.* We'll step through the four Lemmas proved in the finite dimensional case and point out where the results on  $K_v$ ,  $\text{Ad}_\eta$ ,  $\text{Ad}_\eta^*$ ,  $\text{Ad}_\eta^* \text{Ad}_\eta$ , and  $\Omega(t)$  should be used.

Lemma 1: follows just as in the finite dimensional case using Corollary 4, and boundedness and invertibility of  $\text{Ad}_\eta$  and  $\text{Ad}_\eta^*$ .

Lemma 2: follows the finite dimensional proof substituting invertibility of  $\Omega(t)$  in the  $H^r$  topology (Proposition 7); invertibility and boundedness of  $\text{Ad}_{\eta^{-1}} \text{Ad}_{\eta^{-1}}^*$  in the  $H^r$  topology (Proposition 6); and the orthonormal basis of eigenvectors spanning  $H^r$  which is guaranteed by the spectral theorem for compact (skew) self-adjoint operators (Proposition 5).

Lemma 3: follows the finite dimensional proof substituting corollary 4; and the orthonormal basis of eigenvectors spanning  $H^r$  which is guaranteed by the spectral theorem for compact (skew) self-adjoint operators.

The remainder of the proof follows exactly the final part of the proof of Theorem 3 substituting the orthonormal basis of eigenvectors spanning  $H^r$  which is guaranteed by the spectral theorem for compact (skew) self-adjoint operators; boundedness and invertibility of  $\text{Ad}_\eta$  and  $\text{Ad}_\eta^*$ , boundedness of the solution operator  $S(t)$  on  $H^r$ ; and the analogous applications of Lemmas 1 - 3.

□

*Remark 5* If the geodesic  $\eta(t)$  and its velocity field  $v(t)$  are  $C^\infty$  smooth then the formula for the derivative of the exponential map is valid on  $\mathfrak{g}^s$  in the strong  $H^s$  topology.

### Conjugate Points

In contrast to finite dimensions, a linear operator can be injective while failing to be surjective (for example, the shift map in  $l^2$ ) and, consequently, two types of conjugacies can occur in infinite dimensions. A point at which the derivative of the exponential map is not injective will be called a monoconjugate point, while a point at which the derivative is not surjective will be called an epiconjugate point. On an infinite dimensional ellipsoid in  $l^2$  with the induced metric, Grossman [13] showed that there exist finite geodesic segments which contain both types of conjugate points, while on an infinite dimensional sphere in  $l^2$  there exist conjugate points of infinite order. In finite dimensional geometry this phenomenon cannot occur, but Smale [43] introduced a class of maps in infinite dimensions for which this sort of behaviour is excluded.

A bounded linear operator between Banach spaces is said to be Fredholm if it has closed range and its kernel and co-kernel are finite dimensional; the index is defined as the difference between the dimension of the kernel and the dimension of the co-kernel. A smooth map  $f$  is said to be a Fredholm map if its Frechet derivative  $df(p)$  is a Fredholm operator at each  $p$ .

Ebin, Misiolek, and Preston [8] proved that the  $L^2$  ( $H^0$ ) exponential map  $\exp_e$  on  $\mathcal{D}_\mu^s(M^2)$  is a non-linear Fredholm map of index zero, which directly implies that monoconjugate and epiconjugate points coincide, and are isolated and of finite multiplicity along isolated geodesic segments. By the Sard-Smale Theorem the set of singularities form a set of first Baire category; thus the structure of conjugate points of  $\exp_e$  looks like that of a finite dimensional manifold. However, they also showed that the  $L^2$  exponential map on  $\mathcal{D}_\mu^s(M^3)$  for a closed 3-dimensional manifold is *not* Fredholm by constructing a geodesic along which a sequence of monoconjugate points accumulate on a strictly epiconjugate point. These results were extended by Misiolek and Preston [34] to the more general setting of Sobolev  $H^s$  diffeomorphism groups with a weak right-invariant Sobolev  $H^r$  metric, which is also the setting of this section. The same properties of the adjoint operators and  $K_v$  which are used to prove the Fredholm property are those which allow Theorem 3 in the finite dimensional setting to lift to the infinite dimensional setting of Theorem 9. In parallel with Theorem 4 we have

**Theorem 11** *Let  $G^s$ ,  $\mathfrak{g}^s$ ,  $(\cdot, \cdot)_{H^r}$  with  $s$  and  $r$  as in Theorem 9, and suppose the geodesic  $\eta(t)$  and its velocity field  $v(t)$  are  $C^\infty$  smooth. Then  $\eta(t^*)$  is conjugate to  $\eta(0) = e$  if and only if there exists a  $z_o \in (\ker K_{v(0)})^\perp$  such that*

$$\text{Ad}_{\eta^{-1}(t)}^* z_o = S(t^*) z_o$$

The smoothness assumption is not restrictive: almost all known examples of conjugate points in  $\mathcal{D}_\mu^s(M^2)$  and  $\mathcal{D}_\mu^s(M^3)$  occur along smooth stationary geodesics (geodesics whose velocity vector is independent of time) - see for example, [31], [32],[42], [27], [5] - the other examples are constructed along smooth *non-stationary*

geodesics [4]. We are unaware of any examples of conjugate points along non-smooth geodesics.

There is also a growing interest in the exact relationship between conjugate points and the qualitative behaviour of hydrodynamical systems (see the aforementioned references, or Hoffman and Manning [16] and Manning [28] for conjugate points in a different physical setting).

### Eulerian and Lagrangian Stability on Diffeomorphism Groups

For  $k \geq 0$ , a geodesic  $\eta \in \mathcal{D}_\mu^s$  is said to be Lagrangian  $H^k$  (linearly) stable, if every solution of the Jacobi Equation (??) along  $\eta$  remains bounded in the Sobolev  $H^k$  norm.

A solution to the Euler-Arnold equations (53) is Eulerian  $H^k$  stable if every solution of the linearised Euler equations (60) remains bounded in the Sobolev  $H^k$  norm. Intuitively, a solution is Eulerian stable if nearby velocity fields remain nearby.

Here we give a small application to the  $L^2$  Lagrangian and Eulerian stability of geodesics of the  $L^2$  metric on the volume-preserving diffeomorphism group  $\mathcal{D}_\mu^s(M^2)$  of a closed surface with trivial cohomology. A similar study was first given by Preston [41] which related Lagrangian and Eulerian stability assuming some topological restrictions on the solutions considered.

Any  $H^s$  divergence-free vector field on  $M^2$  can be written as the rotated gradient  $v_F = \nabla^\perp F$  of an  $H^{s+1}$  function  $F : M^2 \rightarrow \mathbb{R}$ . In this case the group adjoint operator can be written on functions using the chain rule as

$$\text{Ad}_\eta v_F = v_{F \circ \eta^{-1}}.$$

Define a metric on  $T_e \mathcal{D}_\mu^s$  by

$$\langle v_F, v_G \rangle = \int_M F \cdot G \, d\mu$$

which is related to the standard  $L^2$  metric on vector fields

$$(v_F, v_G)_{L^2} = \int_M g(v_F, v_G) \, d\mu$$

by

$$(v_F, v_G)_{L^2} = -\langle v_{\Delta F}, v_G \rangle$$

where  $\Delta$  is the Laplace operator. A short calculation then shows that the group coadjoint operator is

$$\text{Ad}_\eta^* v_F = v_{\Delta^{-1}(\Delta F \circ \eta)}.$$

**Theorem 12** *Let  $\mathcal{D}_\mu^s(M^2)$  be the volume preserving diffeomorphism group of a closed surface  $M^2$  with trivial cohomology and suppose that the conditions of Theorem 6, and Propositions 5, 6, and 7, are met. Suppose  $\eta(t)$  is a stationary geodesic of the  $L^2$  metric with velocity  $v_o$ .*

1. If  $z_o \in \ker K_{v_o}$  then the norm of  $z(t)$  is uniformly bounded in  $t$  while the corresponding Jacobi field grows at most linearly in  $t$

$$\|z(t)\|_{L^2} \leq C_1 \|z_o\|_{L^2}, \quad \|J(t)\|_{L^2} \leq C_2 t \|z_o\|_{L^2}$$

for finite constants  $C_1$  and  $C_2$ ;

2. If  $z_o \in (\ker K_{v_o})^\perp$  then

$$\|z(t)\|_{L^2} \leq C_2 \cdot \|z_o\|_{L^2} + C_3 \|J(t)\|_{L^2}.$$

for finite constants  $C_2$ , and  $C_3$ ; and

$$\|J(t)\|_{L^2} \approx \mathcal{O}(\|z(t)\|_{L^2}) \quad \text{as } t \rightarrow \infty$$

*Proof.* We'll begin with an observation: for any  $\eta \in \mathcal{D}_\mu^s(M^2)$  and any  $v_F \in T_e \mathcal{D}^s(M^2)$

$$\|\text{Ad}_\eta^* v_F\|_{L^2} \leq C \|v_F\|_{H^s}$$

for some finite constant  $C$ . To see this, let  $v_F \in T_e \mathcal{D}_\mu^s(M^2)$  and use the expression for the coadjoint operator and the relationship between the two metrics:

$$\begin{aligned} \|\text{Ad}_\eta^* v_F\|_{L^2} &= (v_{\Delta^{-1}(\Delta F \circ \eta^{-1})}, v_{\Delta^{-1}(\Delta F \circ \eta^{-1})})_{L^2} = -\langle v_{(\Delta F \circ \eta^{-1})}, v_{\Delta^{-1}(\Delta F \circ \eta^{-1})} \rangle \\ &\leq \left( \int_M (\Delta F \circ \eta) \cdot (\Delta F \circ \eta) d\mu \right)^{\frac{1}{2}} \cdot \left( \int_M (\Delta^{-1}(\Delta F \circ \eta)) \cdot (\Delta^{-1}(\Delta F \circ \eta)) d\mu \right)^{\frac{1}{2}} \\ &\leq C \left( \int_M (\Delta F \circ \eta) \cdot (\Delta F \circ \eta) d\mu \right) \quad \text{since } \Delta^{-1} \text{ is a bounded linear operator} \\ &= C \left( \int_M (\Delta F) \cdot (\Delta F) d\mu \right) \quad \text{since } \eta \text{ is volume preserving} \\ &= C \cdot \|F\|_{H^2} \\ &\leq \tilde{C} \cdot \|v_F\|_{H^2} \quad \text{by the Poincaré inequality} \\ &\leq \tilde{C} \cdot \|v_F\|_{H^s} \quad \text{since } s > 2. \end{aligned}$$

With this observation in hand the remainder of the proof is almost identical to the proof of Theorem 5. Part 1. follows in exactly the same way as part 1 of Theorem 5. Part 2. follows exactly that of the first half of part 3. in Theorem 5 because  $K_{v_o}$  has bounded spectrum. But the second half of the proof of part 3. in Theorem 5 breaks down in this setting because  $K_{v_o}$  is compact and it's spectrum accumulates on 0 - therefore there is no way to bound  $K_{v_o} y(t)$  from below by the smallest eigenvalue. We can, however, assert some long time behaviour: by Proposition 9 and the above observation we have

$$\|K_{v_o} y(t)\|_{L^2} \leq \|\text{Ad}_{\eta(t)^{-1} z_o}^*\|_{L^2} + \|z(t)\|_{L^2} \leq C \cdot \|z_o\|_{L^2} + \|z(t)\|_{L^2}.$$

If we expand  $y(t)$  in the orthonormal basis of eigenvectors  $\{\psi_i\}$  of  $K_{v_o}$  as

$$y(t) = \sum_i h_i(t) \cdot \psi_i$$

and calculate the  $L^2$  norm of  $K_{v_o} y(t)$  we get

$$\|K_{v_o} y(t)\|_{L^2}^2 = \sum_i |\lambda_i|^2 h_i(t)^2.$$

So the  $L^2$  norm of  $K_{v_o}y(t)$  is a re-weighting of the  $L^2$  norm of  $y(t)$  by positive constants and for very large  $t$   $\|y(t)\|_{L^2} \sim \|K_{v_o}y(t)\|_{L^2}$ ; therefore the long time growth is

$$\|J(t)\|_{L^2} = \|y(t)\|_{L^2} \approx \mathcal{O}(\|z(t)\|_{L^2}) \quad \text{as } t \rightarrow \infty.$$

□

□

### Towards a BCH Formula on the Diffeomorphism Group with a Right-Invariant Metric

Poincaré's derivation of the BCH series for the Lie group exponential map begins with his formula for the derivative of the exponential map

$$De^X \cdot Y = dR_{e^X} \frac{e^{\text{ad}_X} - I}{\text{ad}_X}$$

where

$$\frac{e^{\text{ad}_X} - I}{\text{ad}_X} = f(\text{ad}_X) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_X^k$$

is expressed using the functional calculus for the analytic function

$$f(x) = \frac{e^x - 1}{x}.$$

In a sufficiently small neighbourhood  $U$  of 0 in  $\mathfrak{g}$  this function is invertible

$$f^{-1}(x) = \frac{x}{e^x - 1},$$

and using standard properties of the functional calculus

$$f(\text{ad}_X)^{-1} = f^{-1}(\text{ad}_X) = \frac{\text{ad}_X}{I - e^{\text{ad}_X}}$$

which gives an analytic mapping:  $\mathfrak{g} \supset U \ni X \mapsto f^{-1}(\text{ad}_X) \in L(\mathfrak{g}, \mathfrak{g})$ . Setting

$$e^{Z(t)} = e^{tX} \cdot e^Y$$

we can derive a differential equation for  $Z(t)$  and compute the value of  $Z(1)$ , which will express group multiplication using Lie algebraic data. Two good references for the details are [6] and [14]. Continuing formally and ignoring all analytic details of convergence, the time derivative of the left hand side can be computed with Poincaré's formula and gives

$$dR_{e^{Z(t)}} f(\text{ad}_{Z(t)}) \cdot Z'(t) = dR_{e^{tX} \cdot e^Y} f(\text{ad}_{Z(t)}) \cdot Z'(t)$$

while on the right hand side we have, using properties of flows discussed in the Algebra and Geometry section,

$$\frac{d}{dt} e^{tX} \cdot e^Y = dR_{e^{tX} \cdot e^Y} \cdot X.$$

Together, these give

$$Z'(t) = f^{-1}(\text{ad}_{Z(t)}) \cdot X = \frac{\text{ad}_{Z(t)}}{I - e^{\text{ad}_{Z(t)}}} \cdot X.$$

We need to rewrite  $\text{ad}_{Z(t)}$  in terms of  $X$  and  $Y$  and the identity to use is

$$\text{Ad}_{e^V} = e^{\text{ad}_V}.$$

This is not hard to prove and can be deduced from the definition of the group and algebra adjoint operators and the fact that they are representations, but the reader may like to consult the suggested references. Then

$$e^{\text{ad}_{Z(t)}} = \text{Ad}_{e^{Z(t)}} = \text{Ad}_{e^{tX} \cdot e^Y} = \text{Ad}_{e^{tX}} \cdot \text{Ad}_{e^Y} = e^{t\text{ad}_X} \cdot e^{\text{ad}_Y}.$$

Now for any  $V$  we can write

$$V = \log\left(I + (e^V - I)\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} (e^V - I)^{k+1}$$

using the Taylor series for the log, which gives

$$\frac{V}{e^V - I} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} (e^V - I)^k.$$

Substituting this back into the ODE for  $Z$  with  $V = \text{ad}_{Z(t)}$  and integrating gives

$$Z(1) = X + \int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} (e^{t\text{ad}_X} \cdot e^{\text{ad}_Y} - I)^k dt \cdot X.$$

Inserting the series expansions for the exponentials, expanding, collecting terms, and integrating leads to Dynkin's formula which expresses  $Z(1)$  as a series of iterated Lie brackets of  $X$  and  $Y$ .

We would like to mirror this kind of derivation to obtain an analogous formula a right-invariant exponential map and we now outline a program of future work for doing this. The starting point should be a relation of the form

$$\exp_e(sZ(t) + Y) = \exp_e(stX) \star \exp_e(Y) \tag{66}$$

with  $Z(0) = 0$ . Because our formula was derived in terms of Jacobi fields which are zero at time  $t = 0$  we need to consider 2-parameter variations of geodesics - a "shooting" parameter  $s$  and a variation parameter  $t$ . Letting  $\eta_t(s) = \eta(s, t) = \exp_e(sZ(t) + Y)$  we consider the formula for the derivative of the exponential map given in Theorem 3 to be along the geodesic  $\eta_t(s)$ . The time derivative of the opening expression (66) gives

$$D \exp_e(sZ(t) + Y) \cdot sZ'(t) = s dR_{\exp_e(stX) \star \exp_e(Y)} \cdot X(t)$$

and at  $s = 1$  we have

$$D \exp_e(Z(t) + Y) \cdot Z'(t) = dR_{\exp_e(tX) \star \exp_e(Y)} \cdot X(t).$$

Here we hit our first road block: the formulas (64) and (65) for the derivative of the exponential map differ depending on whether or not  $Z'(t)$  belongs to  $\ker K_{sZ(t)+Y}$  or  $\ker K_{sZ(t)+Y}^\perp$ . By Lemma 1 each of these subspaces is closed under the Lie bracket but we either need to know that their images under the exponential map are closed subgroups of  $G$  or if we can place some reasonable assumptions on  $X$  and  $Y$  so that  $Z(t)$  remains in one of the subspaces for all  $t$ :

**Problem 1** Determine if  $\exp_e(\ker K_V)$  and  $\exp_e(\ker K_V)^\perp$  are closed subgroups. Or determine conditions on  $X$  and  $Y$  so that a pre-image of the group product belongs to one subspace.

Assuming we solve Problem 1 and know which formula for  $D\exp_e$  should be applied and inverted, the next problem is the following: although the original BCH problem was to solve an equation in two *non-commuting* indeterminants, only one operator and its infinitesimal generator appears in Poincaré's formula - the group adjoint operator and the algebra adjoint operator, which *do commute*, and are self-adjoint so that a good functional calculus exists for the remaining analysis. There are four operators appearing in our formulas for the derivative of the exponential map, and although they are nicely related it seems like they will never commute; thus our problem involves dealing with a functional calculus of several non-commuting operators. Several approaches to this problem exist: for example, the functional calculus developed via Clifford analysis by Jeffries [18]; functions of non-commuting operators developed by V. P. Maslov and company [29], [35]; the non-commuting variable approach of Taylor [45], [46]; or Feynman's operational calculus for systems of non-commuting operators [9]. It's not obvious which calculus is right for this setting and may even depend on the algebraic or topological structure of  $G$ , and seems like a difficult problem:

**Problem 2** Determine a good spectral representation for differences and products of the non-commuting operators  $\text{Ad}$ ,  $\text{Ad}^*$ ,  $S$ , and  $K$ .

In view of Theorems 5 and 6 a good starting place for understanding problems 1 and 2 would be compact Lie groups or the Heisenberg groups.

## References

1. Achilles, R., Bonfiglioli, A.: The early proofs of the theorem of campbell, baker, hausdorff, and dynkin. *Archive for history of exact sciences* **66**, 295–358 (2012)
2. Arnold, V.I.: Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Annales de L'Institut Fourier (Grenoble)* **16**(1), 319–361 (1966)
3. Banyaga, A.: The structure of classical diffeomorphism groups, vol. 400. Springer Science & Business Media (2013)
4. Benn, J.: Conjugate points in  $d\mu_s$  (s2). *Journal of Geometry and Physics* **170**, 104369 (2021)
5. Drivas, T.D., Misiołek, G., Shi, B., Yoneda, T.: Conjugate and cut points in ideal fluid motion. *Annales mathématiques du Québec* pp. 1–19 (2022)
6. Duistermaat, J.J., Kolk, J.A.: Lie groups. Springer Science & Business Media (2012)
7. Ebin, D.G., Marsden, J.E.: Groups of diffeomorphisms and the motion of an incompressible fluid. *Annals of Mathematics* **92**(1), 102–163 (1970)
8. Ebin, D.G., Misiołek, G., Preston, S.C.: Singularities of the exponential map on the volume-preserving diffeomorphism group. *Geometric and Functional Analysis* **16**(4), 850–868 (2006)



9. Feynman, R.P.: An operator calculus having applications in quantum electrodynamics. *Physical Review* **84**(1), 108 (1951)
10. Grabowski, J.: Groups of diffeomorphisms and lie theory. *Proceedings of the 11th Winter School on Abstract Analysis* pp. 141–147 (1984)
11. Grabowski, J.: Free subgroups of diffeomorphism groups. *Fundamenta Mathematicae* **2**(131), 103–121 (1988)
12. Green, L.W.: Auf weiderschensflächen. *Annals of Mathematics* (1963)
13. Grossman, N.: Hilbert manifolds without epiconjugate points. *Proceedings of the American Mathematical Society* **16**(6), 1365–1371 (1965)
14. Hall, B.C., Hall, B.C.: *Lie groups, Lie algebras, and representations*. Springer (2013)
15. Hebey, E.: *Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities*. American Mathematical Society (200)
16. Hoffman, K.A., Manning, R.S.: An extended conjugate point theory with application to the stability of planar buckling of an elastic rod subject to a repulsive self-potential. *SIAM journal on mathematical analysis* **41**(2), 465–494 (2009)
17. Inci, H., Kappeler, T., Topalov, P.: *On the Regularity of the Composition of Diffeomorphisms*. AMS (2013)
18. Jefferies, B.: *Spectral Properties of Non-Commuting Operators*. Springer (2004)
19. Kaiser, V.V.: Conjugate points of left-invariant metrics on lie groups. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika* (11), 27–37 (1990)
20. Khesin, B., Misiólek, G., Shnirelman, A.: Geometric hydrodynamics in open problems. *Archive for Rational Mechanics and Analysis* **247**(2), 15 (2023)
21. Klingenberg, W.: Manifolds with restricted conjugate locus. *Annals of Mathematics* pp. 527–547 (1963)
22. Klingenberg, W.P.: *Riemannian geometry, vol. 1*. Walter de Gruyter (2011)
23. Kobayashi, S.: Riemannian manifolds without conjugate points. *Annali di Matematica Pura ed Applicata* **53**(1), 149–155 (1961)
24. Koppel, N.: Commuting diffeomorphisms of the circle. In: *Sixth Brazilian Math. Colloq.*, pp. 18–22 (1970)
25. Kriegl, A., Michor, P.W.: *The convenient setting of global analysis, vol. 53*. American Mathematical Soc. (1997)
26. Lee, J.M.: *Smooth Manifolds*. Springer New York, New York, NY (2012)
27. Lichtenfelz, L., Tauchi, T., Yoneda, T.: Existence of a conjugate point in the incompressible euler flow on a three-dimensional ellipsoid. *Arnold Mathematical Journal* (2023)
28. Manning, R.S.: Conjugate points revisited and neumann–neumann problems. *SIAM review* **51**(1), 193–212 (2009)
29. Maslov, V.P., Golo, V., Kulman, N., Voropaeva, G.: *Operational methods*. (No Title) (1976)
30. Milnor, J.: Remarks on infinite-dimensional lie groups. In: *Relativity, groups and topology*. 2 (1984)
31. Misiólek, G.: *Stability of flows of ideal fluids and the geometry of the group of diffeomorphisms*. State University of New York at Stony Brook (1992)
32. Misiólek, G.: Conjugate points in  $d\mu$  (t2). *Proceedings of the American Mathematical Society* pp. 977–982 (1996)
33. Misiólek, G., Preston, S.: Fredholm properties of Riemannian exponential maps on diffeomorphism groups. *Inventiones Mathematicae* (2010)
34. Misiólek, G., Preston, S.C.: Fredholm properties of Riemannian exponential maps on diffeomorphism groups. *Inventiones mathematicae* **179**, 191–227 (2010)
35. Nazaikinskii, V.E., Shatalov, V.E., Sternin, B.Y.: *Methods of noncommutative analysis: theory and applications, vol. 22*. Walter de Gruyter (2011)
36. Neeb, K.H.: Towards a lie theory of locally convex groups. *Japanese Journal of Mathematics* **1**, 291–468 (2006)
37. Omori, H.: *Infinite-dimensional Lie groups, vol. 158*. American Mathematical Soc. (2017)
38. Palis, J.: *Vector fields generate few diffeomorphisms* (1974)
39. Pennec, X., Sommer, S., Fletcher, T.: *Riemannian geometric statistics in medical image analysis*. Academic Press (2019)
40. Pressley, A.N., Segal, G.B.: *Loop groups and their representations*. Oxford University Press, Oxford (1985)
41. Preston, S.C.: For ideal fluids, eulerian and lagrangian instabilities are equivalent. *Geometric and Functional Analysis* **14**(5), 1044–1062 (2004)

42. Preston, S.C.: Conjugate point criteria on the area-preserving diffeomorphism group. *Journal of Geometry and Physics* **183**, 104680 (2023)
43. Smale, S.: An infinite dimensional version of sard's theorem. In: *The Collected Papers of Stephen Smale: Volume 2*, pp. 529–534. World Scientific (2000)
44. Sokolowski, L.M.: On the twin paradox in static spacetimes: I. schwarzschild metric. *General Relativity and Gravitation* **44**, 1267–1283 (2012)
45. Taylor, J.L.: A general framework for a multi-operator functional calculus. *Advances in Mathematics* **9**(2), 183–252 (1972)
46. Taylor, J.L.: *Functions of several noncommuting variables* (1973)
47. Thurston, W.: *Foliations and groups of diffeomorphisms* (1974)
48. Warner, F.: *Foundations of Differentiable Manifolds and Lie Groups*. Springer (1983)
49. Warner, F.W.: Conjugate loci of constant order. *Annals of Mathematics* pp. 192–212 (1967)
50. Zajtz, A.: Calculus of flows on convenient manifolds. *Archivum Mathematicum* **32**(4), 355–372 (1996)