

## Geodesics and connexions on matrix Lie groups (2015 ed.)

**Citation for published version (APA):**

Graaf, de, J. (2015). *Geodesics and connexions on matrix Lie groups (2015 ed.)*. (CASA-report; Vol. 1521). Technische Universiteit Eindhoven.

**Document status and date:**

Published: 01/01/2015

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

**EINDHOVEN UNIVERSITY OF TECHNOLOGY**  
Department of Mathematics and Computer Science

CASA-Report 15-21  
May 2015

Geodesics and connexions on matrix Lie groups  
2015 ed.

by

J. de Graaf



Centre for Analysis, Scientific computing and Applications  
Department of Mathematics and Computer Science  
Eindhoven University of Technology  
P.O. Box 513  
5600 MB Eindhoven, The Netherlands  
ISSN: 0926-4507



# Geodesics and Connexions on Matrix Lie Groups

2015 ed.

J. de GRAAF

## Abstract

These notes present some elementary considerations on Geodesics and Connexions on Matrix Lie Groups. As motivating examples two sections with special attention to the mechanics of solid bodies are included.

This text is originally intended as and grew out of an Appendix to my Lecture Notes on 'Tensor Calculus and Differential Geometry' [dG]. The only prerequisites are some skill in 'Calculus in  $\mathbb{R}^n$ ' and 'Elementary Linear Algebra'. The text mainly concerns 'Advanced Matrix Theory'. It could be useful for numerical calculations.

In this new (2015)-edition the proofs in Ch. 5-7 have been somewhat extended. Also some alternative viewpoints have been added there.

## Notation

THREE DIFFERENT FONTS ARE USED FOR  $n \times n$ -MATRICES:

- Capitals of type P, Q, X, Y, ... for matrices belonging to the Lie group  $\mathbf{G}_{\mathcal{M}} \subset \mathbb{R}^{n \times n}$  under consideration.
- Capitals of type  $\mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{F}, \mathcal{G}, \dots$  for matrices belonging to the Lie algebra  $\mathfrak{g}_{\mathcal{M}} \subset \mathbb{R}^{n \times n}$  under consideration.
- Ordinary capitals like A, B, J, K, L, ... for non-specified matrices belonging to  $\mathbb{R}^{n \times n}$ .

FURTHER

- Script capitals like  $\mathcal{J}, \mathcal{H}, \mathcal{K}, \mathcal{R}, \dots$  are used for linear mappings  $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ ,  $\mathbb{R}^{n \times n} \rightarrow \mathfrak{g}_{\mathcal{M}}$ , or  $\mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$ .
- Gothic capitals like  $\mathfrak{A}, \mathfrak{B}, \mathfrak{F}, \mathfrak{G}, \mathfrak{H}, \dots$  are used for tangent vector fields (= 'tangent matrix fields') to  $\mathbf{G}_{\mathcal{M}} \subset \mathbb{R}^{n \times n}$ .

## Contents

### 1. Rigid Body Motion p 4

First the full algebraic correspondence between  $\mathbb{R}^3$  and the (Lie algebra of) skew symmetric matrices in  $\mathbb{R}^{3 \times 3}$  is established. Next, starting from Newton's 2nd Law, the equations of motion for a freely rotating rigid body are derived. Concepts like *objective* and *relative angular velocity* are clarified by using both vector and matrix descriptions. Finally on the orthogonal group  $\mathbf{SO}(3)$ , equations for the 'shortest path' (geodesics) with respect to a left invariant inner product are derived. They turn out to be the same as the equations of motion for the rigid body if the inner product is suitably related to the *Inertia Tensor*.

### 2. A special class of Matrix Lie Groups p 8

The subject of study is the Matrix Lie Group

$$\mathbf{G}(n, \mathbb{R}, \Delta) = \{ \mathbf{P} \in \mathbb{R}^{n \times n} \mid \mathbf{P}^\top \Delta \mathbf{P} = \Delta \},$$

for  $\Delta \in \mathbb{R}^{n \times n}$  with  $\det \Delta \neq 0$ , and the corresponding Matrix Lie-Algebra

$$\mathfrak{g}(n, \mathbb{R}, \Delta) = \{ \mathcal{A} \in \mathbb{R}^{n \times n} \mid \mathcal{A}^\top \Delta + \Delta \mathcal{A} = \mathbf{O} \},$$

The treatment is elementary and straightforward. Some attention is given to special invariant inner products.

### 3. Traces and Bilinear Forms on Spaces of Matrices p 12

The sole purpose of this section is to show that inner products, which are both left and right invariant, are available.

### 4. General Matrix Lie Groups p 13

The Lie Algebra  $\mathfrak{g}_{\mathcal{M}} \subset \mathbb{R}^{n \times n}$  (and its properties) of a general Lie Group  $\mathbf{G}_{\mathcal{M}} \subset \mathbb{R}^{n \times n}$  are obtained by means of the limit  $\lim_{n \rightarrow \infty} \{ \mathbf{X}(\frac{t}{n}) \}^n = e^{t\mathbf{X}'(0)}$  for any curve  $t \mapsto \mathbf{X}(t) \in \mathbf{G}_{\mathcal{M}}$ , with  $\mathbf{X}(0) = \mathbf{I}$ .

### 5. Left and Right Geodesics on Matrix Lie Groups p 16

Starting from the standard inner product  $\text{Trace}[X^\top Y]$  on  $\mathbb{R}^{n \times n}$ , the orthogonal projection  $\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{g}_{\mathcal{M}}$  and a symmetric linear map  $\mathcal{J} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$ , two versions of geodesic equations are derived from a variational principle. Only left (or right) invariant inner products on the tangent spaces of  $\mathbf{G}_{\mathcal{M}}$  are considered. The abstract version uses concepts like  $\text{ad}$  and  $\text{Ad}$ . The more down to earth version only involves the mappings  $\mathcal{J}$  and  $\mathcal{P}_{\mathfrak{g}}$ . For 6 distinct types of Matrix Lie Groups the geodesic equations are written out in full.

**6. Functions and Vectorfields on Matrix Lie Groups** p 28

For tangent vectorfields on  $\mathbf{G}_{\mathcal{M}}$  the left (or right) description is introduced. The action of a vectorfield on a function is defined by differentiation along an exponential curve. The Lie bracket and its properties now follow from very elementary considerations.

**7. Parallel Transport. Connexions** p 31

All possible left invariant Koszul Connexions  $\nabla$  on  $\mathbf{G}_{\mathcal{M}}$  are characterised by means of differentiation along an exponential curve  $t \mapsto \mathbf{P}e^{t\mathbf{A}}$  in combination with a fixed  $\mathfrak{g}_{\mathcal{M}}$ -valued 2-tensor  $\mathcal{B}(\cdot, \cdot)$  on  $\mathfrak{g}_{\mathcal{M}}$ . Properties of  $\nabla$  like 'torsionfree', 'inner product preserving', Riemannian, are explained in terms of the 'Christoffel tensor'  $\mathcal{B}$ . The results are related to section 5 on geodesics.

**A. Appendix. Note on the Symmetric Top** p 38

- **Acknowledgement** p 40
- **References** p 40

# 1 Rigid Body Motion

To a vector  $\underline{a} \in \mathbb{R}^3$  we relate an anti-symmetric matrix  $\mathcal{M}(\underline{a}) \in \mathbb{R}^{3 \times 3}$  in the usual way. Note the following correspondences and identities

$$\begin{aligned}
\underline{a} = \text{column}[a_1, a_2, a_3] &\longleftrightarrow \mathcal{M}(\underline{a}) = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \\
\underline{a} \times \underline{x} &= \mathcal{M}(\underline{a}) \underline{x} \\
\underline{a} \times \underline{b} &\longleftrightarrow \mathcal{M}(\underline{a} \times \underline{b}) = \mathcal{M}(\underline{a}) \mathcal{M}(\underline{b}) - \mathcal{M}(\underline{b}) \mathcal{M}(\underline{a}) = \\
&= \{\underline{b}\underline{a}^\top - \underline{a}\underline{b}^\top\} - \{\underline{a}\underline{b}^\top - \underline{b}\underline{a}^\top\} = \underline{b}\underline{a}^\top - \underline{a}\underline{b}^\top \\
\underline{a} \times (\underline{b} \times \underline{c}) &= (\underline{a} \cdot \underline{c})\underline{b} - (\underline{a} \cdot \underline{b})\underline{c} = \underline{b} [\underline{a}^\top \underline{c} - \underline{c} \underline{a}^\top] \\
(\underline{a} \cdot \underline{b}) &= \frac{1}{2} \text{Trace}[\mathcal{M}(\underline{a})^\top \mathcal{M}(\underline{b})] = -\frac{1}{2} \text{Trace}[\mathcal{M}(\underline{a}) \mathcal{M}(\underline{b})] \\
K\underline{a} \times K\underline{b} &= \det[K] K^{-\top} (\underline{a} \times \underline{b}), \quad K \in \mathbb{R}^{3 \times 3} \text{ invertible} \\
K\underline{a} &\longleftrightarrow \mathcal{M}(K\underline{a}) = \det[K] K^{-\top} \mathcal{M}(\underline{a}) K^{-1}
\end{aligned} \tag{1.1}$$

Note that the latter correspondence follows from  $\mathcal{M}(K\underline{a})\underline{x} = K\underline{a} \times \underline{x} = K\underline{a} \times K(K^{-1}\underline{x}) = \det[K] K^{-\top} (\underline{a} \times K^{-1}\underline{x}) = \det[K] K^{-\top} \mathcal{M}(\underline{a}) K^{-1} \underline{x}$ .

In  $\mathbb{R}^3$  we now consider  $N$  points situated at  $\underline{\xi}_1, \dots, \underline{\xi}_N$  and carrying mass  $m_i$ , respectively. Those mass points are connected by massless rigid rods so as to make them, as a whole, into a rigid system. We allow rigid *rotations* around  $\underline{0} \in \mathbb{R}^3$ . The general motion of  $\underline{\xi}_i$  as a function of time  $t$  is then given by  $t \mapsto \underline{x}_i(t) = \mathbf{R}(t)\underline{\xi}_i$ , with  $t \mapsto \mathbf{R}(t) \in \mathbf{SO}(3)$ , which is the manifold of all  $3 \times 3$ -matrices  $\mathbf{P}$ , with  $\mathbf{P}^\top = \mathbf{P}^{-1}$  and  $\det \mathbf{P} = +1$ . Keep in mind that the set  $\mathbf{SO}(3)$  establishes a smooth 3-dimensional surface in  $\mathbb{R}^{3 \times 3}$ .

Note that from  $\mathbf{R}^\top(t)\mathbf{R}(t) = I$  it follows that  $\mathbf{R}^\top(t)\dot{\mathbf{R}}(t) = -\dot{\mathbf{R}}^\top(t)\mathbf{R}(t)$ . As a consequence

$$\Omega(t) = \dot{\mathbf{R}}(t)\mathbf{R}(t)^\top = \dot{\mathbf{R}}(t)\mathbf{R}(t)^{-1} \tag{1.2}$$

and

$$\mathcal{A}(t) = \mathbf{R}^\top(t)\dot{\mathbf{R}}(t) = \mathbf{R}^{-1}\dot{\mathbf{R}}(t) = \mathbf{R}^\top(t)\Omega(t)\mathbf{R}(t) = \mathbf{R}^{-1}(t)\Omega(t)\mathbf{R}(t) \tag{1.3}$$

are anti-symmetric. From this we introduce two angular velocities and the correspondence between them.

$$\begin{aligned}
\text{The objective angular velocity } \underline{\omega} &\longleftrightarrow \Omega = \mathcal{M}(\underline{\omega}), \\
\text{The relative angular velocity } \underline{\alpha} &\longleftrightarrow \mathcal{A} = \mathcal{M}(\underline{\alpha}), \\
\underline{\alpha} &= \mathbf{R}^{-1}\underline{\omega}, \\
\dot{\underline{\alpha}} = \mathbf{R}^{-1}\dot{\underline{\omega}} &\longleftrightarrow \dot{\mathcal{A}} = \mathbf{R}^{-1}\dot{\Omega}\mathbf{R}.
\end{aligned} \tag{1.4}$$

The latter follows by differentiating (1.3) again and the observation  $\Omega\underline{\omega} = \underline{\omega} \times \underline{\omega} = \underline{0}$ .

We now derive an expression for the kinetic energy. The velocity of the  $i$ -th particle can be written  $\dot{\underline{x}}_i(t) = \dot{\mathbf{R}}(t)\underline{\xi}_i = \Omega(t)\mathbf{R}(t)\underline{\xi}_i = \underline{\omega}(t) \times \mathbf{R}(t)\underline{\xi}_i$ . We calculate the *angular momentum*

$$\begin{aligned} \underline{L} &= \sum_{j=1}^N m_j \underline{x}_j \times \dot{\underline{x}}_j = \sum_{j=1}^N m_j \mathbf{R}\underline{\xi}_j \times \dot{\mathbf{R}}\underline{\xi}_j = \sum_{j=1}^N m_j \mathbf{R}\underline{\xi}_j \times (\underline{\omega} \times \mathbf{R}\underline{\xi}_j) = \\ &= \mathbf{R} \left\{ \sum_{j=1}^N (m_j (\underline{\xi}_j^\top \underline{\xi}_j) I - (\underline{\xi}_j \underline{\xi}_j^\top)) \right\} \mathbf{R}^\top \underline{\omega} = \mathbf{R} J \mathbf{R}^\top \underline{\omega}, \end{aligned}$$

with  $J = \{ \cdot \}$ , the *inertia tensor* in the co-moving frame. Summarizing notations,

$$\underline{L}(t) = \mathbf{R}(t) J \mathbf{R}^\top(t) \underline{\omega}(t) = \mathbf{R}(t) J \underline{\alpha}(t) = J_{\mathbf{R}}(t) \underline{\omega}(t). \quad (1.5)$$

For the time derivative  $\dot{\underline{L}}$  we find

$$\dot{\underline{L}} = \sum_{j=1}^N m_j \dot{\underline{x}}_j \times \dot{\underline{x}}_j + \sum_{j=1}^N m_j \underline{x}_j \times \ddot{\underline{x}}_j = \sum_{j=1}^N \underline{x}_j \times (m_j \ddot{\underline{x}}_j) = \sum_{j=1}^N \underline{x}_j \times \underline{K}_j,$$

where  $\underline{K}_j$  is the (total) *external force* acting on the  $j$ -th particle. The *external torque* exerted on the rigid body is defined to be  $\underline{\tau} = \sum_{j=1}^N \underline{x}_j \times \underline{K}_j$ . The total *internal torque* is supposed to be zero. This always happens if  $\underline{\xi}_i$  and  $\underline{\xi}_j$  exert equal, but opposite, forces on each other in the direction  $\underline{\xi}_i - \underline{\xi}_j$ .

Rephrasing Newton's 2nd Law, the equation of motion becomes

$$\dot{\underline{L}}(t) = \underline{\tau}.$$

Differentiate the first identity in (1.5),

$$\begin{aligned} \dot{\underline{L}}(t) &= \dot{\mathbf{R}}(t) J \mathbf{R}^\top(t) \underline{\omega}(t) + \mathbf{R}(t) J \dot{\mathbf{R}}^\top(t) \underline{\omega}(t) + \mathbf{R}(t) J \mathbf{R}^\top(t) \dot{\underline{\omega}}(t) = \\ &= (\dot{\mathbf{R}}(t) \mathbf{R}^{-1}(t)) J_{\mathbf{R}}(t) \underline{\omega}(t) - J_{\mathbf{R}}(t) \Omega(t) \underline{\omega}(t) + J_{\mathbf{R}}(t) \dot{\underline{\omega}}(t) = \\ &= \underline{\omega} \times J_{\mathbf{R}} \underline{\omega} + J_{\mathbf{R}} \dot{\underline{\omega}} = \\ &= \mathbf{R} [\underline{\alpha}(t) \times J \underline{\alpha}(t) + J \dot{\underline{\alpha}}(t)] = \\ &= \underline{\tau}. \end{aligned} \quad (1.6)$$

If it happens that the torque  $\underline{\tau} = \underline{0}$  we find the famous Euler Equations

$$\underline{\alpha}(t) \times J \underline{\alpha}(t) + J \dot{\underline{\alpha}}(t) = \underline{0}. \quad (1.7)$$

Remind that  $\underline{\alpha}$  is the 'angular velocity in the co-moving frame'. Once, for given initial  $\underline{\alpha}(0)$ , the solution  $t \mapsto \underline{\alpha}(t)$  has been found and given the initial position  $\mathbf{R}(0)$ , the actual position  $t \mapsto \mathbf{R}(t)$  can be solved from

$$\dot{\mathbf{R}}(t) = \mathbf{R}(t) \mathcal{M}(\underline{\alpha}(t)) = \mathbf{R}(t) \mathcal{A}(t). \quad (1.8)$$



Note that, if  $\underline{\tau} = \underline{0}$  or co-moving, which means that  $\mathbf{R}^{-1}\underline{\tau}$  does not depend on  $\mathbf{R}$ , the differential equations for  $\underline{\alpha}$  and for  $\mathbf{R}$  are decoupled and both of 1st order. The equation (1.6) for  $\underline{\omega}$  is never decoupled from  $\mathbf{R}$ !

For later reference we rewrite (1.6) in matrix language, cf. (1.1),

$$\begin{aligned}\dot{\underline{\Omega}} + [J_{\mathbf{R}}, \underline{\Omega} J_{\mathbf{R}}^{-1} \underline{\Omega}] &= \dot{\underline{\Omega}} + J_{\mathbf{R}} \underline{\Omega} J_{\mathbf{R}}^{-1} \underline{\Omega} - \underline{\Omega} J_{\mathbf{R}}^{-1} \underline{\Omega} J_{\mathbf{R}} = \frac{1}{\det J} J_{\mathbf{R}} \mathcal{M}(\underline{\tau}) J_{\mathbf{R}}, \\ \dot{\underline{A}} + [J, \underline{A} J^{-1} \underline{A}] &= \dot{\underline{A}} + J \underline{A} J^{-1} \underline{A} - \underline{A} J^{-1} \underline{A} J = \frac{1}{\det J} J \mathbf{R}^{-1} \mathcal{M}(\underline{\tau}) \mathbf{R} J.\end{aligned}\quad (1.9)$$

The *kinetic energy* of the moving body is defined by

$$T = \sum_{j=1}^N m_j |\dot{\underline{x}}_j|^2 = \sum_{j=1}^N m_j (\underline{\omega} \times \mathbf{R} \underline{\xi}_j)^\top (\underline{\omega} \times \mathbf{R} \underline{\xi}_j),$$

which can be written ( use the identity  $|\underline{a} \times \underline{b}|^2 = (\underline{a}^\top \underline{a})(\underline{b}^\top \underline{b}) - (\underline{a}^\top \underline{b})^2$  )

$$\begin{aligned}T &= \frac{1}{2} \underline{\omega}^\top J_{\mathbf{R}} \underline{\omega} = \frac{1}{2} \underline{\alpha}^\top J \underline{\alpha} = \\ &= \frac{\det J}{4} \text{Trace}[\underline{\mathcal{A}}^\top J^{-1} \underline{\mathcal{A}} J^{-1}] = \frac{\det J}{4} \text{Trace}[(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J^{-1} (\mathbf{R}^{-1} \dot{\mathbf{R}}) J^{-1}].\end{aligned}\quad (1.10)$$

Inner multiplication of (1.7) by  $\underline{\alpha}$  shows that 'the kinetic energy  $T$  is conserved'. If we introduce an inner product of two skew symmetric matrices  $\underline{\mathcal{A}}$  and  $\underline{\mathcal{B}}$  by

$$\langle \underline{\mathcal{A}}, \underline{\mathcal{B}} \rangle = \frac{\det J}{4} \text{Trace}[(J^{-1} \underline{\mathcal{A}} J^{-1})^\top \underline{\mathcal{B}}], \quad (1.11)$$

it will be clear that the tangent vector (= 'tangent matrix')  $\dot{\mathbf{R}}(t)$  to the trajectory  $t \mapsto \mathbf{R}(t) \in \mathbf{SO}(3)$  has constant length with respect to this inner product if we take  $\underline{\mathcal{A}}(t) = \underline{\mathcal{B}}(t) = \mathbf{R}^{-1}(t) \dot{\mathbf{R}}(t)$ .

'Left invariance' of the inner product means that its value does not change if the trajectory  $t \mapsto \mathbf{R}(t)$  is replaced by  $t \mapsto \mathbf{Q} \mathbf{R}(t)$ , with  $\mathbf{Q}$  any fixed matrix in  $\mathbf{SO}(3)$ . Note there is *additional right invariance* only if  $J = I$ , the identity matrix. See section 2 for more details.

The trajectory  $t \mapsto \mathbf{R}(t)$  turns out to be a geodesic.

**Theorem 1.1** *Any solution  $t \mapsto \mathbf{R}(t)$  of (1.8), where  $\underline{\mathcal{A}}$  satisfies (1.9), with  $\underline{\tau} = \underline{0}$  describes a geodesic curve in  $\mathbf{SO}(3)$  provided with the left invariant inner product produced by (1.11). Thus, for the squared length of the tangent vector  $\dot{\mathbf{R}}(t)$  we write*

$$\left\langle \dot{\mathbf{R}}(t) \cdot \dot{\mathbf{R}}(t) \right\rangle_{\mathbf{R}(t)} = \frac{1}{2} \det[J] \text{Trace}[(J^{-1} \mathbf{R}^{-1}(t) \dot{\mathbf{R}}(t) J^{-1})^\top (\mathbf{R}^{-1}(t) \dot{\mathbf{R}}(t))] = \underline{\alpha}^\top(t) J \underline{\alpha}(t). \quad (1.12)$$

Here  $\underline{\alpha}$  is such that  $\mathcal{M}(\underline{\alpha}) = \mathbf{R}^{-1}(t) \dot{\mathbf{R}}(t)$ .

**Proof** Consider an arclength parametrized curve  $s \mapsto \mathbf{R}(s)$  of length  $L$  which connects two points in  $\mathbf{SO}(3)$ . We suppose arclength parametrization. That means

$$\forall s : \frac{1}{2} \det[J] \text{Trace}[(\mathbf{R}^{-1}\dot{\mathbf{R}})^\top J^{-1}(\mathbf{R}^{-1}\dot{\mathbf{R}})J^{-1}] = 1. \quad (1.13)$$

Calculate the length of the perturbed curve  $\mathbf{R}_{\varepsilon; \mathcal{H}} : s \mapsto e^{\varepsilon \mathcal{H}(s)} \mathbf{R}(s)$ , with any fixed  $s \mapsto \mathcal{H}(s) = -\mathcal{H}^\top(s)$ ,  $\mathcal{H}(0) = \mathcal{H}(L) = 0$ :

$$L(\varepsilon; \mathcal{H}) = \int_0^L \sqrt{\frac{1}{2} \det[J] \text{Trace}[(\mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}})^\top J^{-1}(\mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}})J^{-1}]} ds. \quad (1.14)$$

We first want to find the Frechet derivative  $\left. \frac{\partial}{\partial \varepsilon} L(\varepsilon; \mathcal{H}) \right|_{\varepsilon=0}$ .

For that purpose we use the expansions

$$\begin{aligned} \mathbf{R}_{\varepsilon; \mathcal{H}} &= e^{\varepsilon \mathcal{H}} \mathbf{R} = \mathbf{R} + \varepsilon \mathcal{H} \mathbf{R} + \mathcal{O}(\varepsilon^2), \\ \mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} &= \mathbf{R}^{-1} e^{-\varepsilon \mathcal{H}} = \mathbf{R}^{-1} - \varepsilon \mathbf{R}^{-1} \mathcal{H} + \mathcal{O}(\varepsilon^2), \\ \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}} &= \dot{\mathbf{R}} + \varepsilon \{ \mathcal{H} \dot{\mathbf{R}} + \dot{\mathcal{H}} \mathbf{R} \} + \mathcal{O}(\varepsilon^2), \\ \mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}} &= \mathbf{R}^{-1} \dot{\mathbf{R}} + \varepsilon \{ \mathbf{R}^{-1} \dot{\mathcal{H}} \mathbf{R} \} + \mathcal{O}(\varepsilon^2), \\ \left. \frac{\partial}{\partial \varepsilon} (\mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}}) \right|_{\varepsilon=0} &= \mathbf{R}^{-1} \dot{\mathcal{H}} \mathbf{R} + \mathcal{O}(\varepsilon), \end{aligned} \quad (1.15)$$

It becomes clear now that

$$\begin{aligned} \left. \frac{\partial}{\partial \varepsilon} \text{Trace}[(\mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}})^\top J^{-1}(\mathbf{R}_{\varepsilon; \mathcal{H}}^{-1} \dot{\mathbf{R}}_{\varepsilon; \mathcal{H}})J^{-1}] \right|_{\varepsilon=0} &= 2 \text{Trace}[(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J^{-1}(\mathbf{R}^{-1} \dot{\mathcal{H}} \mathbf{R})J^{-1}] = \\ &= 2 \text{Trace}[\{ \mathbf{R} J^{-1}(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J^{-1} \mathbf{R}^{-1} \} \dot{\mathcal{H}}] \end{aligned}$$

From this, (1.13)-(1.15) and partial integration, we find for all  $\mathcal{H}$

$$\left. \frac{\partial}{\partial \varepsilon} L(\varepsilon; \mathcal{H}) \right|_{\varepsilon=0} = -\det[J] \cdot \int_0^L \text{Trace} \left[ \frac{d}{ds} \{ \mathbf{R} J^{-1}(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J^{-1} \mathbf{R}^{-1} \} \mathcal{H} \right] ds \quad (1.16)$$

If the length of our curve is required to be stationary with respect to small perturbations (=geodesic curve), we have to impose

$$\left. \frac{\partial}{\partial \varepsilon} L(\varepsilon; \mathcal{H}) \right|_{\varepsilon=0} = 0, \quad \text{for all functions } \mathcal{H}. \quad (1.17)$$

Because of (1.16) this means  $\frac{d}{ds} \{ \mathbf{R} J^{-1}(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J^{-1} \mathbf{R}^{-1} \} = 0$ , or, after some arrangements

$$\frac{d}{ds} (\mathbf{R}^{-1} \dot{\mathbf{R}}) - J(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J^{-1}(\mathbf{R}^{-1} \dot{\mathbf{R}}) + (\mathbf{R}^{-1} \dot{\mathbf{R}}) J^{-1}(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top J = 0, \quad (1.18)$$

which corresponds to (1.9), taking into account  $(\mathbf{R}^{-1} \dot{\mathbf{R}})^\top = -(\mathbf{R}^{-1} \dot{\mathbf{R}})$ . ■

**Remarks**

- If the square root  $\sqrt{\cdot}$  in (1.14) is replaced by any smooth strictly monotone increasing function  $\Theta : (0, \infty) \rightarrow (0, \infty)$ , one still arrives at (1.18).
- The interpretation of 'conservation of kinetic energy' typically belongs to the left invariant inner product. The right invariant inner product does not admit such interpretation!
- In the Appendix we deal with the motion of a rigid body acted on by an external torque. Such as a spinning top or the globe!

## 2 A special class of Matrix Lie Groups

For  $\Delta \in \mathbb{R}^{n \times n}$  with  $\det \Delta \neq 0$ , consider the Matrix Lie-Group

$$\mathbf{G}(n, \mathbb{R}, \Delta) = \{ P \in \mathbb{R}^{n \times n} \mid P^\top \Delta P = \Delta \}, \quad (2.1)$$

and the corresponding Matrix Lie-Algebra

$$\mathfrak{g}(n, \mathbb{R}, \Delta) = \{ \mathcal{A} \in \mathbb{R}^{n \times n} \mid \mathcal{A}^\top \Delta + \Delta \mathcal{A} = O \}, \quad (2.2)$$

**Notation:**

- Matrices belonging to  $\mathbf{G}(n, \mathbb{R}, \Delta)$  are denoted by capitals of type  $P, F, X, Y, \dots$
- Matrices belonging to  $\mathfrak{g}(n, \mathbb{R}, \Delta)$  are rendered by capitals of type  $\mathcal{A}, \mathcal{B}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}, \dots$
- Non specified matrices are rendered by ordinary capitals  $P, A, R, J, \dots$

**Note 1:** If  $\mathcal{A} \in \mathfrak{g}(n, \mathbb{R}, \Delta)$  it follows that for all  $t \mapsto e^{t\mathcal{A}}$  is a curve in  $\mathbf{G}(n, \mathbb{R}, \Delta)$ . To see this show that  $\frac{d}{dt}[e^{t\mathcal{A}\top} \Delta e^{t\mathcal{A}} - \Delta] = 0$ .

**Note 2:** Let  $\mathfrak{d} = \dim \mathfrak{g}(n, \mathbb{R}, \Delta)$  and let  $\{\mathcal{A}_1, \dots, \mathcal{A}_{\mathfrak{d}}\}$  be a basis of  $\mathfrak{g}(n, \mathbb{R}, \Delta)$ .

A straightforward application of the Implicit Function Theorem leads to the fact that  $\mathbf{G}(n, \mathbb{R}, \Delta)$  is, near  $P = I$ , a smooth surface of dimension  $\mathfrak{d}$ . Indeed, the derivative of  $Q \mapsto Q^\top \Delta Q - \Delta$  is, at  $Q = I$ , given by the linear mapping  $H \mapsto H^\top \Delta + \Delta H$ . The latter mapping has rank  $n^2 - \mathfrak{d}$ .

Local charts, near  $I$ , for  $\mathbf{G}(n, \mathbb{R}, \Delta)$  are given by

$$\{t_1, \dots, t_{\mathfrak{d}}\} \mapsto e^{t_1 \mathcal{A}_1} \dots e^{t_{\mathfrak{d}} \mathcal{A}_{\mathfrak{d}}}, \quad \text{and also} \quad \{t_1, \dots, t_{\mathfrak{d}}\} \mapsto e^{t_1 \mathcal{A}_1 + \dots + t_{\mathfrak{d}} \mathcal{A}_{\mathfrak{d}}}, \quad (2.3)$$

at least for  $\{t_1, \dots, t_{\mathfrak{d}}\}$  near  $\{0, \dots, 0\} \in \mathbb{R}^{\mathfrak{d}}$ .

By means of the group multiplication local charts near any  $P \in \mathbf{G}(n, \mathbb{R}, \Delta)$  are produced by

$$\{t_1, \dots, t_{\mathfrak{d}}\} \mapsto P e^{t_1 \mathcal{A}_1} \dots e^{t_{\mathfrak{d}} \mathcal{A}_{\mathfrak{d}}}, \quad \{t_1, \dots, t_{\mathfrak{d}}\} \mapsto e^{t_1 \mathcal{A}_1 + \dots + t_{\mathfrak{d}} \mathcal{A}_{\mathfrak{d}}} P, \quad \text{etc.} \quad (2.4)$$

Thus an atlas can be provided for the whole of  $\mathbf{G}(n, \mathbb{R}, \Delta)$ .

**Note 3:** From taking the determinant of  $\mathbf{P}^\top \Delta \mathbf{P} = \Delta$  it follows that

$$|\det \mathbf{P}| = 1, \text{ for all } \mathbf{P} \in \mathbf{G}(n, \mathbb{R}, \Delta).$$

From taking the trace on both sides of  $\Delta^{-1} \mathcal{A}^\top \Delta = -\mathcal{A}$ , it follows that

$$\text{Trace} \mathcal{A} = 0, \text{ for all } \mathcal{A} \in \mathfrak{g}(n, \mathbb{R}, \Delta).$$

**Note 4:** Suppose  $\Delta^2 = \gamma I$ ,  $\gamma \in \mathbb{R}$ . Multiply  $\mathcal{A}^\top \Delta + \Delta \mathcal{A} = O$  on both sides by  $\Delta$ , then it follows:

$$\mathcal{A} \in \mathfrak{g}(n, \mathbb{R}, \Delta) \Rightarrow \mathcal{A}^\top \in \mathfrak{g}(n, \mathbb{R}, \Delta),$$

and a fortiori:

$$\mathbf{P} \in \mathbf{G}(n, \mathbb{R}, \Delta) \Rightarrow \mathbf{P}^\top \in \mathbf{G}(n, \mathbb{R}, \Delta).$$

**Note 5:** If  $\Delta = I$  one denotes  $\mathbf{G}(n, \mathbb{R}, I) = \mathbf{O}(n)$ .

Its *subgroup* with  $\det \mathbf{P} = 1$  is denoted by  $\mathbf{SO}(n)$ .

If  $\Delta = \text{diag}[\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_s]$ , one denotes  $\mathbf{G}(n, \mathbb{R}, \Delta) = \mathbf{O}(r, s)$ .

Its *subgroup* with  $\det \mathbf{P} = 1$  is denoted by  $\mathbf{SO}(r, s)$ .

**Note 6:** If  $\Delta \in \mathbb{R}^{2n \times 2n}$  and  $\Delta = \begin{bmatrix} O & I \\ -I & O \end{bmatrix}$ , one gets the *symplectic group*  $\mathbf{Sp}(2n)$ .

**Note 7:** If we multiply  $\mathcal{A}^\top \Delta + \Delta \mathcal{A} = O$  from the left with  $\mathbf{P}^\top$  and from the right with  $\mathbf{P}$ , we get the identity  $(\mathbf{P}^\top \mathcal{A}^\top \mathbf{P}^{-\top}) \mathbf{P}^\top \Delta \mathbf{P} + \mathbf{P}^\top \Delta \mathbf{P} (\mathbf{P}^{-1} \mathcal{A} \mathbf{P}) = O$ . This can be written

$$(\mathbf{P}^{-1} \mathcal{A} \mathbf{P})^\top \Delta + \Delta (\mathbf{P}^{-1} \mathcal{A} \mathbf{P}) = O. \quad (2.5)$$

It follows that for all  $\mathbf{Q} \in \mathbf{G}(n, \mathbb{R}, \Delta)$ , the mapping  $\text{Ad}_{\mathbf{Q}}$

$$\mathcal{A} \mapsto \text{Ad}_{\mathbf{Q}} \mathcal{A} = \mathbf{Q} \mathcal{A} \mathbf{Q}^{-1}, \quad (2.6)$$

is a linear mapping and a Lie-Algebra homomorphism in  $\mathfrak{g}(n, \mathbb{R}, \Delta)$ .

Let  $t \mapsto \mathbf{F}(t) \in \mathbf{G}(n, \mathbb{R}, \Delta)$ , with  $t \in \mathbb{R}$  taken from some open interval which contains  $t = 0$ . From differentiation of  $\mathbf{F}^\top \Delta \mathbf{F} = \Delta$ , we learn  $\dot{\mathbf{F}}^\top \Delta \mathbf{F} + \mathbf{F}^\top \Delta \dot{\mathbf{F}} = O$  and hence both  $(\dot{\mathbf{F}} \mathbf{F}^{-1})^\top \Delta + \Delta (\dot{\mathbf{F}} \mathbf{F}^{-1}) = O$  and  $(\mathbf{F}^{-1} \dot{\mathbf{F}})^\top \Delta + \Delta (\mathbf{F}^{-1} \dot{\mathbf{F}}) = O$ .

Therefore both  $t \mapsto \dot{\mathbf{F}}(t) \mathbf{F}^{-1}(t)$  and  $t \mapsto \mathbf{F}^{-1}(t) \dot{\mathbf{F}}(t)$  are curves in  $\mathfrak{g}(n, \mathbb{R}, \Delta)$ .

For the moment we describe smooth local curves through  $\mathbf{P} \in \mathbf{G}(n, \mathbb{R}, \Delta)$  by  $t \mapsto \mathbf{F}_{\mathbf{P}}(t)$ , with  $\mathbf{F}_{\mathbf{P}}(0) = \mathbf{P}$ . If  $\mathbf{P} = I$ , we omit  $I$  and write  $t \mapsto \mathbf{X}_I(t) = \mathbf{X}(t)$ . With  $\mathbf{X}, \mathbf{Y}, \dots$  instead of  $\mathbf{P}, \mathbf{Q}, \dots$ .

**Examples:**  $t \mapsto \mathbf{F}_{\mathbf{P}}(t) = \mathbf{P} e^{t\mathcal{A}}$ , with  $\mathcal{A} \in \mathfrak{g}(n, \mathbb{R}, \Delta)$ , and  $t \mapsto \mathbf{G}_{\mathbf{P}}(t) = e^{t\mathcal{A}} \mathbf{P}$ .

Next, two bilinear forms on the tangent spaces of  $\mathbf{G}(n, \mathbb{R}, \Delta)$  will be introduced.

### 1. A right invariant inner product

Let  $J \in \mathbb{R}^{n \times n}$ , with  $\det J \neq 0$ . Define the inner product of the tangent vectors  $\dot{F}_P, \dot{G}_P$  at  $P$  by

$$(\dot{F}_P \cdot \dot{G}_P)_P = \text{Trace}[\Delta^{-1} \dot{F}_P^\top(0) J \dot{G}_P(0)]. \quad (2.7)$$

This inner product is right invariant: If we replace  $t \mapsto F_P(t)$  by  $t \mapsto F_{PQ}(t) = F_P(t)Q$ , and  $t \mapsto G_P(t)$  by  $t \mapsto G_{PQ}(t) = G_P(t)Q$ , we find

$$(\dot{F}_{PQ} \cdot \dot{G}_{PQ})_{PQ} = \text{Trace}[\Delta^{-1} (\dot{F}_P(0)Q)^\top J \dot{G}_P(0)Q] = \text{Trace}[Q \Delta^{-1} Q^\top \dot{F}_P^\top(0) J \dot{G}_P(0)] = (\dot{F}_P \cdot \dot{G}_P)_P.$$

If we take in particular  $Q = P^{-1}$ , then  $\dot{F}_P P^{-1}$  and  $\dot{G}_P P^{-1}$  are elements from  $\mathfrak{g}(n, \mathbb{R}, \Delta)$  and

$$(\dot{F}_P \cdot \dot{G}_P)_P = (\dot{F}_P P^{-1} \cdot \dot{G}_P P^{-1})_I. \quad (2.8)$$

If we use (2.8) as a definition we get right invariant inner products starting from **any** inner product  $(\cdot, \cdot)_I$  on  $\mathfrak{g}(n, \mathbb{R}, \Delta)$ .

The inner product (2.7) is symmetric if  $J$  and  $\Delta$  are both symmetric or both anti symmetric. The inner product (2.7) is anti symmetric if  $J$  is symmetric and  $\Delta$  is anti symmetric or vice versa. In the important special case  $\Delta = I$  and  $J = J^\top > 0$ , the orthogonal group, the inner product is non-degenerate and positive.

### 2. A left invariant inner product

Let again  $J \in \mathbb{R}^{n \times n}$ , with  $\det J \neq 0$ . Define the inner product of the tangent vectors  $\dot{F}_P(0), \dot{G}_P(0)$  at  $P$  by

$$\left\langle \dot{F}_P \cdot \dot{G}_P \right\rangle_P = \text{Trace}[\Delta \dot{F}_P(0) J \dot{G}_P^\top(0)]. \quad (2.9)$$

This inner product is left invariant. If we replace  $t \mapsto F_P(t)$  by  $t \mapsto F_{QP}(t) = QF_P(t)$ , and  $t \mapsto G_P(t)$  by  $t \mapsto G_{QP}(t) = QG_P(t)$ , we find

$$\left\langle \dot{F}_{QP} \cdot \dot{G}_{QP} \right\rangle_{QP} = \text{Trace}[\Delta Q \dot{F}_P(0) J (Q \dot{G}_P(0))^\top] = \text{Trace}[Q^\top \Delta Q \dot{F}_P(0) J \dot{G}_P^\top(0)] = \left\langle \dot{F}_P \cdot \dot{G}_P \right\rangle_P.$$

If we take in particular  $Q = P^{-1}$ , then  $\dot{F}_P P^{-1}$  and  $\dot{G}_P P^{-1}$  are elements from  $\mathfrak{g}(n, \mathbb{R}, \Delta)$  and

$$\left\langle \dot{F}_P \cdot \dot{G}_P \right\rangle_P = \left\langle P^{-1} \dot{F}_P \cdot P^{-1} \dot{G}_P \right\rangle_I. \quad (2.10)$$

If we use (2.10) as a definition we get left invariant inner products starting from any inner product  $\left\langle \cdot, \cdot \right\rangle_I$  on  $\mathfrak{g}(n, \mathbb{R}, \Delta)$ .

The inner product (2.9) can be written

$$\left\langle \dot{F}_P \cdot \dot{G}_P \right\rangle_P = \text{Trace}[\Delta (\dot{F}_P(0)P^{-1}) J_P (\dot{G}_P(0)P^{-1})^\top(0)], \quad \text{with} \quad J_P = P J P^\top. \quad (2.11)$$

### 3. Left/right invariant inner products

First we compare the inner products (2.7) and (2.9) at I. Because of  $\mathcal{A}^\top = -\Delta\mathcal{A}\Delta^{-1}$ ,

$$(\mathcal{A} \cdot \mathcal{B})_I = \text{Trace}[\Delta^{-1}\mathcal{A}^\top J\mathcal{B}] = -\text{Trace}[\Delta^{-1}\Delta\mathcal{A}\Delta^{-1}J\mathcal{B}] = -\text{Trace}[\mathcal{A}\Delta^{-1}J\mathcal{B}],$$

and

$$\langle \mathcal{A} \cdot \mathcal{B} \rangle_I = \text{Trace}[\Delta\mathcal{A}J\mathcal{B}^\top] = -\text{Trace}[\Delta\mathcal{A}J\Delta\mathcal{B}\Delta^{-1}] = -\text{Trace}[\mathcal{A}J\Delta\mathcal{B}].$$

Both inner products are the same at I iff  $J$  satisfies  $J\Delta = \Delta^{-1}J$ .

Next, we observe that the right invariant inner product (2.7) is also left invariant if

$$\forall \mathbf{Q} \in \mathbf{G}(n, \mathbb{R}, \Delta) : \quad \text{Trace}[\Delta^{-1}\dot{\mathbf{F}}_P^\top(0) \mathbf{Q}^\top J \mathbf{Q} \dot{\mathbf{G}}_P(0)] = \text{Trace}[\Delta^{-1}\dot{\mathbf{F}}_P^\top(0) J \dot{\mathbf{G}}_P(0)].$$

This happens to be the case if  $J = \Delta$  is taken.

Finally, the left invariant invariant inner product (2.9) is also right invariant if

$$\forall \mathbf{Q} \in \mathbf{G}(n, \mathbb{R}, \Delta) : \quad \text{Trace}[\Delta\dot{\mathbf{F}}_P(0) \mathbf{Q} J \mathbf{Q}^\top \dot{\mathbf{G}}_P^\top(0)] = \text{Trace}[\Delta\dot{\mathbf{F}}_P(0) J \dot{\mathbf{G}}_P^\top(0)]$$

This happens to be the case if  $J = \Delta^{-1}$  is taken. Similarly, the right invariant invariant inner product (2.7) is also left invariant if  $J = \Delta$  is taken.

We conclude that, if  $J = \Delta = \Delta^{-1}$  both inner products (2.7) and (2.9) are the same *and* left/right invariant.

If we start from an inner product  $\langle \cdot, \cdot \rangle_I$  on the Lie Algebra  $\mathfrak{g}(n, \mathbb{R}, \Delta)$ , with the property

$$\forall \mathcal{A}, \mathcal{B} \in \mathfrak{g}(n, \mathbb{R}, \Delta) \quad \forall \mathbf{Q} \in \mathbf{G}(n, \mathbb{R}, \Delta) :$$

$$\langle \mathcal{A}, \mathcal{B} \rangle_I = \langle \text{Ad}_Q \mathcal{A}, \text{Ad}_Q \mathcal{B} \rangle_I = \langle \mathbf{Q}\mathcal{A}\mathbf{Q}^{-1}, \mathbf{Q}\mathcal{B}\mathbf{Q}^{-1} \rangle_I, \quad (2.12)$$

then the definitions (2.8) and (2.10) lead to the same left/right invariant inner product.

As will be shown in the next section, examples of such inner products (bilinear forms) are

$$\langle \mathcal{A}, \mathcal{B} \rangle_I = \lambda \text{Trace}[\mathcal{A}\mathcal{B}] + \mu \text{Trace}[\mathcal{A}]\text{Trace}[\mathcal{B}], \quad \text{with } \lambda, \mu \in \mathbb{R}. \quad (2.13)$$

They have the property (2.12) for *all* matrix groups.

### 3 Traces and Bilinear Forms on Spaces of Matrices

For a fixed  $A = [a_{ij}] \in \mathbb{C}^{n \times n}$  we consider the linear map

$$\text{ad}_A : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n} \quad : \quad X \mapsto \text{ad}_A X = AX - XA. \quad (3.1)$$

We want to express  $\text{Trace}[\text{ad}_B \text{ad}_A]$  in the traces of  $A$ ,  $B$ ,  $BA$ .

For this purpose we identify  $\mathbb{C}^{n \times n}$  and  $\mathbb{C}^{n^2}$  by constructing one long column from the successive columns of  $X$ :

$$X = [x_{ij}] \quad \longleftrightarrow \quad \text{column}[x_{11}, x_{21}, \dots, x_{n1}, \dots, \dots, x_{1n}, x_{2n}, \dots, x_{nn}]. \quad (3.2)$$

This column will again be denoted by  $X$ .

The linear mapping  $X \mapsto AX$  is represented by the  $n^2 \times n^2$ -matrix

$$\mathcal{A}_L = \begin{bmatrix} A & O & \cdots & O \\ O & A & \cdots & O \\ O & O & \cdots & O \\ \vdots & \vdots & \vdots & \vdots \\ O & O & \cdots & A \end{bmatrix}, \quad (3.3)$$

with the  $n \times n$  diagonal matrix  $A$  repeated on its 'diagonal' and  $O$  the zero matrix. Further, the linear mapping  $X \mapsto XA$  is represented by the  $n^2 \times n^2$ -matrix

$$\mathcal{A}_R = \begin{bmatrix} a_{11}I & a_{21}I & \cdots & a_{n1}I \\ a_{12}I & a_{22}I & \cdots & a_{n2}I \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n}I & a_{2n}I & \cdots & a_{nn}I \end{bmatrix}. \quad (3.4)$$

Note that  $\mathcal{A}_R$  appears to be subdivided in  $n \times n$ -matrices which are all multiples of the identity matrix  $I$ . Note also that the 'multipliers' appear in the order of the transposed  $A^\top$  of  $A$ .

We have  $[\text{ad}_B \text{ad}_A]X = BAX - BXA - AXB + XAB$ . From this we find for the  $n^2 \times n^2$ -matrix representation of  $[\text{ad}_B \text{ad}_A]$

$$\mathcal{B}_L \mathcal{A}_L - \mathcal{B}_L \mathcal{A}_R - \mathcal{A}_L \mathcal{B}_R + \mathcal{B}_R \mathcal{A}_R. \quad (3.5)$$

The respective 'block diagonals', built from  $n \times n$ -matrices, are

$$\begin{aligned} \mathcal{B}_L \mathcal{A}_L &\supset \text{diag}[BA, BA, \dots, BA] \\ \mathcal{B}_L \mathcal{A}_R &\supset \text{diag}[a_{11}B, a_{22}B, \dots, a_{nn}B] \\ \mathcal{A}_L \mathcal{B}_R &\supset \text{diag}[b_{11}A, b_{22}A, \dots, b_{nn}A] \\ \mathcal{B}_R \mathcal{A}_R &\supset \text{diag}[[B^\top A^\top]_{11}I, [B^\top A^\top]_{22}I, \dots, [B^\top A^\top]_{nn}I] \end{aligned} \quad (3.6)$$

This enables us to calculate the traces of the matrices in (3.5). This leads to

$$\text{Trace}[\text{ad}_A \text{ad}_B] = 2n\text{Trace}[AB] - 2\text{Trace}[A] \text{Trace}[B]. \quad (3.7)$$

**Note 1.** If  $A, B \in \mathfrak{g}(n, \mathbb{R}, \Delta)$ , cf (2.1), then  $\text{Trace}[A] = \text{Trace}[B] = 0$ , since from  $A^\top + \Delta A \Delta^{-1} = O$ , it follows that  $2\text{Trace}[A] = 0$ .

**Note 2.** If  $A, B \in \mathfrak{g}(n, \mathbb{R}, \Delta)$ , with  $\Delta$  symmetric, then because of  $A = -\Delta^{-1}A^\top\Delta$ , we have

$$\text{Trace}[\text{ad}_A \text{ad}_B] = 2n\text{Trace}[AB] = -2n\text{Trace}[\Delta A \Delta^{-1} B^\top] = -2n\text{Trace}[\Delta^{-1} A^\top \Delta B]. \quad (3.8)$$

**Note 3.** The inner product  $\text{Trace}[AB]$  on any matrix Lie Algebra leads to both a left and right invariant inner product on the corresponding matrix Lie group.

## 4 General Matrix Lie-Groups

Let  $\mathbf{GL}(\mathbb{R}, n)$  denote the Group of all *invertible* matrices in  $\mathbb{R}^{n \times n}$ .

A subgroup  $\mathbf{G}_M \subset \mathbf{GL}(\mathbb{R}, n)$  is defined to be a *Matrix Lie Group* if the intersection of  $\mathbf{G}_M$  with a sufficiently small neighbourhood of the identity  $I \in \mathbb{R}^{n \times n}$  is a  $\mathfrak{d}$ -dimensional smooth surface for some  $\mathfrak{d} \leq n^2$ . Starting from a chart of the neighbourhood of  $I$  it is then straightforward to construct an atlas for  $\mathbf{G}_M$  as a whole.

**Notation:**

- Denote the tangent space of the chosen  $\mathbf{G}_M$  at  $I$  by  $\mathfrak{g}_M$ .
- Matrices belonging to the Lie group  $\mathbf{G}_M$  under consideration are denoted by capitals of type  $P, F, X, \dots$
- Matrices belonging to the corresponding 'Lie Algebra'  $\mathfrak{g}_M$  are rendered by capitals of type  $\mathcal{A}, \mathcal{X}, \mathcal{B}, \dots$
- Non specified matrices are rendered by ordinary capitals  $P, A, R, \dots$
- If we want to emphasize that a smooth curve  $t \mapsto F(t)$  passes through  $P$  at  $t = 0$ , we write  $t \mapsto F(P; t)$  instead.

Not all Matrix Lie Groups are of type (2.1)!

**Examples:**

- Full  $\mathbf{GL}(\mathbb{R}, n)$ .
- The set all Upper Triangular Matrices.
- Affine groups: Any set of matrices of type

$$\begin{bmatrix} \mathbf{G} & \underline{x} \\ \underline{0}^\top & 1 \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}, \quad \underline{x} \in \mathbb{R}^n, \mathbf{G} \in \mathbf{G}_M.$$

For the latter  $\mathbf{G}_M$  one can take any fixed matrix Lie Group situated in  $\mathbf{GL}(\mathbb{R}, n) \subset \mathbb{R}^{n \times n}$ .



Observe that  $P^{-1}\dot{G}(P; 0)$  and  $\dot{G}(P; 0)P^{-1}$  both belong to the linear space  $\mathfrak{g}_{\mathcal{M}}$  since  $t \mapsto P^{-1}G(P; t)$  and  $t \mapsto G(P; t)P^{-1}$  are curves in  $\mathbf{G}_{\mathcal{M}}$  passing through  $I$  at  $t = 0$ .

Let  $t \mapsto X(t)$  and  $t \mapsto Y(t)$  be 'local curves' in  $\mathbf{G}_{\mathcal{M}}$ , with  $X(0) = Y(0) = I$ . Denote their respective tangent matrices at  $I$  by  $\dot{X}(0)$  and  $\dot{Y}(0)$ .

**Theorem 4.1** *Let  $t \mapsto X(t)$ ,  $-T < t < T$  be a differentiable curve in  $\mathbf{G}_{\mathcal{M}}$ , with  $X(0) = I$ . Then on some interval  $[-T_1, T_1]$  with  $T_1 > 0$  we have*

$$\lim_{n \rightarrow \infty} \left\{ X\left(\frac{t}{n}\right) \right\}^n = e^{tX'(0)}, \quad (4.1)$$

uniformly on  $[-T_1, T_1]$ . Therefore  $t \mapsto e^{tX'(0)}$  is a 1-parameter subgroup in  $\mathbf{G}_{\mathcal{M}}$ .

**Proof** Put  $X'(0) = \mathcal{A}$  and write  $X(t) = I + t\mathcal{A} + E(t)$ , with  $\frac{E(t)}{t} \rightarrow 0$  as  $t \rightarrow 0$ . We employ the 'operator norm' on  $\mathbb{R}^{n \times n}$  with property  $\|AB\| \leq \|A\| \|B\|$ . Denote  $\varepsilon_{t,n} = \frac{n}{t} \|E(\frac{t}{n})\|$ . Beware of non-commuting matrices and write

$$\begin{aligned} \left\{ X\left(\frac{t}{n}\right) \right\}^n - \left\{ I + \frac{t}{n} \mathcal{A} \right\}^n &= \left\{ I + \frac{t}{n} \mathcal{A} + E\left(\frac{t}{n}\right) \right\}^n - \left\{ I + \frac{t}{n} \mathcal{A} \right\}^n = \\ &= I + n \left\{ \frac{t}{n} \mathcal{A} + E\left(\frac{t}{n}\right) \right\} - \left\{ I + n \frac{t}{n} \mathcal{A} \right\} + \sum_{k=2}^n \binom{n}{k} \left(\frac{t}{n}\right)^k \left\{ \left( \mathcal{A} + \frac{n}{t} E\left(\frac{t}{n}\right) \right)^k - \mathcal{A}^k \right\}. \end{aligned} \quad (4.2)$$

The norm of the first 3 terms is smaller than  $t\varepsilon_{t,n}$ . Choose  $N \in \mathbb{N}$  such that for  $n > tN$  one has  $\varepsilon_{t,n} \leq 1$ . Then

$$\left\| \left( \mathcal{A} + \frac{n}{t} E\left(\frac{t}{n}\right) \right)^k - \mathcal{A}^k \right\| \leq \left\| \frac{n}{t} E\left(\frac{t}{n}\right) \right\| \sum_{\ell=1}^k \binom{k}{\ell} \|\mathcal{A}\|^{k-\ell} \left\| \frac{n}{t} E\left(\frac{t}{n}\right) \right\|^{\ell-1} \leq \varepsilon_{t,n} (\|\mathcal{A}\| + 1)^k$$

The norm of the last term in (4.2) is smaller than

$$\varepsilon_{t,n} \sum_{k=2}^n \binom{n}{k} \left(\frac{t}{n}\right)^k (\|\mathcal{A}\| + 1)^k \leq \varepsilon_{t,n} \left( \frac{t(\|\mathcal{A}\| + 1)}{n} + 1 \right)^n.$$

For  $n \rightarrow \infty$  we have  $\varepsilon_{t,n} \rightarrow 0$  and  $\left( \frac{t(\|\mathcal{A}\| + 1)}{n} + 1 \right)^n \rightarrow e^{t(\|\mathcal{A}\| + 1)}$ . ■

**Theorem 4.2** *The tangent space  $\mathfrak{g}_{\mathcal{M}}$  is a Lie Algebra with respect to the Lie product  $[\mathcal{X}, \mathcal{Y}] = \mathcal{X}\mathcal{Y} - \mathcal{Y}\mathcal{X}$ . We have the linear mappings  $\text{ad}$  and  $\text{Ad}$ :*

- $\forall \mathcal{X}, \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}} : \mathcal{Y} \mapsto \text{ad}_{\mathcal{X}} \mathcal{Y} = \mathcal{X}\mathcal{Y} - \mathcal{Y}\mathcal{X} \in \mathfrak{g}_{\mathcal{M}}$ .
- $\forall P \in \mathbf{G}_{\mathcal{M}} \forall \mathcal{X} \in \mathfrak{g}_{\mathcal{M}} : \mathcal{X} \mapsto \text{Ad}_P \mathcal{X} = P\mathcal{X}P^{-1} \in \mathfrak{g}_{\mathcal{M}}$ .

**Proof**

- Differentiation of  $t \mapsto X(t)X^{-1}(t) = I$  leads to  $(X^{-1})'(t) = -X^{-1}(t)\dot{X}(t)X^{-1}(t)$ . Evaluation at  $t = 0$  gives us  $(X^{-1})'(0) = -\dot{X}(0)$ .
  - Let  $P \in \mathbf{G}_{\mathcal{M}}$  be fixed. The tangent vector at  $t = 0$  to the curve  $t \mapsto PX(t)P^{-1}$ , with  $X(0) = I$  is given by  $P\dot{X}(0)P^{-1}$ . So for any  $P \in \mathbf{G}_{\mathcal{M}}$  mapping  $\mathcal{A} \mapsto \text{Ad}_P \mathcal{A} = P\mathcal{A}P^{-1}$  is a linear mapping in  $\mathfrak{g}_{\mathcal{M}}$ .
  - It now follows that for  $\mathcal{A} \in \mathfrak{g}_{\mathcal{M}}$  the curve of matrices  $t \mapsto X(t)\mathcal{A}X^{-1}(t)$  is a curve in  $\mathfrak{g}_{\mathcal{M}}$ . The tangent to this curve  $t \mapsto \dot{X}(t)\mathcal{A}X^{-1}(t) + X(t)\mathcal{A}(X^{-1})'(t)$  is, for each  $t$ , a vector in  $\mathfrak{g}_{\mathcal{M}}$ . For  $t = 0$  we find  $\dot{X}(0)\mathcal{A} - \mathcal{A}\dot{X}(0) \in \mathfrak{g}_{\mathcal{M}}$ . Take  $X(t) = e^{t\mathcal{B}}$ .
- The conclusion is that for any  $\mathcal{B}, \mathcal{A} \in \mathfrak{g}_{\mathcal{M}}$  the matrix  $\text{ad}_{\mathcal{B}}\mathcal{A} = \mathcal{B}\mathcal{A} - \mathcal{A}\mathcal{B} \in \mathfrak{g}_{\mathcal{M}}$ . ■

**Note:** Let  $\mathfrak{d} = \dim \mathfrak{g}_{\mathcal{M}}$  and let  $\{\mathcal{A}_1, \dots, \mathcal{A}_{\mathfrak{d}}\}$  be a basis of  $\mathfrak{g}_{\mathcal{M}}$ . Local charts (parametrizations), near  $I$ , for  $\mathbf{G}_{\mathcal{M}}$  are given by

$$\{t_1, \dots, t_{\mathfrak{d}}\} \mapsto e^{t_1\mathcal{A}_1} \dots e^{t_{\mathfrak{d}}\mathcal{A}_{\mathfrak{d}}}, \quad \text{and also} \quad \{t_1, \dots, t_{\mathfrak{d}}\} \mapsto e^{t_1\mathcal{A}_1 + \dots + t_{\mathfrak{d}}\mathcal{A}_{\mathfrak{d}}}, \quad (4.3)$$

at least for  $\{t_1, \dots, t_{\mathfrak{d}}\}$  near  $\{0, \dots, 0\} \in \mathbb{R}^{\mathfrak{d}}$ .

Local charts (parametrizations), near  $P \in \mathbf{G}_{\mathcal{M}}$  are given by

$$\{t_1, \dots, t_{\mathfrak{d}}\} \mapsto P e^{t_1\mathcal{A}_1} \dots e^{t_{\mathfrak{d}}\mathcal{A}_{\mathfrak{d}}}, \quad \text{and also} \quad \{t_1, \dots, t_{\mathfrak{d}}\} \mapsto e^{t_1\mathcal{A}_1 + \dots + t_{\mathfrak{d}}\mathcal{A}_{\mathfrak{d}}} P, \dots \quad (4.4)$$

The tangent plane (or tangent space) at  $P \in \mathbf{G}_{\mathcal{M}}$  is spanned by the matrices  $\{P\mathcal{A}_1, \dots, P\mathcal{A}_{\mathfrak{d}}\}$  and also by  $\{\mathcal{A}_1P, \dots, \mathcal{A}_{\mathfrak{d}}P\}$ . The tangent space at  $P \in \mathbf{G}_{\mathcal{M}}$  will be denoted by  $P\mathfrak{g}_{\mathcal{M}}$ .

## 5 Left and Right Geodesics on Matrix Lie Groups

We consider a general matrix Lie-Group  $\mathbf{G}_{\mathcal{M}} \subset \mathbf{GL}(\mathbb{R}, n)$ . The tangent space (= the Lie -Algebra) of  $\mathbf{G}_{\mathcal{M}} \subset \mathbf{GL}(\mathbb{R}, n)$  at the identity matrix  $\mathbf{I}$  is denoted by  $\mathfrak{g}_{\mathcal{M}}$ . On  $\mathfrak{g}_{\mathcal{M}}$  we choose a symmetric, non-degenerate inner product  $\langle \cdot, \cdot \rangle_{\mathbf{I}}$ .

As before we introduce the left invariant inner product

$$\langle \dot{\mathbf{F}}_{\mathbf{P}} \cdot \dot{\mathbf{G}}_{\mathbf{P}} \rangle_{\mathbf{P}} = \langle \mathbf{P}^{-1} \dot{\mathbf{F}}_{\mathbf{P}} \cdot \mathbf{P}^{-1} \dot{\mathbf{G}}_{\mathbf{P}} \rangle_{\mathbf{I}},$$

and the right invariant inner product

$$(\dot{\mathbf{F}}_{\mathbf{P}} \cdot \dot{\mathbf{G}}_{\mathbf{P}})_{\mathbf{P}} = \langle \dot{\mathbf{F}}_{\mathbf{P}} \mathbf{P}^{-1} \cdot \dot{\mathbf{G}}_{\mathbf{P}} \mathbf{P}^{-1} \rangle_{\mathbf{I}}.$$

We will also need the mappings  $\text{Ad}_{\mathbf{Q}}, \text{ad}_{\mathcal{A}}, \mathbf{Q} \in \mathbf{G}_{\mathcal{M}}, \mathcal{A} \in \mathfrak{g}_{\mathcal{M}}$ , we mentioned before,

$$\text{Ad}_{\mathbf{Q}} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}} : \mathcal{X} \mapsto \text{Ad}_{\mathbf{Q}} \mathcal{X} = \mathbf{Q} \mathcal{X} \mathbf{Q}^{-1}, \quad \text{ad}_{\mathcal{A}} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}} : \mathcal{X} \mapsto \text{ad}_{\mathcal{A}} \mathcal{X} = \mathcal{A} \mathcal{X} - \mathcal{X} \mathcal{A}.$$

Their adjoints  $\text{Ad}_{\mathbf{Q}}^*$ ,  $\text{ad}_{\mathcal{A}} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$  depend on the inner product on  $\mathfrak{g}_{\mathcal{M}}$  and are such that

$$\forall \mathcal{X}, \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}} : \langle \text{Ad}_{\mathbf{Q}}^* \mathcal{X} \cdot \mathcal{Y} \rangle_{\mathbf{I}} = \langle \mathcal{X} \cdot \text{Ad}_{\mathbf{Q}} \mathcal{Y} \rangle_{\mathbf{I}}, \quad \text{and} \quad \langle \text{ad}_{\mathcal{A}}^* \mathcal{X} \cdot \mathcal{Y} \rangle_{\mathbf{I}} = \langle \mathcal{X} \cdot \text{ad}_{\mathcal{A}} \mathcal{Y} \rangle_{\mathbf{I}}.$$

### Lemma 5.1

For differentiable  $t \mapsto \mathcal{Y}(t) \in \mathfrak{g}_{\mathcal{M}}$  and  $t \mapsto \mathbf{Q}(t) \in \mathbf{G}_{\mathcal{M}}$ , we have

$$\frac{d}{dt} \text{Ad}_{\mathbf{Q}(t)}^* \mathcal{Y}(t) = \text{Ad}_{\mathbf{Q}(t)}^* \dot{\mathcal{Y}}(t) + \text{ad}_{\mathbf{Q}^{-1}(t) \dot{\mathbf{Q}}(t)}^* \text{Ad}_{\mathbf{Q}(t)}^* \mathcal{Y}(t) = \text{Ad}_{\mathbf{Q}(t)}^* \dot{\mathcal{Y}}(t) + \text{Ad}_{\mathbf{Q}(t)}^* \text{ad}_{\mathbf{Q}(t) \mathbf{Q}^{-1}(t)}^* \mathcal{Y}(t).$$

**Proof** For all fixed  $\mathcal{A} \in \mathfrak{g}_{\mathcal{M}}$

$$\begin{aligned} & \frac{d}{dt} \langle \text{Ad}_{\mathbf{Q}(t)}^* \mathcal{Y}(t) \cdot \mathcal{A} \rangle_{\mathbf{I}} = \frac{d}{dt} \langle \mathcal{Y}(t) \cdot \mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} = \\ & = \langle \dot{\mathcal{Y}}(t) \cdot \mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} + \langle \mathcal{Y}(t) \cdot \dot{\mathbf{Q}}(t) \mathcal{A} \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} - \langle \mathcal{Y}(t) \cdot \mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t) \dot{\mathbf{Q}}(t) \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} = \\ & = \langle \dot{\mathcal{Y}}(t) \cdot \mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} + \langle \mathcal{Y}(t) \cdot \mathbf{Q}(t) \left( \mathbf{Q}^{-1}(t) \dot{\mathbf{Q}}(t) \mathcal{A} - \mathcal{A} \mathbf{Q}^{-1}(t) \dot{\mathbf{Q}}(t) \right) \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} = \\ & = \langle \dot{\mathcal{Y}}(t) \cdot \mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t) \rangle_{\mathbf{I}} + \langle \mathcal{Y}(t) \cdot \left( \dot{\mathbf{Q}}(t) \mathbf{Q}^{-1}(t) (\mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t)) - (\mathbf{Q}(t) \mathcal{A} \mathbf{Q}^{-1}(t)) \dot{\mathbf{Q}}(t) \mathbf{Q}^{-1}(t) \right) \rangle_{\mathbf{I}}. \end{aligned}$$

Moving the operators to the left factor of the inner products in the latter two lines leads to the wanted result. ■

## Theorem 5.2 Geodesic Equations 1st version

Assume the inner product to be strictly positive.

- A necessary condition for a curve  $t \mapsto \mathbf{F}(t) \in \mathbf{G}_{\mathcal{M}}$ , with arc length parametrization, to be a geodesic with respect to a left invariant inner product is

$$\frac{d}{ds} \left[ \text{Ad}_{\mathbf{F}^{-1}(s)}^* (\mathbf{F}^{-1}(s) \dot{\mathbf{F}}(s)) \right] = 0. \quad (5.1)$$

Or, equivalently,

$$\frac{d}{ds} (\mathbf{F}^{-1}(s) \dot{\mathbf{F}}(s)) - \text{ad}_{\mathbf{F}^{-1}(s) \dot{\mathbf{F}}(s)}^* (\mathbf{F}^{-1}(s) \dot{\mathbf{F}}(s)) = 0. \quad (5.2)$$

- A necessary condition for a curve  $t \mapsto \mathbf{G}(t) \in \mathbf{G}_{\mathcal{M}}$ , with arc length parametrization, to be a geodesic with respect to a right invariant inner product is

$$\frac{d}{ds} \left[ \text{Ad}_{\mathbf{G}(s)}^* (\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)) \right] = 0. \quad (5.3)$$

Or, equivalently,

$$\frac{d}{ds} (\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)) + \text{ad}_{\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)} (\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)) = 0. \quad (5.4)$$

- A curve  $t \mapsto \mathbf{F}(t) \in \mathbf{G}_{\mathcal{M}}$  is a left invariant geodesic if and only if the curve  $t \mapsto \mathbf{F}^{-1}(t) \in \mathbf{G}_{\mathcal{M}}$  is a right invariant geodesic.

**Proof** In the previous section we considered a special left invariant case. So let us only do the right invariant case here.

Consider an arclength parametrized curve  $s \mapsto \mathbf{G}(s)$  of length  $L$  which connects two points in  $\mathbf{G}_{\mathcal{M}}$ . We suppose arclength parametrization. That means

$$\forall s : \left\langle \dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s) \cdot \dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s) \right\rangle_1 = 1. \quad (5.5)$$

Calculate the length of the perturbed curve  $\mathbf{G}_{\varepsilon; \mathcal{H}} : s \mapsto \mathbf{G}(s) e^{\varepsilon \mathcal{H}(s)}$ , with any fixed  $s \mapsto \mathcal{H}(s) \in \mathfrak{g}_{\mathcal{M}}$ ,  $\mathcal{H}(0) = \mathcal{H}(L) = 0$ :

$$L(\varepsilon; \mathcal{H}) = \int_0^L \sqrt{\left\langle \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}}(s) \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1}(s) \cdot \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}}(s) \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1}(s) \right\rangle_1} ds. \quad (5.6)$$

We first want to find the Frechet derivative  $\left. \frac{\partial}{\partial \varepsilon} L(\varepsilon; \mathcal{H}) \right|_{\varepsilon=0}$ .

For that purpose we use the expansions

$$\begin{aligned} \mathbf{G}_{\varepsilon; \mathcal{H}} &= \mathbf{G} e^{\varepsilon \mathcal{H}} = \mathbf{G} + \varepsilon \mathcal{H} \mathbf{G} + \mathcal{O}(\varepsilon^2), \\ \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1} &= e^{-\varepsilon \mathcal{H}} \mathbf{G}^{-1} = \mathbf{G}^{-1} - \varepsilon \mathcal{H} \mathbf{G}^{-1} + \mathcal{O}(\varepsilon^2), \\ \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}} &= \dot{\mathbf{G}} + \varepsilon \{ \dot{\mathbf{G}} \mathcal{H} + \mathbf{G} \dot{\mathcal{H}} \} + \mathcal{O}(\varepsilon^2), \\ \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}} \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1} &= \dot{\mathbf{G}} \mathbf{G}^{-1} + \varepsilon \{ \dot{\mathbf{G}} \mathcal{H} \mathbf{G}^{-1} \} + \mathcal{O}(\varepsilon^2), \\ \frac{\partial}{\partial \varepsilon} \left( \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}} \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1} \right) &= \mathbf{G} \dot{\mathcal{H}} \mathbf{G}^{-1} + \mathcal{O}(\varepsilon), \end{aligned} \quad (5.7)$$

It becomes clear now that

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} \left\langle \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}}(s) \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1}(s) \cdot \dot{\mathbf{G}}_{\varepsilon; \mathcal{H}}(s) \mathbf{G}_{\varepsilon; \mathcal{H}}^{-1}(s) \right\rangle \Big|_{\varepsilon=0} &= 2 \left\langle \dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s) \cdot \mathbf{G}(s) \dot{\mathcal{H}}(s) \mathbf{G}^{-1}(s) \right\rangle_1 = \\ &= 2 \left\langle \dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s) \cdot \text{Ad}_{\mathbf{G}(s)} \dot{\mathcal{H}}(s) \right\rangle_1 = 2 \left\langle \text{Ad}_{\mathbf{G}(s)}^* (\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)) \cdot \dot{\mathcal{H}}(s) \right\rangle_1. \end{aligned} \quad (5.8)$$

From this, (5.2), (5.3) and partial integration, we find for all  $\mathcal{H}$

$$\frac{\partial}{\partial \varepsilon} L(\varepsilon; \mathcal{H}) \Big|_{\varepsilon=0} = \int_0^L \left\langle \frac{d}{ds} \{ \text{Ad}_{\mathbf{G}(s)}^* (\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)) \} \cdot \mathcal{H}(s) \right\rangle_1 ds \quad (5.9)$$

If the length of our curve is required to be stationary with respect to small perturbations (=geodesic curve), we have to impose

$$\frac{\partial}{\partial \varepsilon} L(\varepsilon; \mathcal{H}) \Big|_{\varepsilon=0} = 0, \quad \text{for all functions } \mathcal{H}. \quad (5.10)$$

Because of (5.9) this implies

$$\frac{d}{ds} \left[ \text{Ad}_{\mathbf{G}(s)}^* (\dot{\mathbf{G}}(s) \mathbf{G}^{-1}(s)) \right] = 0.$$

Note that we could have taken  $e^{\varepsilon \mathcal{H}} \mathbf{G}$  for the perturbed curve. It would, of course, have led to the same result, be it at the cost of some more involved bookkeeping.  $\blacksquare$

### Remark

• If the square root  $\sqrt{\cdot}$  in (5.6) is replaced by any smooth strictly monotone increasing function  $\Theta : (0, \infty) \rightarrow (0, \infty)$ , one still arrives at (5.1)-(5.4).

In order to rewrite the differential equations (5.1), (5.2) in a less abstract way we will employ the standard inner product on  $\mathbb{R}^{n \times n}$ , given by

$$\text{Trace}[Y^\top X], \quad \text{for } Y, X \in \mathbb{R}^{n \times n}.$$

Since this inner product remains positive and non-degenerate when restricted to any linear subspace of  $\mathbb{R}^{n \times n}$ , it is possible to represent the inner product of our choice  $\langle \cdot, \cdot \rangle_I$  on  $\mathfrak{g}_{\mathcal{M}}$ , by a uniquely defined bijective linear map  $\mathcal{J} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$ , in such a way that

$$\langle \mathcal{Y}, \mathcal{X} \rangle_1 = \text{Trace}[(\mathcal{J} \mathcal{Y})^\top \mathcal{X}], \quad \text{for } \mathcal{Y}, \mathcal{X} \in \mathfrak{g}_{\mathcal{M}}. \quad (5.11)$$

We will also need the projection

$$\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{g}_{\mathcal{M}} : \mathcal{Y} \mapsto \mathcal{P}_{\mathfrak{g}} \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}}, \quad (5.12)$$

which is chosen to be orthogonal with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ .

**Lemma 5.3**

Fix a bijective linear map  $\mathcal{J} : \mathfrak{g}_M \rightarrow \mathfrak{g}_M$ .

On  $\mathfrak{g}_M$  consider the bilinear form  $\langle \mathcal{Y} \cdot \mathcal{X} \rangle_1 = \text{Trace}[(\mathcal{J}\mathcal{Y})^\top \mathcal{X}]$ .

(If the latter equals  $\text{Trace}[\mathcal{Y}^\top (\mathcal{J}\mathcal{X})]$ , it represents a symmetric inner product.)

We have

- $\text{Ad}_Q^* \mathcal{Y} = \mathcal{J}^{-1} \mathcal{P}_\mathfrak{g} (\mathcal{Q}^\top (\mathcal{J}\mathcal{Y}) \mathcal{Q}^{-\top})$ .
- $\text{ad}_A^* \mathcal{Y} = \mathcal{J}^{-1} \mathcal{P}_\mathfrak{g} (\mathcal{A}^\top (\mathcal{J}\mathcal{Y}) - (\mathcal{J}\mathcal{Y}) \mathcal{A}^\top)$ .
- If  $\mathfrak{g}_M$  has the special property  $\mathcal{A} \in \mathfrak{g}_M \Rightarrow \mathcal{A}^\top \in \mathfrak{g}_M$ , the projection  $\mathcal{P}_\mathfrak{g}$  in the expressions for  $\text{Ad}_A^*$  and  $\text{ad}_A^*$  can be omitted. This happens to be if we take  $\mathbf{G}_M = \mathbf{G}(n, \mathbb{R}, \Delta)$ , with  $\Delta^2 = \alpha I$ ,  $\alpha \in \mathbb{R}$ .

**Proof** • For  $\mathcal{X}, \mathcal{Y} \in \mathfrak{g}_M$  and  $\mathcal{Q} \in \mathbf{G}_M$ , compare

$$\begin{aligned} \langle \mathcal{Y}, \text{Ad}_Q \mathcal{X} \rangle_1 &= \langle \mathcal{Y}, \mathcal{Q} \mathcal{X} \mathcal{Q}^{-1} \rangle_1 = \text{Trace}[(\mathcal{J}\mathcal{Y})^\top \mathcal{Q} \mathcal{X} \mathcal{Q}^{-1}] = \text{Trace}[\mathcal{Q}^{-1} (\mathcal{J}\mathcal{Y})^\top \mathcal{Q} \mathcal{X}] = \\ &= \text{Trace}[(\mathcal{Q}^\top (\mathcal{J}\mathcal{Y}) \mathcal{Q}^{-\top})^\top \mathcal{X}] = \text{Trace}[(\mathcal{P}_\mathfrak{g} (\mathcal{Q}^\top (\mathcal{J}\mathcal{Y}) \mathcal{Q}^{-\top}))^\top \mathcal{X}], \end{aligned}$$

and

$$\langle \text{Ad}_Q^* \mathcal{Y}, \mathcal{X} \rangle_1 = \text{Trace}[(\mathcal{J} \text{Ad}_Q^* \mathcal{Y})^\top \mathcal{X}].$$

It follows that

$$\text{Ad}_Q^* \mathcal{Y} = \mathcal{J}^{-1} \mathcal{P}_\mathfrak{g} (\mathcal{Q}^\top (\mathcal{J}\mathcal{Y}) \mathcal{Q}^{-\top}).$$

- For  $\mathcal{X}, \mathcal{Y} \in \mathfrak{g}_M$

$$\begin{aligned} \langle \text{ad}_A^* \mathcal{Y} \cdot \mathcal{X} \rangle &= \langle \mathcal{Y} \cdot \text{ad}_A \mathcal{X} \rangle = \\ &= \text{Trace}[(\mathcal{J}\mathcal{Y})^\top (\mathcal{A}\mathcal{X} - \mathcal{X}\mathcal{A})] = \text{Trace}[(\mathcal{J}\mathcal{Y})^\top \mathcal{A} - \mathcal{A} (\mathcal{J}\mathcal{Y})^\top] \mathcal{X} = \\ &= \text{Trace}[(\mathcal{J} \mathcal{J}^{-1} \mathcal{P}_\mathfrak{g} (\mathcal{A}^\top (\mathcal{J}\mathcal{Y}) - (\mathcal{J}\mathcal{Y}) \mathcal{A}^\top))^\top \mathcal{X}]. \end{aligned}$$

- We use Note 4 of section 2, which gives a condition for:  $\mathcal{A} \in \mathfrak{g}_M \Rightarrow \mathcal{A}^\top \in \mathfrak{g}_M$ .

First,  $(\mathcal{Q}^\top (\mathcal{J}\mathcal{Y}) \mathcal{Q}^{-\top})^\top = \mathcal{Q}^{-1} (\mathcal{J}\mathcal{Y})^\top \mathcal{Q}$  belongs to  $\mathfrak{g}_M$ , because  $(\mathcal{J}\mathcal{Y})^\top$  does.

Second,  $\mathcal{A}^\top (\mathcal{J}\mathcal{Y}) - (\mathcal{J}\mathcal{Y}) \mathcal{A}^\top \in \mathfrak{g}_M$ , because  $\mathcal{J}\mathcal{Y}$  and  $\mathcal{A}^\top$  do. ■

The following Lemma will be convenient.

**Lemma 5.4**

Let  $\mathbf{V}$  be a finite dimensional vectorspace over  $\mathbb{R}$  or  $\mathbb{C}$  endowed with an inner product  $(\cdot, \cdot)$ . Let  $\mathbf{U} \subset \mathbf{V}$  be a linear subspace. By  $\mathcal{P}_U$  the orthogonal projection on  $\mathbf{U}$  is denoted.

Let  $\mathcal{K}$  be a linear map with the property  $\mathcal{K}(\mathbf{U}) \subset \mathbf{U}$ . Then the adjoint  $\mathcal{K}^*$  has the properties

$$\mathcal{P}_U \mathcal{K}^* \mathcal{P}_U^\perp = 0, \quad \mathcal{P}_U \mathcal{K}^* = \mathcal{P}_U \mathcal{K}^* \mathcal{P}_U, \quad \mathcal{P}_U \mathcal{K}^{-*} \mathcal{P}_U \mathcal{K}^* = \mathcal{P}_U. \quad (5.13)$$

**Proof** For all  $\underline{u}, \underline{v} \in \mathbf{V}$  we have

$$(\mathcal{P}_{\mathbf{U}} \mathcal{K}^* \mathcal{P}_{\mathbf{U}}^\perp \underline{u}, \underline{v}) = (\mathcal{K}^* \mathcal{P}_{\mathbf{U}}^\perp \underline{u}, \mathcal{P}_{\mathbf{U}} \underline{v}) = (\mathcal{P}_{\mathbf{U}}^\perp \underline{u}, \mathcal{K} \mathcal{P}_{\mathbf{U}} \underline{v}) = 0$$

The second statement follows from  $\mathcal{P}_{\mathbf{U}} \mathcal{K}^* = \mathcal{P}_{\mathbf{U}} \mathcal{K}^* (\mathcal{P}_{\mathbf{U}} + \mathcal{P}_{\mathbf{U}}^\perp)$ . ■

Note that this Lemma has some obvious generalisations to infinite dimensional Hilbertspaces.

### Theorem 5.5 Geodesic Equations 2nd version

Assume the inner product to be strictly positive.

• A necessary condition for a curve  $t \mapsto F(t) \in \mathbf{G}_{\mathcal{M}}$ , with arc length parametrization, to be a geodesic with respect to a left invariant inner product, is

$$\left. \begin{aligned} & \frac{d}{ds} \mathcal{P}_{\mathfrak{g}} \{ F^{-\top}(s) \mathcal{J}(F^{-1} \dot{F})(s) F^\top(s) \} = 0, \\ \text{or, with } & \frac{d}{ds} F(s) = F(s) \mathcal{A}(s) \text{ and } \frac{d}{ds} F^{-1}(s) = -\mathcal{A}(s) F^{-1}(s), \\ & \mathcal{J}(\dot{\mathcal{A}}) - \mathcal{P}_{\mathfrak{g}} [\mathcal{A}^\top \mathcal{J}(\mathcal{A}) - \mathcal{J}(\mathcal{A}) \mathcal{A}^\top] = 0. \end{aligned} \right\} \quad (5.14)$$

• A necessary condition for a curve  $t \mapsto G(t) \in \mathbf{G}_{\mathcal{M}}$ , with arc length parametrization, to be a geodesic with respect to a right invariant inner product, is

$$\left. \begin{aligned} & \frac{d}{ds} \mathcal{P}_{\mathfrak{g}} \{ G^\top(s) \mathcal{J}(\dot{G} G^{-1})(s) G^{-\top}(s) \} = 0, \\ \text{or, with } & \frac{d}{ds} G(s) = \Omega(s) G(s) \text{ and } \frac{d}{ds} G^{-1}(s) = -G^{-1}(s) \Omega(s), \\ & \mathcal{J}(\dot{\Omega}) + \mathcal{P}_{\mathfrak{g}} [\Omega^\top \mathcal{J}(\Omega) - \mathcal{J}(\Omega) \Omega^\top] = 0. \end{aligned} \right\} \quad (5.15)$$

**Proof 1** Apply Lemma 5.3 to (5.1)-(5.3).

**Proof 2** Without Lemma 5.3. Consider the right invariant case. Starting from (5.8) we derive

$$\begin{aligned} \left\langle \dot{G}(s) G^{-1}(s) \cdot G(s) \dot{\mathcal{H}}(s) G^{-1}(s) \right\rangle_I &= \text{Trace} [\mathcal{J}(\dot{G}(s) G^{-1}(s))^\top G(s) \dot{\mathcal{H}}(s) G^{-1}(s)] = \\ &= \text{Trace} \left[ \left( G^\top \mathcal{J}(\dot{G}(s) G^{-1}(s)) G^{-\top} \right)^\top \dot{\mathcal{H}}(s) \right]. \end{aligned}$$

Take the integral  $\int_0^L \dots ds$  of this expression and integrate partially. The integrand becomes

$$-\text{Trace} \left[ \left( \mathcal{P}_{\mathfrak{g}} \frac{d}{ds} G^\top \mathcal{J}(\dot{G}(s) G^{-1}(s)) G^{-\top} \right)^\top \mathcal{H}(s) \right],$$

which should be 0 for all  $\mathcal{H}(\cdot)$ . This gives the 1st formula in (5.15). Writing out the derivatives leads to

$$\mathcal{P}_{\mathfrak{g}} [ G^\top (\mathcal{J}(\dot{\Omega}) + \Omega^\top \mathcal{J}(\Omega) - \mathcal{J}(\Omega) \Omega^\top) G^{-\top} ] = 0.$$

Then also

$$\mathcal{P}_{\mathfrak{g}} \left[ G^{-\top} \left\{ \mathcal{P}_{\mathfrak{g}} \left[ G^{\top} \left( \mathcal{J}(\dot{\Omega}) + \Omega^{\top} \mathcal{J}(\Omega) - \mathcal{J}(\Omega)\Omega^{\top} \right) G^{-\top} \right] \right\} G^{\top} \right] = 0.$$

Finally, with the 3rd formula in (5.13) of Lemma 5.4 we arrive at

$$\mathcal{P}_{\mathfrak{g}} \left( \mathcal{J}(\dot{\Omega}) + \Omega^{\top} \mathcal{J}(\Omega) - \mathcal{J}(\Omega)\Omega^{\top} \right) = 0,$$

the 3rd formula in (5.15), since  $\mathcal{P}_{\mathfrak{g}} \mathcal{J}(\dot{\Omega}) = \mathcal{J}(\dot{\Omega}) \in \mathfrak{g}_{\mathcal{M}}$ . ■

Note again that, if it happens that for every  $F \in \mathbf{G}_{\mathcal{M}}$  also  $F^{\top} \in \mathbf{G}_{\mathcal{M}}$ , the projection  $\mathcal{P}_{\mathfrak{g}}$  in (5.14) and (5.15) acts as the identity map and we are left with a generalization sort of Euler's Equations in section 1.

The inner products that we want to consider in the sequel of this section are of type

$$\left\langle \mathcal{Y}, \mathcal{X} \right\rangle_{\mathfrak{g}} = \text{Trace}[K\mathcal{Y}^{\top}L\mathcal{X}], \quad (5.16)$$

with  $K, L \in \mathbb{R}^{n \times n}$ , both *symmetric* and *invertible*.

This inner product is symmetric:

$$\begin{aligned} \left\langle \mathcal{Y}, \mathcal{X} \right\rangle_{\mathfrak{g}} &= \text{Trace}[K\mathcal{Y}^{\top}L\mathcal{X}] = \text{Trace}[L\mathcal{X}K\mathcal{Y}^{\top}] = \text{Trace}[\mathcal{Y}K\mathcal{X}^{\top}L] = \\ &= \text{Trace}[K\mathcal{X}^{\top}L\mathcal{Y}] = \left\langle \mathcal{X}, \mathcal{Y} \right\rangle_{\mathfrak{g}}. \end{aligned}$$

Note that for the type of inner product (5.16)  $\mathcal{J}$  can be expressed with  $\mathcal{P}_{\mathfrak{g}}$

$$\text{Trace}[K\mathcal{Y}^{\top}L\mathcal{X}] = \text{Trace}[(\mathcal{P}_{\mathfrak{g}}(LYK))^{\top}\mathcal{X}], \quad \mathcal{X}, \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}},$$

hence

$$\mathcal{J}\mathcal{Y} = \mathcal{P}_{\mathfrak{g}}(LYK), \quad \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}}. \quad (5.17)$$

If  $K$  and  $L$  happen to be diagonal matrices,  $D$  and  $\Lambda$ , say, we have

$$\text{Trace}[D\mathcal{Y}^{\top}\Lambda\mathcal{X}] = \sum_i \sum_j D_{jj}\Lambda_{ii}\mathcal{Y}_{ij}\mathcal{X}_{ij}.$$

This inner product is non-degenerate iff all  $D_{ii} \neq 0$  and all  $\Lambda_{jj} \neq 0$ . The inner product is positive iff all  $D_{ii} > 0$  and all  $\Lambda_{jj} > 0$ . In general we have, for suitable orthogonal  $U, V$  that  $K = UDU^{\top}$  and  $L = V\Lambda V^{\top}$  and

$$\left\langle \mathcal{Y}, \mathcal{X} \right\rangle_{\mathfrak{g}} = \text{Trace}[K\mathcal{Y}^{\top}L\mathcal{X}] = \text{Trace}[UDU^{\top}\mathcal{Y}^{\top}V\Lambda V^{\top}\mathcal{X}] = \text{Trace}[D(V^{\top}\mathcal{Y}U)^{\top}\mathcal{Y}^{\top}\Lambda(V^{\top}\mathcal{X}U)].$$

On  $\mathbb{R}^{n \times n}$  as a whole we conclude non-degeneracy iff,  $K, L$  are invertible and positivity iff  $K, L > 0$ . In the non-positive case we have to be careful about the non-degeneracy if we



restrict this inner product to a Lie algebra subspace  $\mathfrak{g}_{\mathcal{M}}$ .

We want to write down the geodesic equations for a series of special cases. In each case the mapping  $\mathcal{J}$  and the projection  $\mathcal{P}_{\mathfrak{g}}$  have to be made explicit.

**Case 1.**  $\mathfrak{g}_{\mathcal{M}} = \mathfrak{so}(n)$

First we find conditions on  $K, L$  such that the inner product (5.16) is non-degenerate on  $\mathfrak{so}(n)$ , the space of skew-symmetric matrices. It is required that for  $\mathcal{Y} \in \mathfrak{so}(n)$  and  $\text{Trace}[K\mathcal{Y}^{\top}L\mathcal{X}] = 0$ , for all  $\mathcal{X} \in \mathfrak{so}(n)$ , it should follow that  $\mathcal{Y} = 0$ . From the latter property it follows that  $K\mathcal{Y}^{\top}L$  is symmetric:  $K\mathcal{Y}^{\top}L = L\mathcal{Y}K$ .

Hence  $\mathcal{Y}$  satisfies  $K\mathcal{Y}L + L\mathcal{Y}K = 0$ , or  $(L^{-1}K)\mathcal{Y} + \mathcal{Y}(KL^{-1}) = 0$ . Gantmacher tells us, [G] Ch VIII.3, that the latter equation has  $\mathcal{Y} = 0$  as its only solution iff the matrices  $L^{-1}K$  and  $-KL^{-1}$  have no characteristic values in common. Since  $(L^{-1}K)^{\top} = KL^{-1}$ , it is forbidden for  $(L^{-1}K)$  to have a set of characteristic values which contains a number together with its opposite.

The orthogonal projection  $\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{so}(n)$ , with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ , is obviously given by

$$\mathcal{P}_{\mathfrak{g}} Z = \frac{1}{2}\{Z - Z^{\top}\}, \quad Z \in \mathbb{R}^{n \times n}$$

The mapping  $\mathcal{J} : \mathfrak{so}(n) \rightarrow \mathfrak{so}(n)$  is given by

$$\mathcal{J}\mathcal{Y} = \mathcal{P}_{\mathfrak{g}}(L\mathcal{Y}K) = \frac{1}{2}\{L\mathcal{Y}K - K\mathcal{Y}^{\top}L\}, \quad \mathcal{Y} \in \mathfrak{so}(n).$$

- In our special Case 1. the geodesic equation (5.14) thus becomes

$$\frac{1}{2} \frac{d}{ds} \mathcal{P}_{\mathfrak{g}} \{F[L(F^{-1}\dot{F})K + K(F^{-1}\dot{F})L]F^{-1}\} = \frac{1}{2} \frac{d}{ds} \{F[L(F^{-1}\dot{F})K + K(F^{-1}\dot{F})L]F^{-1}\} = 0.$$

With  $\mathcal{A} = F^{-1}\dot{F}$  this can be written

$$L\dot{\mathcal{A}}K + K\dot{\mathcal{A}}L = LAKA - AKAL + KALA - ALAK, \quad (5.18)$$

which corresponds to (1.18) if  $n = 3$  and  $K = L = J^{-1}$  is taken.

Once  $\mathcal{A}$  has been found,  $F$  can be solved from  $\dot{F} = F\mathcal{A}$ .

Note that, because of the condition we imposed on the pair  $K, L$ , the derivative  $\dot{\mathcal{A}}$  is solvable from (5.18).

Suppose that there exists invertible  $S \in \mathbb{R}^{n \times n}$  such that  $KL^{-1} = SDS^{-1}$ , with  $D = \text{diag}[D_{11}, \dots, D_{nn}]$ , then  $\dot{\mathcal{A}}$  can be 'explicitly' solved from from (5.17) via

$$[S^{\top}\dot{\mathcal{A}}S]_{ij} = \frac{[S^{\top}L^{-1}\{LAKA - AKAL + KALA - ALAK\}L^{-1}S]_{ij}}{D_{ii} + D_{jj}}. \quad (5.19)$$

Note that the diagonalizability assumption applies if, e.g.,  $K > 0$  or  $L > 0$ .

**Case 2.**  $\mathfrak{g}_{\mathcal{M}} = \mathfrak{so}(r, s)$

First we find conditions on  $K, L$  such that the inner product (5.16) is non-degenerate on  $\mathfrak{so}(r, s)$ , that is the space of matrices  $\mathcal{X} \in \mathbb{R}^{n \times n}$ , such that  $\Delta \mathcal{X} + \mathcal{X}^\top \Delta = O$ . Hence  $\Delta \mathcal{X}$  is a skew-symmetric matrix. Remind that the matrix  $\Delta$  has the properties  $\Delta = \Delta^\top = \Delta^{-1}$ . For non-degeneracy of the inner product it is required that for  $\mathcal{Y} \in \mathfrak{so}(r, s)$  and  $\text{Trace}[K \mathcal{Y}^\top L \mathcal{X}] = 0$ , for all  $\mathcal{X} \in \mathfrak{so}(r, s)$ , it should follow that  $\mathcal{Y} = 0$ . From the latter property it follows that  $K(\mathcal{Y}^\top \Delta) \Delta L \Delta$  has to be symmetric:  $K(\mathcal{Y}^\top \Delta) \Delta L \Delta = \Delta L \Delta (\Delta \mathcal{Y}) K$ , or  $\Delta L \Delta (\Delta \mathcal{Y}) K + K (\Delta \mathcal{Y}) \Delta L \Delta = 0$ . As in the preceding case we conclude that it is forbidden that both  $\lambda$  and  $-\lambda$  belong to the set of characteristic values of  $(\Delta L^{-1} \Delta) K$ .

Let  $Z \in \mathbb{R}^{n \times n}$ . Write  $Z = \frac{1}{2}\{Z - \Delta Z^\top \Delta\} + \frac{1}{2}\{Z + \Delta Z^\top \Delta\}$ . The first term belongs to  $\mathfrak{so}(r, s)$  since it satisfies (1.2). The second term is orthogonal to  $\mathfrak{so}(r, s)$  with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ . Indeed  $\text{Trace}[\{Z + \Delta Z^\top \Delta\}^\top X] = \text{Trace}[\{\Delta Z + Z^\top \Delta\}^\top \Delta \mathcal{X}]$ , which contains a product of a symmetric and an antisymmetric matrix.

Therefore the orthogonal projection  $\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{so}(r, s)$ , with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ , is given by

$$\mathcal{P}_{\mathfrak{g}} Z = \frac{1}{2}\{Z - \Delta Z^\top \Delta\}, \quad Z \in \mathbb{R}^{n \times n}$$

The mapping  $\mathcal{J} : \mathfrak{so}(n) \rightarrow \mathfrak{so}(r, s)$  is given by

$$\mathcal{J} \mathcal{Y} = \mathcal{P}_{\mathfrak{g}} (L \mathcal{Y} K) = \frac{1}{2}\{L \mathcal{Y} K - \Delta K \mathcal{Y}^\top L \Delta\} = \frac{1}{2}\{L \mathcal{Y} K + (\Delta K \Delta) \mathcal{Y} (\Delta L \Delta)\}, \quad \mathcal{Y} \in \mathfrak{so}(r, s).$$

As in the previous case the Lie group  $\mathfrak{so}(r, s)$  contains the transposed of each of its elements. To see this, take the inverse of both sides of the identity  $\mathbf{P}^{-\top} \Delta \mathbf{P}^{-1} = \Delta$ . This implies that in the equations (5.14) and (5.15) the projection  $\mathcal{P}_{\mathfrak{g}}$  acts as the identity map.

- In the underlying Case 2. the geodesic equation (5.14) reads

$$L \dot{\mathcal{A}} K + (\Delta K \Delta) \dot{\mathcal{A}} (\Delta L \Delta) = [\mathcal{A}^\top L \mathcal{A} K - L \mathcal{A} K \mathcal{A}^\top] + \Delta [\mathcal{A} K \mathcal{A}^\top L - K \mathcal{A}^\top L \mathcal{A}] \Delta. \quad (5.20)$$

- In the very special case where  $K = \Delta L \Delta$  is taken, the geodesic equation can be written as an 'explicit' differential equation:

$$\dot{\mathcal{A}} = L^{-1} \mathcal{A}^\top L \mathcal{A} - \Delta \mathcal{A}^\top L \mathcal{A} L^{-1} \Delta, \quad (5.21)$$

**Case 3.**  $\mathfrak{g}_{\mathcal{M}} = \mathfrak{gl}(n) = \mathbb{R}^{n \times n}$

The projection  $\mathcal{P}_{\mathfrak{g}}$  in (5.17) is the identity map in this special case. As a consequence

$$\mathcal{J}\mathcal{Y} = \mathcal{P}_{\mathfrak{g}}(L\mathcal{Y}K) = L\mathcal{Y}K.$$

- In our special Case 3. the geodesic equation (5.14) therefore becomes

$$\dot{A} = L^{-1}A^{\top}LA - \mathcal{A}K\mathcal{A}^{\top}K^{-1}, \quad (5.22)$$

**Case 4.**  $\mathfrak{g}_{\mathcal{M}} = \mathfrak{sgl}(n)$ .

First we find conditions on  $K, L$  such that the inner product (5.16) is non-degenerate on  $\mathfrak{sgl}(n)$ , the space of matrices having trace equal 0. It is required that for  $\mathcal{Y} \in \mathfrak{sgl}(n)$  and  $\text{Trace}[K\mathcal{Y}^{\top}L\mathcal{X}] = 0$ , for all  $\mathcal{X} \in \mathfrak{sgl}(n)$ , it should follow that  $\mathcal{Y} = 0$ . From the latter property it follows that  $K\mathcal{Y}^{\top}L$  is a multiple of the identity matrix:  $K\mathcal{Y}^{\top}L = \beta I$ . Hence  $\mathcal{Y} = \beta L^{-1}K^{-1}$ . Since  $\text{Trace}[\mathcal{Y}] = 0$ , we have  $\beta = 0$  precisely if  $\text{Trace}[L^{-1}K^{-1}] \neq 0$ . It is the latter condition that we have to impose on  $K, L$ !

Let  $Z \in \mathbb{R}^{n \times n}$ . Write  $Z = \{Z - \frac{1}{n}\text{Trace}[Z]I\} + \frac{1}{n}\text{Trace}[Z]I$ . The first term belongs to  $\mathfrak{sgl}(n)$ , it has trace equal 0. The second term, being a multiple of the identity matrix, is orthogonal to  $\mathfrak{sgl}(n)$  with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ .

Therefore the orthogonal projection  $\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{sgl}(n)$ , with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ , is given by

$$\mathcal{P}_{\mathfrak{g}} Z = Z - \frac{1}{n}\text{Trace}[Z]I$$

The mapping  $\mathcal{J} : \mathfrak{sgl}(n) \rightarrow \mathfrak{sgl}(n)$  is given by

$$\mathcal{J}\mathcal{Y} = \mathcal{P}_{\mathfrak{g}}(L\mathcal{Y}K) = L\mathcal{Y}K - \frac{1}{n}\text{Trace}[L\mathcal{Y}K]I.$$

We will need the solution of the equation  $\mathcal{J}\mathcal{Y} = W$ , with  $\text{Trace}[\mathcal{Y}] = \text{Trace}[W] = 0$ . Multiplication from the left by  $L^{-1}$  and from the right by  $K^{-1}$  and taking the trace leads to an expression for  $\text{Trace}[L\mathcal{Y}K]$ . The solution  $\mathcal{Y}$  turns out to be

$$\mathcal{Y} = L^{-1}WK^{-1} - \frac{\text{Trace}[L^{-1}WK^{-1}]}{\text{Trace}[L^{-1}K^{-1}]}L^{-1}K^{-1}. \quad (5.23)$$

As in the previous cases the Lie group  $\mathfrak{sgl}(n)$  contains the transposed of each of its elements. The equation (5.14) for the left geodesic becomes in this case

$$L\dot{A}K - \frac{1}{n}\text{Trace}[L\dot{A}K]I = \mathcal{A}^{\top}L\mathcal{A}K - L\mathcal{A}K\mathcal{A}^{\top}.$$

Via (5.21)  $\dot{A}$  can be solved from this identity. Which leads to the 'explicit' differential equation for the left geodesics of our special case 4.:

$$\dot{\mathcal{A}} = L^{-1}\mathcal{A}^\top L\mathcal{A} - \mathcal{A}K\mathcal{A}^\top K^{-1} - \frac{\text{Trace}[L^{-1}\mathcal{A}^\top L\mathcal{A}] - \text{Trace}[\mathcal{A}K\mathcal{A}^\top K^{-1}]}{\text{Trace}[L^{-1}K^{-1}]} L^{-1}K^{-1}. \quad (5.24)$$

**Case 5.**  $\mathfrak{g}_{\mathcal{M}} = \mathfrak{heis}(2)$ .

We denote the elements of the matrix Lie-group **Heis**(2) and the corresponding Lie Algebra  $\mathfrak{heis}(2)$  by

$$\mathbf{F} = \begin{bmatrix} 1 & f_1 & f_3 \\ 0 & 1 & f_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 & g_1 & g_3 \\ 0 & 1 & g_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{resp } \mathcal{A} = \begin{bmatrix} 0 & \alpha_1 & \alpha_3 \\ 0 & 0 & \alpha_2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 & \beta_1 & \beta_3 \\ 0 & 0 & \beta_2 \\ 0 & 0 & 0 \end{bmatrix}$$

Calculate

$$\begin{aligned} \mathbf{FG} &= \begin{bmatrix} 1 & f_1 + g_1 & f_3 + g_3 + f_1g_2 \\ 0 & 1 & f_2 + g_2 \\ 0 & 0 & 1 \end{bmatrix}, \\ \mathbf{F}^{-1} &= \begin{bmatrix} 1 & -f_1 & f_1f_2 - f_3 \\ 0 & 1 & -f_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{F}^\top = \begin{bmatrix} 1 & 0 & 0 \\ f_1 & 1 & 0 \\ f_3 & f_2 & 1 \end{bmatrix}, \quad \mathbf{F}^{-\top} = \begin{bmatrix} 1 & 0 & 0 \\ -f_1 & 1 & 0 \\ f_1f_2 - f_3 & -f_2 & 1 \end{bmatrix}, \\ \mathbf{F}^{-\top}\dot{\mathcal{B}}\mathbf{F}^\top &= \begin{bmatrix} \star & \dot{\beta}_1 + \dot{\beta}_3f_2 & \dot{\beta}_3 \\ \star & \star & -f_1\dot{\beta}_3 + \dot{\beta}_2 \\ \star & \star & \star \end{bmatrix} \\ \mathcal{A} = \mathbf{F}^{-1}\dot{\mathcal{F}} &= \begin{bmatrix} 0 & \dot{f}_1 & \dot{f}_3 - f_1\dot{f}_2 \\ 0 & 0 & \dot{f}_2 \\ 0 & 0 & 0 \end{bmatrix} \\ \mathcal{B}\mathcal{A}^\top - \mathcal{A}^\top\mathcal{B} &= \begin{bmatrix} \beta_1\alpha_1 + \beta_3\alpha_3 & \beta_3\alpha_2 & 0 \\ \beta_2\alpha_3 & \beta_2\alpha_2 - \beta_1\alpha_1 & -\beta_3\alpha_1 \\ 0 & -\beta_1\alpha_3 & -\beta_3\alpha_3 - \beta_2\alpha_2 \end{bmatrix} \\ \mathbf{F}^{-\top}[\mathcal{B}\mathcal{A}^\top - \mathcal{A}^\top\mathcal{B}]\mathbf{F}^\top &= \begin{bmatrix} \star & \beta_3\alpha_2 & 0 \\ \star & \star & -\beta_3\alpha_1 \\ \star & \star & \star \end{bmatrix} \end{aligned}$$

Note that because of the projection  $\mathcal{P}_{\mathfrak{g}}$ , in this case the entries  $\star$  need not be calculated.

Next we write  $\mathcal{B} = \mathcal{J}\mathcal{A}$  or  $\mathcal{A} = \mathcal{J}^{-1}\mathcal{B}$ . The equations for left geodesics become

$$\dot{\beta}_1 + \dot{\beta}_3f_2 + \beta_3\alpha_2 = 0, \quad \dot{\beta}_2 - \dot{\beta}_3f_1 - \beta_3\alpha_1 = 0, \quad \dot{\beta}_3 = 0.$$

We find  $\beta_3 = C$ , a constant equal to its initial value, and we are left with the geodesic equations expressed in the entries of  $\mathcal{B}$ .

$$\begin{cases} \dot{\beta}_1 + \beta_3\alpha_2 = 0 \\ \dot{\beta}_2 - \beta_3\alpha_1 = 0 \end{cases} \quad (5.25)$$

By means of  $\mathcal{J}$  the values of  $\alpha_j$  can be linearly expressed in  $\beta_j$  and the equations (5.21) have global explicit solutions  $s \mapsto \beta_j(s)$ . The solution curve lies in a plane, which is, a fortiori, also true for  $s \mapsto \alpha_j(s)$ . Finally the left geodesics  $s \mapsto f_k(s)$  can be found from

$$\begin{cases} \dot{f}_1 = \alpha_1 \\ \dot{f}_2 = \alpha_2 \\ \dot{f}_3 = \alpha_3 + f_1\alpha_2 \end{cases} \quad (5.26)$$

So *all* geodesics with respect to *any* inner product can be found explicitly.

**Case 6.**  $\mathfrak{g}_{\mathcal{M}} = \mathfrak{se}(n)$ .

We denote the elements of the matrix Lie-group  $\mathbf{SE}(n)$  and the corresponding Lie Algebra  $\mathfrak{se}(n)$  by

$$\mathbf{F} = \begin{bmatrix} \Phi & \underline{f} \\ \underline{0}^\top & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \Psi & \underline{g} \\ \underline{0}^\top & 1 \end{bmatrix}, \quad \text{respectively,} \quad \mathcal{A} = \begin{bmatrix} \mathbf{A} & \underline{a} \\ \underline{0}^\top & 0 \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} \mathbf{B} & \underline{b} \\ \underline{0}^\top & 0 \end{bmatrix},$$

with  $\mathbf{F}, \mathbf{G}, \mathcal{A}, \mathcal{B} \in \mathbb{R}^{(n+1) \times (n+1)}$ ,  $\Phi, \Psi, \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ ,  $\underline{f}, \underline{g}, \underline{a}, \underline{b} \in \mathbb{R}^n$  and  $\Phi^\top = \Phi^{-1}$ ,  $\Psi^\top = \Psi^{-1}$ ,  $\mathbf{A} = -\mathbf{A}^\top$ ,  $\mathbf{B} = -\mathbf{B}^\top$ .

Calculate

$$\begin{aligned} \mathbf{F}^{-1} &= \begin{bmatrix} \Phi^{-1} & -\Phi^{-1}\underline{f} \\ \underline{0}^\top & 1 \end{bmatrix}, \quad \mathbf{F}^\top = \begin{bmatrix} \Phi^{-1} & \underline{0} \\ \underline{f}^\top & 1 \end{bmatrix}, \quad \mathbf{F}^{-\top} = \begin{bmatrix} \Phi & \underline{0} \\ -\underline{f}^\top\Phi & 1 \end{bmatrix}, \\ \mathbf{F}^{-\top}\dot{\mathbf{B}}\mathbf{F}^\top &= \begin{bmatrix} \Phi\dot{\mathbf{B}}\Phi^{-1} + \Phi\dot{\underline{b}}\underline{f}^\top & \Phi\dot{\underline{b}} \\ -\underline{f}^\top\Phi\dot{\mathbf{B}}\Phi^{-1} - \underline{f}^\top\Phi\dot{\underline{b}}\underline{f}^\top & -\underline{f}^\top\Phi\dot{\underline{b}} \end{bmatrix}, \\ \mathcal{A} = \mathbf{F}^{-1}\dot{\mathbf{F}} &= \begin{bmatrix} \Phi^{-1}\dot{\Phi} & \Phi^{-1}\dot{\underline{f}} \\ 0 & 0 \end{bmatrix}, \\ \mathcal{B}\mathcal{A}^\top &= \begin{bmatrix} \mathbf{B}\mathbf{A}^\top + \underline{b}\underline{a}^\top & 0 \\ \underline{0}^\top & 0 \end{bmatrix}, \quad \mathcal{A}^\top\mathcal{B} = \begin{bmatrix} \mathbf{A}^\top\mathbf{B} & \mathbf{A}^\top\underline{b} \\ \underline{a}^\top\mathbf{B} & \underline{a}^\top\underline{b} \end{bmatrix}, \\ \mathbf{F}^{-\top}[\mathcal{B}\mathcal{A}^\top - \mathcal{A}^\top\mathcal{B}]\mathbf{F}^\top &= \begin{bmatrix} \Phi(\mathbf{B}\mathbf{A}^\top - \mathbf{A}^\top\mathbf{B})\Phi^{-1} - \Phi\mathbf{A}^\top\underline{b}\underline{f}^\top + \Phi\underline{b}\underline{a}^\top\Phi^{-1} & -\Phi\mathbf{A}^\top\underline{b} \\ \star & \star \end{bmatrix}. \end{aligned}$$

Note that because of the projection  $\mathcal{P}_{\mathfrak{g}}$ , in this case the entries  $\star$  need not be calculated.

Represent the vector space  $\mathfrak{se}(n)$  as a direct sum  $\mathfrak{so}(n) \oplus \mathbb{R}^n$ . Correspondingly we split the linear (inner product) map  $\mathcal{J} : \mathfrak{se}(n) \rightarrow \mathfrak{se}(n)$  as

$$\begin{cases} \mathcal{J}_1 : \mathfrak{so}(n) \oplus \mathbb{R}^n \rightarrow \mathfrak{so}(n), & \mathbf{A} \oplus \underline{a} \mapsto \mathcal{J}_1(\mathbf{A} \oplus \underline{a}) \\ \mathcal{J}_2 : \mathfrak{so}(n) \oplus \mathbb{R}^n \rightarrow \mathbb{R}^n, & \mathbf{A} \oplus \underline{a} \mapsto \mathcal{J}_2(\mathbf{A} \oplus \underline{a}) \end{cases}, \quad (5.27)$$

such that

$$\mathcal{I} \mathcal{A} = \mathcal{B} = \begin{bmatrix} \underline{\mathbf{B}} & \underline{\mathbf{b}} \\ \underline{\mathbf{0}}^\top & 0 \end{bmatrix} = \begin{bmatrix} \mathcal{I}_1(\mathbf{A} \oplus \underline{\mathbf{a}}) & \mathcal{I}_2(\mathbf{A} \oplus \underline{\mathbf{a}}) \\ \underline{\mathbf{0}}^\top & 0 \end{bmatrix}.$$

The projection  $\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{(n+1) \times (n+1)} \rightarrow \mathfrak{se}(n)$  does

$$\mathcal{P}_{\mathfrak{g}} \begin{bmatrix} \underline{\mathbf{C}} & \underline{\mathbf{c}} \\ \underline{\mathbf{d}}^\top & \gamma \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(\underline{\mathbf{C}} - \underline{\mathbf{C}}^\top) & \underline{\mathbf{c}} \\ \underline{\mathbf{0}}^\top & 0 \end{bmatrix}.$$

We gather

$$\begin{aligned} & \mathcal{P}_{\mathfrak{g}} \mathbf{F}^{-\top} \{ \dot{\underline{\mathbf{B}}} + \underline{\mathbf{B}} \mathbf{A}^\top - \mathbf{A}^\top \underline{\mathbf{B}} \} \mathbf{F}^\top = \\ & = \begin{bmatrix} \Phi \left\{ \dot{\underline{\mathbf{B}}} + \underline{\mathbf{B}} \mathbf{A}^\top - \mathbf{A}^\top \underline{\mathbf{B}} + \frac{1}{2}(\underline{\mathbf{b}} \underline{\mathbf{f}}^\top \Phi - \Phi^{-1} \underline{\mathbf{f}} \underline{\mathbf{b}}^\top) - \frac{1}{2}(\mathbf{A}^\top \underline{\mathbf{b}} \underline{\mathbf{f}}^\top \Phi - \Phi^{-1} \underline{\mathbf{f}} \underline{\mathbf{b}}^\top \mathbf{A}) \right\} \Phi^{-1} & \Phi \left\{ \underline{\mathbf{b}} - \mathbf{A}^\top \underline{\mathbf{b}} \right\} \\ \underline{\mathbf{0}}^\top & 0 \end{bmatrix} \end{aligned}$$

For the geodesics the entries of this matrix have to be put equal to 0. Because of the upper right entry equal to zero, the left upper entry reduces considerably. For the left geodesic equations we find

$$\begin{cases} \mathcal{I}_1(\dot{\mathbf{A}} \oplus \dot{\underline{\mathbf{a}}}) + \mathcal{I}_1(\mathbf{A} \oplus \underline{\mathbf{a}}) \mathbf{A}^\top - \mathbf{A}^\top \mathcal{I}_1(\mathbf{A} \oplus \underline{\mathbf{a}}) = 0 \\ \mathcal{I}_2(\dot{\mathbf{A}} \oplus \dot{\underline{\mathbf{a}}}) - \mathbf{A}^\top \mathcal{I}_2(\mathbf{A} \oplus \underline{\mathbf{a}}) = \underline{\mathbf{0}} \end{cases}. \quad (5.28)$$

Note that the two equations are decoupled if the chosen inner product is 'decoupled', that means  $\mathcal{I}_1(\mathbf{A} \oplus \underline{\mathbf{a}}) = \mathcal{I}_1(\mathbf{A})$  and  $\mathcal{I}_2(\mathbf{A} \oplus \underline{\mathbf{a}}) = \mathcal{I}_2(\underline{\mathbf{a}})$ .

## 6 Functions and Vectorfields on Matrix Lie Groups

For smooth  $\mathbb{R}$ -valued functions  $M \mapsto \phi(M)$ , with  $M = [M_{ij}]$  in an open set of the Euclidean space  $\mathbb{R}^{n \times n}$ , we employ the standard freshman's concepts of 1st and 2nd order derivatives.

### Notation

- The 1st order derivative of  $\phi$  at  $M \in \mathbb{R}^{n \times n}$  is written as a matrix  $d\phi(M) = [d\phi_{ij}(M)]$  with  $d\phi_{ij}(M) = \frac{\partial \phi}{\partial X_{ij}}(M)$ . For the directional derivative of  $\phi$  at  $M$  in the direction  $V$  we denote

$$\langle d\phi(M): V \rangle = \sum_{i,j=1}^n \frac{\partial \phi}{\partial X_{ij}}(M) V_{ij} = \text{Trace}[d\phi(M)^\top V].$$

- The 2nd order derivative of  $\phi$  at  $M \in \mathbb{R}^{n \times n}$  is denoted by  $d^2\phi(M)$ , which can be looked upon as a  $n^2 \times n^2$ -matrix  $[\frac{\partial^2 \phi}{\partial X_{ij} \partial X_{kl}}(M)]$ . For matrices  $V, W \in \mathbb{R}^{n \times n}$  we denote

$$\langle V: d^2\phi(M): W \rangle = \langle W: d^2\phi(M): V \rangle = \sum_{i,j,k,\ell=1}^n \frac{\partial^2 \phi}{\partial X_{ij} \partial X_{kl}}(M) V_{ij} W_{kl}.$$

For the function  $\phi$  we will usually take a local extension  $f_e$  of a function  $f$  defined on  $\mathbf{G}_M$ . The specific choice of this extension will not matter.

### Notation

- Let  $f: \mathbf{G}_M \rightarrow \mathbb{R}$  be a smooth function. By  $f_e$  we denote a smooth extension of  $f$  to an open neighborhood of  $P \in \mathbf{G}_M$  in  $\mathbb{R}^{n \times n}$ . There are many such extensions available. As an example: First choose any basis  $\{\mathcal{A}_1, \dots, \mathcal{A}_\mathfrak{d}\} \subset \mathfrak{g}_M$  and extend it to a basis  $\{\mathcal{A}_1, \dots, \mathcal{A}_\mathfrak{d}, \mathcal{A}_{\mathfrak{d}+1}, \dots, \mathcal{A}_{n^2}\}$  of  $\mathbb{R}^{n \times n}$ . Then, define

$$f_e(P \exp(t_1 \mathcal{A}_1 + \dots + t_\mathfrak{d} \mathcal{A}_\mathfrak{d} + t_{\mathfrak{d}+1} \mathcal{A}_{\mathfrak{d}+1} + \dots + t_{n^2} \mathcal{A}_{n^2})) = f(P \exp(t_1 \mathcal{A}_1 + \dots + t_\mathfrak{d} \mathcal{A}_\mathfrak{d})).$$

We come to the concept of *vector field*  $\mathfrak{F}$  whereby at each point  $X$  of (a subset of)  $\mathbf{G}_M$  a tangent vector (= 'tangent matrix')  $\mathfrak{F}(X)$  is attached. Note that  $X^{-1}\mathfrak{F}(X)$  and  $\mathfrak{F}(X)X^{-1}$  belong to  $\mathfrak{g}_M$ , for all  $X \in \mathbf{G}_M$ . Also note that for all  $P \in \mathbf{G}_M$  the matrix  $\mathfrak{F}(P)$  is tangent to both curves  $t \mapsto P e^{tP^{-1}\mathfrak{F}(P)}$  and  $t \mapsto e^{t\mathfrak{F}(P)P^{-1}}P$  at  $t = 0$ .

### Definition 6.1

- By the *left description* of  $\mathfrak{F}$  we mean the mapping  $\mathbf{G}_M \ni X \mapsto \mathcal{F}(X) \in \mathfrak{g}_M$ , such that  $\mathfrak{F}(X) = X\mathcal{F}(X)$ , for all  $X$ . In the important special case that  $\mathcal{F}(X)$  is constant, i.e. the same matrix for all  $X$ , we name  $\mathfrak{F}$  a *left invariant* vectorfield.
- By the *right description* of  $\mathfrak{F}$  we mean the mapping  $\mathbf{G}_M \ni X \mapsto \mathcal{F}_r(X) \in \mathfrak{g}_M$ , such that  $\mathfrak{F}(X) = \mathcal{F}_r(X)X$ , for all  $X$ . Note that  $\mathcal{F}_r(X) = X\mathcal{F}(X)X^{-1}$ .

In the important special case that  $\mathcal{F}_r(X)$  is the same matrix for all  $X$  we name  $\mathfrak{F}$  a *right invariant* vectorfield.

• By an *integral curve of  $\mathfrak{F}$  through  $P$*  we mean the (local) solution  $t \mapsto X(P; t)$  of the ordinary differential equation

$$\dot{X} = \mathfrak{F}(X) = X\mathcal{F}(X), \quad \text{with } X(0) = X(P; 0) = P. \quad (6.1)$$

If we want to emphasize the dependence on  $\mathfrak{F}$ , we write  $X_{\mathfrak{F}}(P, t)$  instead of  $X(P, t)$ . Note that if  $\mathcal{F}(X) = \mathcal{A} = \text{constant}$ , the solution is  $t \mapsto P e^{t\mathcal{A}}$ .

We now want to talk about vector fields acting as operators on functions and on (other) vector fields on  $\mathbf{G}_{\mathcal{M}}$ . Locally this means something like differentiating the object under consideration along an integral curve of the vector field that is supposed to 'operate'.

**Definition 6.2**

Let  $\mathfrak{F}$  be a vector field on  $\mathbf{G}_{\mathcal{M}}$ . We use the left description  $X \mapsto X\mathcal{F}(X)$ .

Let  $f : \mathbf{G}_{\mathcal{M}} \rightarrow \mathbb{R}$  be a smooth function.

The function  $\mathfrak{F}f : \mathbf{G}_{\mathcal{M}} \rightarrow \mathbb{R}$  is defined by

$$P \mapsto \mathfrak{F}f(P) = \left. \frac{d}{dt} f(X_{\mathfrak{F}}(P; t)) \right|_{t=0} = \left. \frac{d}{dt} f(P e^{t\mathcal{F}(P)}) \right|_{t=0}. \quad (6.2)$$

Straightforward calculations lead to

**Theorem 6.3** Consider the vector fields  $\mathfrak{F}$ ,  $\mathfrak{G}$  and  $\mathfrak{H}$  with respective left descriptions  $X \mapsto X\mathcal{F}(X)$ ,  $X \mapsto X\mathcal{G}(X)$  and  $X \mapsto X\mathcal{H}(X)$ .

Let  $f : \mathbf{G}_{\mathcal{M}} \rightarrow \mathbb{R}$  be a smooth function. Let  $f_e$  denote a smooth extension of  $f$  to an open neighborhood of  $P \in \mathbf{G}_{\mathcal{M}}$  in  $\mathbb{R}^{n \times n}$ . We have

$$\begin{aligned} 1. \quad \mathfrak{F}f(P) &= \left\langle df(P) : P\mathcal{F}(P) \right\rangle = \left\langle df_e(P) : P\mathcal{F}(P) \right\rangle \\ 2. \quad \mathfrak{F}\mathfrak{G}f(P) &= \left. \frac{\partial^2}{\partial t \partial s} f(X_{\mathfrak{G}}(X_{\mathfrak{F}}(P; t); s)) \right|_{(s,t)=(0,0)} = \\ &= \left. \frac{\partial^2}{\partial t \partial s} f(P \exp(s\mathcal{G}(P)) \exp(t\mathcal{F}(P \exp(s\mathcal{G}(P)))) \right) \right|_{(s,t)=(0,0)} = \\ &= \left\langle df_e(P) : P\mathcal{F}(P)\mathcal{G}(P) \right\rangle + \left\langle df_e(P) : (P\mathcal{D}\mathcal{G}(P) : P\mathcal{F}(P)) \right\rangle + \left\langle P\mathcal{F}(P) : d^2 f_e(P) : P\mathcal{G}(P) \right\rangle. \end{aligned} \quad (6.3)$$

In the second term of the latter expression we can write

$$\mathcal{D}\mathcal{G}(P) : P\mathcal{F}(P) = \left. \frac{d}{ds} \mathcal{G}(P e^{s\mathcal{F}(P)}) \right|_{s=0} \in \mathfrak{g}_{\mathcal{M}}. \quad (6.4)$$



3.  $[\mathfrak{F}, \mathfrak{G}] = \mathfrak{F}\mathfrak{G} - \mathfrak{G}\mathfrak{F}$  acts on functions as a vectorfield. It has left description

$$\begin{aligned}
[\mathfrak{F}, \mathfrak{G}](P) &= P \left[ \mathcal{F}(P)\mathcal{G}(P) - \mathcal{G}(P)\mathcal{F}(P) + \mathbf{D}\mathcal{G}(P) : P\mathcal{F}(P) - \mathbf{D}\mathcal{F}(P) : P\mathcal{G}(P) \right] = \\
&= P \left[ \mathcal{F}(P)\mathcal{G}(P) - \mathcal{G}(P)\mathcal{F}(P) + \frac{d}{ds}\mathcal{G}(Pe^{s\mathcal{F}(P)}) \Big|_{s=0} - \frac{d}{ds}\mathcal{F}(Pe^{s\mathcal{G}(P)}) \Big|_{s=0} \right] = \\
&= P \frac{d}{ds} \left[ e^{s\mathcal{F}(P)}\mathcal{G}(Pe^{s\mathcal{F}(P)}) - e^{s\mathcal{G}(P)}\mathcal{F}(Pe^{s\mathcal{G}(P)}) \right] \Big|_{s=0} = \tag{6.5} \\
&= \frac{d}{ds} \left[ \mathfrak{G}(Pe^{s\mathcal{F}(P)}) - \mathfrak{F}(Pe^{s\mathcal{G}(P)}) \right] \Big|_{s=0}. \quad (\text{dangerous notation!})
\end{aligned}$$

4.  $[\mathfrak{F}, [\mathfrak{G}, \mathfrak{H}]] + [\mathfrak{G}, [\mathfrak{H}, \mathfrak{F}]] + [\mathfrak{H}, [\mathfrak{F}, \mathfrak{G}]] = 0$  (Jacobi Identity)

5. The left invariant vector fields are a sub Lie algebra of the Lie algebra of vector fields on  $\mathbf{G}_M$ . For left invariant vector fields there is the pointwise correspondence

$$[\mathfrak{F}, \mathfrak{G}](P) = P \left[ \mathcal{F}(P)\mathcal{G}(P) - \mathcal{G}(P)\mathcal{F}(P) \right],$$

where  $\mathcal{F}(P) = \mathcal{F} = \text{constant}$ ,  $\mathcal{G}(P) = \mathcal{G} = \text{constant}$ .

### Proof

Straightforward calculations lead to all results.

**ad 2.** Since  $\mathfrak{G}f(Q) = \frac{d}{ds}f(Qe^{s\mathcal{G}(Q)}) \Big|_{s=0}$ , we find, with  $Q = Pe^{t\mathcal{F}(P)}$ .

$$\mathfrak{G}f(Pe^{t\mathcal{F}(P)}) = \frac{d}{ds}f(Pe^{t\mathcal{F}(P)}e^{s\mathcal{G}(Pe^{t\mathcal{F}(P)})}) \Big|_{s=0} = \left\langle df_e(Pe^{t\mathcal{F}(P)}) : Pe^{t\mathcal{F}(P)}\mathcal{G}(Pe^{t\mathcal{F}(P)}) \right\rangle$$

Taking  $\frac{d}{dt} \Big|_{t=0}$  of the latter gives (6.3).

**ad 4.** This is a very general property of vector fields on manifolds. In our special case it easily follows from the associativity of successive actions of vector fields on functions. Indeed,

$$(\mathfrak{X}\mathfrak{Y})(\mathfrak{Z}f)(P) = \frac{\partial^2}{\partial t \partial s}(\mathfrak{Z}f)(G(F(P; t); s)) \Big|_{(s,t)=(0,0)} = \frac{\partial^3}{\partial t \partial s \partial r}f(H(G(F(P; t); s); r)) \Big|_{(r,s,t)=(0,0,0)},$$

since

$$(\mathfrak{Y}\mathfrak{Z})f(Q) = \frac{\partial^2}{\partial s \partial r}f(H(G(Q; s); r)) \Big|_{(r,s)=(0,0)},$$

it also follows

$$\mathfrak{X}(\mathfrak{Y}\mathfrak{Z}f)(P) = \frac{\partial}{\partial t} \frac{\partial^2}{\partial s \partial r}f(H(G(F(P; t); s); r)) \Big|_{(r,s,t)=(0,0,0)}.$$

If the Jacobi identity is written out all terms cancel because of this associativity property. ■

## 7 Parallel Transport. Connexions

Given a left invariant inner product and with an eye on Theorems 5.2-5.5 the concept of (left) covariant derivative along a curve is introduced in the following way.

**Definition 7.1** Consider a smooth parametrised curve  $t \mapsto \mathbf{K}(t) \in \mathbf{G}_{\mathcal{M}}$ , with  $t \in (a, b)$ , an open real interval. Suppose that at each point  $\mathbf{K}(t)$  of the curve a 'tangent to  $\mathbf{G}_{\mathcal{M}}$ ' vector ('tangent matrix')  $\mathfrak{Y}(t)$  is attached in a smooth way.

We introduce the curve  $t \mapsto \mathcal{Y}(t) = \mathbf{K}(t)^{-1}\mathfrak{Y}(t) \in \mathfrak{g}_{\mathcal{M}}$ , the 'left-description' of the tangent vectorfield along the curve  $t \mapsto \mathbf{K}(t) \in \mathbf{G}_{\mathcal{M}}$ .

Based on the mapping  $\mathcal{J} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$  introduce a left invariant inner product as before.

The left-covariant derivative of  $t \mapsto \mathfrak{Y}(t)$  along the curve  $t \mapsto \mathbf{K}(t)$  is defined by

$$\begin{aligned} t \mapsto \frac{\nabla_{\mathbf{K}(t)}}{dt} \mathfrak{Y}(t) &= \mathbf{K}(t) \left\{ \frac{d\mathcal{Y}(t)}{dt} + \mathcal{J}^{-1} \mathcal{P}_{\mathfrak{g}} \left[ \mathcal{J}(\mathbf{K}^{-1}(t)\dot{\mathbf{K}}(t))\mathcal{Y}^{\top}(t) - \mathcal{Y}^{\top}(t)\mathcal{J}(\mathbf{K}^{-1}(t)\dot{\mathbf{K}}(t)) \right] \right\} = \\ &= \mathbf{K}(t) \left\{ \frac{d\mathcal{Y}(t)}{dt} - \text{ad}_{\mathcal{Y}(t)}^*(\mathbf{K}^{-1}(t)\dot{\mathbf{K}}(t)) \right\}. \end{aligned} \quad (7.1)$$

**Theorem 7.2** The left-covariant derivative is left invariant. That means

$$\frac{\nabla_{\mathbf{Q}\mathbf{K}(t)}}{dt} \mathbf{Q}\mathfrak{Y}(t) = \mathbf{Q} \frac{\nabla_{\mathbf{K}(t)}}{dt} \mathfrak{Y}(t), \quad (7.2)$$

for any fixed  $\mathbf{Q} \in \mathbf{G}_{\mathcal{M}}$ .

**Proof** Straightforward substitution. ■

### Definition 7.3

• A vectorfield  $t \mapsto \mathfrak{Y}(t)$  along a curve  $t \mapsto \mathbf{K}(t) \in \mathbf{G}_{\mathcal{M}}$  is said to be left-parallel transported if

$$\frac{\nabla_{\mathbf{K}(t)}}{dt} \mathfrak{Y}(t) = 0, \quad \text{for all } t. \quad (7.3)$$

• A curve  $t \mapsto \mathbf{K}(t) \in \mathbf{G}_{\mathcal{M}}$  is said to be left geodesic if its own tangent vector  $t \mapsto \frac{d\mathbf{K}(t)}{dt}$  is left-parallel transported. That means, corresponding to (5.14),

$$\frac{\nabla_{\mathbf{K}(t)}}{dt} \dot{\mathbf{K}}(t) = 0, \quad \text{for all } t. \quad (7.4)$$

### Theorem 7.4

Suppose a left invariant inner product

$$\left\langle \mathbf{P}\mathcal{X} \cdot \mathbf{P}\mathcal{Y} \right\rangle_{\mathbf{p}} = \left\langle \mathcal{X} \cdot \mathcal{Y} \right\rangle_1 = \text{Trace}[(\mathcal{J}\mathcal{X})^{\top}\mathcal{Y}], \quad \text{with symmetric } \mathcal{J} : \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}} \quad (7.5)$$

Let  $t \mapsto \mathbf{K}(t) \in \mathbf{G}_{\mathcal{M}}$  be a parametrized curve in  $\mathbf{G}_{\mathcal{M}}$ .

- If  $t \mapsto \mathfrak{Y}(t)$  and  $t \mapsto \mathfrak{Z}(t)$  are parallel transported tangent vectors along  $t \mapsto \mathbf{K}(t)$ , then

$$t \mapsto \left\langle \mathfrak{Y}(t) \cdot \mathfrak{Z}(t) \right\rangle_{\mathbf{K}(t)} = \text{constant}. \quad (7.6)$$

- In particular the 'squared length' of a parallel transported vector is conserved

$$t \mapsto \left\langle \mathfrak{Y}(t) \cdot \mathfrak{Y}(t) \right\rangle_{\mathbf{K}(t)} = \text{constant}. \quad (7.7)$$

- In particular, if  $t \mapsto \mathbf{K}(t)$  happens to be a left geodesic it necessarily has arc-length parametrization (up to a constant).

**Proof** Because of the symmetry of  $\mathcal{J}$ ,

$$\frac{d}{dt} \text{Trace}[(\mathcal{J} \mathcal{Y}(t))^\top \mathcal{Z}(t)] = \text{Trace}[(\mathcal{J} \dot{\mathcal{Y}}(t))^\top \mathcal{Z}(t)] + \text{Trace}[(\mathcal{Y}(t))^\top \mathcal{J} \dot{\mathcal{Z}}(t)].$$

Omitting the argument  $t$  this becomes, with (7.1) put equal to zero,

$$-\text{Trace}[(\mathcal{J}(\mathbf{K}^{-1}\dot{\mathbf{K}})\mathcal{Y})^\top - \mathcal{Y}^\top \mathcal{J}(\mathbf{K}^{-1}\dot{\mathbf{K}})]^\top \mathcal{Z} - \text{Trace}[\mathcal{Y}^\top (\mathcal{J}(\mathbf{K}^{-1}\dot{\mathbf{K}})\mathcal{Z})^\top - \mathcal{Z}^\top \mathcal{J}(\mathbf{K}^{-1}\dot{\mathbf{K}})].$$

This expression equals 0 because of the properties  $\text{Trace}[AB] = \text{Trace}[BA]$ ,

$$\text{Trace}[C] = \text{Trace}[C^\top] \quad \text{and} \quad \text{Trace}\left[[Y, J]Z + J[Y, Z]\right] = 0. \quad \blacksquare$$

Next we intend to consider and to construct a class of 'directional derivatives' of a vectorfield  $\mathfrak{G}$  in the direction of a vectorfield  $\mathfrak{F}$ . Cf. Hicks [H], p.57.

### Definition 7.5

Let  $\mathfrak{F}, \mathfrak{G}, \mathfrak{H}$  be smooth vector fields on  $\mathbf{G}_{\mathcal{M}}$ . Let  $f : \mathbf{G}_{\mathcal{M}} \rightarrow \mathbb{R}$  be a smooth function.

- A recipe  $\nabla$  that builds a new vectorfield  $\nabla_{\mathfrak{F}}\mathfrak{G}$  on  $\mathbf{G}_{\mathcal{M}}$  from any pair  $\mathfrak{F}, \mathfrak{G}$  is named a *Koszul connexion* on  $\mathbf{G}_{\mathcal{M}}$  if the following properties hold

$$\left\{ \begin{array}{l} \nabla_{\mathfrak{F}}(\mathfrak{G} + \mathfrak{H}) = \nabla_{\mathfrak{F}}\mathfrak{G} + \nabla_{\mathfrak{F}}\mathfrak{H} \\ \nabla_{(\mathfrak{F}+\mathfrak{G})}\mathfrak{H} = \nabla_{\mathfrak{F}}\mathfrak{H} + \nabla_{\mathfrak{G}}\mathfrak{H} \\ \nabla_{(f\mathfrak{F})}\mathfrak{G} = f\nabla_{\mathfrak{F}}\mathfrak{G} \\ \nabla_{\mathfrak{F}}(f\mathfrak{G}) = (\mathfrak{F}f)\mathfrak{G} + f\nabla_{\mathfrak{F}}\mathfrak{G} \end{array} \right. \quad (7.8)$$

Here, as before, the directional derivative  $\mathfrak{F}f(\mathbf{P})$  of a function  $f$  at  $\mathbf{P}$ , is given by

$$\mathfrak{F}f(\mathbf{P}) = df(\mathbf{P}) \cdot \mathfrak{F}(\mathbf{P}) = \left. \frac{d}{dt} f(\mathbf{P}e^{t\mathfrak{F}(\mathbf{P})}) \right|_{t=0} \in \mathbb{R}.$$

- A connexion  $\nabla$  is named *torsion free* if for all  $\mathfrak{F}, \mathfrak{G}$ ,

$$\nabla_{\mathfrak{F}}\mathfrak{G} - \nabla_{\mathfrak{G}}\mathfrak{F} = [\mathfrak{F}, \mathfrak{G}], \quad (7.9)$$

the Lie product of  $\mathfrak{F}$  and  $\mathfrak{G}$ .

- A connexion  $\nabla$  is said to *preserve the inner product*  $P \mapsto \left\langle \cdot \right\rangle_P$  on the tangent vectors at  $P \in \mathbf{G}_{\mathcal{M}}$ , if

$$\mathfrak{F} \left\langle \mathfrak{G} \cdot \mathfrak{H} \right\rangle_P = \left\langle \nabla_{\mathfrak{F}} \mathfrak{G}(P) \cdot \mathfrak{H}(P) \right\rangle_P + \left\langle \mathfrak{G}(P) \cdot \nabla_{\mathfrak{F}} \mathfrak{H}(P) \right\rangle_P. \quad (7.10)$$

for any triple of vectorfields  $\mathfrak{F}, \mathfrak{G}, \mathfrak{H}$  at all  $P \in \mathbf{G}_{\mathcal{M}}$ .

- If the properties (7.8), (7.9), (7.10) all hold true, the connexion  $\nabla$  is named a *Riemannian Connexion*.
- The vectorfield  $\mathfrak{T}_{\nabla}$  associated with a given connexion  $\nabla$  and any pair  $\mathfrak{F}, \mathfrak{G}$ , defined by

$$\mathfrak{T}_{\nabla}(\mathfrak{F}, \mathfrak{G}) = \nabla_{\mathfrak{F}} \mathfrak{G} - \nabla_{\mathfrak{G}} \mathfrak{F} - [\mathfrak{F}, \mathfrak{G}], \quad (7.11)$$

is named *Torsion tensor*. Remind that  $\mathfrak{T}_{\nabla}$  being a *tensorfield* means that for all functions  $f, g$  on  $\mathbf{G}_{\mathcal{M}}$  and all vectorfields  $\mathfrak{F}, \mathfrak{G}$  one has  $\mathfrak{T}_{\nabla}(f\mathfrak{F}, g\mathfrak{G}) = fg\mathfrak{T}_{\nabla}(\mathfrak{F}, \mathfrak{G})$ .  $\blacksquare$

### Theorem 7.6

Let  $\mathfrak{F}, \mathfrak{G}$  be vectorfields with respective left-description  $X \mapsto X\mathcal{F}(X)$  and  $X \mapsto X\mathcal{G}(X)$ .

Let  $\mathcal{B}(\cdot, \cdot) : \mathfrak{g}_{\mathcal{M}} \times \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$  be any fixed bilinear map.

Define the vector field  $\nabla_{\mathfrak{F}} \mathfrak{G}$  on  $\mathbf{G}_{\mathcal{M}}$  by

$$\nabla_{\mathfrak{F}} \mathfrak{G}(X) = X \left( \left. \frac{d}{dt} \mathcal{G}(Xe^{t\mathcal{F}(X)}) \right|_{t=0} + \mathcal{B}(\mathcal{F}(X), \mathcal{G}(X)) \right). \quad (7.12)$$

a. Expression (7.12) defines a connexion on  $\mathbf{G}_{\mathcal{M}}$ .

b. The connexion (7.12) is left translation invariant.

That means, with  $(L_Q \mathfrak{F})(X) = X\mathcal{F}(Q^{-1}X)$ ,

$$\nabla_{(L_Q \mathfrak{F})} (L_Q \mathfrak{G})(X) = L_Q \left( \nabla_{\mathfrak{F}} \mathfrak{G} \right) (X) = X(\nabla_{\mathfrak{F}} \mathfrak{G})(Q^{-1}X),$$

for all  $X \in \mathbf{G}_{\mathcal{M}}$ .

c. All left invariant connexions on  $\mathbf{G}_{\mathcal{M}}$  are of the form (7.12).

d. If one takes  $\mathcal{B}(\mathcal{F}(X), \mathcal{G}(X)) = \mathcal{B}_{\mathfrak{g}}(\mathcal{F}(X)\mathcal{G}(X))$ , with any  $\mathcal{B}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{g}_{\mathcal{M}}$  such that  $\mathcal{B}_{\mathfrak{g}} \mathcal{X} = \mathcal{X}$ , for all  $\mathcal{X} \in \mathfrak{g}_{\mathcal{M}}$ , the obtained connexion can be written

$$\nabla_{\mathfrak{F}} \mathfrak{G}(X) = X \mathcal{B}_{\mathfrak{g}} X^{-1} \left. \frac{d}{dt} \left( X e^{t\mathcal{F}(X)} \mathcal{G}(X e^{t\mathcal{F}(X)}) \right) \right|_{t=0} = X \mathcal{B}_{\mathfrak{g}} X^{-1} \left. \frac{d}{dt} \left( \mathfrak{G}(X e^{t\mathcal{F}(X)}) \right) \right|_{t=0}. \quad (7.13)$$

### Proof

**ad a.** Use, as before  $\left. \frac{d}{dt} \mathcal{G}(X e^{t\mathcal{F}(X)}) \right|_{t=0} = D\mathcal{G}(X)X\mathcal{F}(X)$ , which is the derivative of  $\mathcal{G} : \mathbf{G}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$  at  $X \in \mathbf{G}_{\mathcal{M}}$  applied to the tangent vector  $X\mathcal{F}(X)$  at  $X \in \mathbf{G}_{\mathcal{M}}$ , cf. (6.4).

All properties (7.8) follow by straightforward calculation.

**ad b.** Compare

$$\left(\nabla_{(L_Q \mathfrak{F})}(L_Q \mathfrak{G})\right)(X) = X \left( D\mathcal{G}(Q^{-1}X):Q^{-1}X \mathcal{F}(Q^{-1}X) + \mathcal{B}(\mathcal{F}(Q^{-1}X), \mathcal{G}(Q^{-1}X)) \right),$$

and

$$L_Q \left( \nabla_{\mathfrak{F}} \mathfrak{G} \right)(X) = \left( \nabla_{\mathfrak{F}} \mathfrak{G} \right)(Q^{-1}X) = X \left( D\mathcal{G}(Q^{-1}X):Q^{-1}X \mathcal{F}(Q^{-1}X) + \mathcal{B}(\mathcal{F}(Q^{-1}X), \mathcal{G}(Q^{-1}X)) \right).$$

Note: If  $L_Q$  is applied to the vector field  $X \mapsto X\mathcal{B}(\mathcal{F}(X), \mathcal{G}(X)) = X\mathcal{B}(X^{-1}\mathfrak{F}(X), X^{-1}\mathfrak{G}(X))$ , one finds

$$X \mapsto X\mathcal{B}(X^{-1}L_Q\mathfrak{F}(X), X^{-1}L_Q\mathfrak{G}(X)) = X\mathcal{B}(\mathcal{F}(Q^{-1}X), \mathcal{G}(Q^{-1}X)).$$

**ad c.** Let  $\nabla'$  be any other left invariant connexion and consider the difference  $\nabla - \nabla'$ . Because of the properties (7.8) we have for all vectorfields  $\mathfrak{F}, \mathfrak{G}$  and all smooth functions  $f, g$ ,

$$\nabla_{(f\mathfrak{F})}(g\mathfrak{G}) - \nabla'_{(f\mathfrak{F})}(g\mathfrak{G}) = fg(\nabla_{\mathfrak{F}}\mathfrak{G} - \nabla'_{\mathfrak{F}}\mathfrak{G}).$$

This means that, together with the other properties (7.8) of  $\nabla$ , the function  $\mathfrak{F}, \mathfrak{G} \mapsto \mathfrak{R}(\mathfrak{F}, \mathfrak{G}) = (\nabla_{\mathfrak{F}}\mathfrak{G} - \nabla'_{\mathfrak{F}}\mathfrak{G})$  is a vector valued left invariant 2-tensorfield.

It can be written  $\mathfrak{R}(\mathfrak{F}, \mathfrak{G})(X) = X\mathcal{R}(X; \mathcal{F}(X), \mathcal{G}(X))$  where  $\mathcal{R} : \mathbf{G}_{\mathcal{M}} \times \mathfrak{g}_{\mathcal{M}} \times \mathfrak{g}_{\mathcal{M}} \rightarrow \mathfrak{g}_{\mathcal{M}}$  which is bi-linear in the 2nd and 3rd arguments. From the supposed left-translation invariance of both  $\nabla$  and  $\nabla'$ , it follows that  $(L_Q\mathfrak{R}(\mathfrak{F}, \mathfrak{G}))(X) = X\mathcal{R}(Q^{-1}X; \mathcal{F}(Q^{-1}X), \mathcal{G}(Q^{-1}X))$  should equal  $\mathfrak{R}(L_Q\mathfrak{F}, L_Q\mathfrak{G})(X) = X\mathcal{R}(X; \mathcal{F}(Q^{-1}X), \mathcal{G}(Q^{-1}X))$ .

This can only happen if  $\mathcal{R}$  is constant in its 1st argument. Therefore  $\nabla$  and  $\nabla'$  can differ only because of a distinct choice of  $\mathcal{B}(\cdot, \cdot)$ . ■

**Remark** The expression (7.13) looks very much like the historical connexions on surfaces in  $\mathbb{R}^n$  defined via projections on tangent planes to those surfaces. Especially if one takes  $\mathcal{B}_{\mathfrak{g}} = \mathcal{P}_{\mathfrak{g}}$  or, more general, any 'skew' projection on  $\mathfrak{g}_{\mathcal{M}}$ , the linear mapping  $X\mathcal{B}_{\mathfrak{g}}X^{-1}$  is a projection which maps a possibly protruding vector into a tangent vector. It 'chops off' the protruding part of  $\partial_t\mathfrak{G}(Xe^{t\mathcal{F}(X)})$ , so to say.

**Definition 7.7** Consider the connexion  $\nabla$  from (7.12).

- i. The (with  $\nabla$ ) *associated covariant derivative* of  $\mathfrak{G}$  along a curve  $t \mapsto K(t)$  is defined by

$$t \mapsto \frac{\nabla_{K(t)}}{dt} \mathfrak{G}(K(t)) = K(t) \left\{ \frac{d}{dt} \mathcal{G}(K(t)) + \mathcal{B}(K^{-1}(t)\dot{K}(t), \mathcal{G}(K(t))) \right\}. \quad (7.14)$$

- i'. If  $\mathcal{B}(\mathcal{X}, \mathcal{Y}) = \mathcal{B}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y})$ , for  $\mathcal{X}, \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}}$ , the latter can be written

$$t \mapsto K(t)\mathcal{B}_{\mathfrak{g}}K^{-1}(t) \frac{d}{dt} \{K(t)\mathcal{G}(K(t))\}. \quad (7.15)$$

ii. The (with  $\nabla$ ) *associated geodesics* are the curves  $t \mapsto \mathbf{K}(t)$  which satisfy

$$t \mapsto \frac{\nabla_{\mathbf{K}(t)} \dot{\mathbf{K}}(t)}{dt} = \mathbf{K}(t) \left\{ \frac{d}{dt} \mathbf{K}^{-1}(t) \dot{\mathbf{K}}(t) + \mathbf{B}(\mathbf{K}^{-1}(t) \dot{\mathbf{K}}(t), \mathbf{K}^{-1}(t) \dot{\mathbf{K}}(t)) \right\} = 0. \quad (7.16)$$

**Remark i.** If we take, in particular,  $\mathbf{K}(t) = \mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t)$ , a curve which fits  $\mathfrak{F}$ , cf. (6.1), we find

$$\begin{aligned} \nabla_{\mathfrak{F}} \mathfrak{G}(\mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t)) &= \frac{\nabla_{\mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t)} \mathfrak{G}(\mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t))}{dt} = \\ &= \mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t) \left\{ \frac{d\mathcal{G}(\mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t))}{dt} + \mathbf{B}(\mathcal{F}(\mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t)), \mathcal{G}(\mathbf{X}_{\mathfrak{F}}(\mathbf{P}, t))) \right\}. \end{aligned} \quad (7.17)$$

**Remark ii.** A special case of (7.16) is given by (7.1). Note that only the 'symmetric part' of  $\mathbf{B}$  plays a role in the geodesic equation.

**Remark iii.** The exponential curves  $t \mapsto \mathbf{P}e^{t\mathbf{A}}$  or  $t \mapsto e^{t\mathbf{B}}\mathbf{Q}$  are (the only) geodesics iff  $\mathbf{B}$  is anti-symmetric (or vanishes).

We now gather conditions on  $\mathbf{B}$  such that the connexion  $\nabla$  has special properties as mentioned in Definition 7.5.

**Theorem 7.8** *Let  $\mathcal{J}, \mathcal{K}, \mathcal{L} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  be linear mappings. Require  $\mathcal{J}^* = \mathcal{J}$ , with respect to the standard inner product on  $\mathbb{R}^{n \times n}$ . Suppose  $\mathcal{J}$  bijective and, moreover,  $\mathcal{J}(\mathfrak{g}_{\mathcal{M}}) = \mathfrak{g}_{\mathcal{M}}$ , where  $\mathfrak{g}_{\mathcal{M}}$  is the Lie Algebra of a Matrix Lie Group  $\mathbf{G}_{\mathcal{M}} \subset \mathbb{R}^{n \times n}$ . Let  $\mathcal{P}_{\mathfrak{g}} : \mathbb{R}^{n \times n} \rightarrow \mathfrak{g}_{\mathcal{M}}$  be the orthogonal projection.*

*Let the (symmetric) inner product  $\langle \cdot \rangle_1$  on  $\mathfrak{g}_{\mathcal{M}}$  be given by  $\langle \mathcal{X} \cdot \mathcal{Y} \rangle_1 = \text{Trace}[(\mathcal{J}\mathcal{X})^\top \mathcal{Y}]$ , and let it be extended to a left invariant inner product on the tangent spaces of  $\mathbf{G}_{\mathcal{M}}$ .*

*Consider the connexion  $\nabla$  as given by (7.12).*

1.  $\nabla$  is torsionfree if, for all  $\mathcal{X}, \mathcal{Y} \in \mathfrak{g}_{\mathcal{M}}$

$$\mathbf{B}(\mathcal{X}, \mathcal{Y}) = \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) + \mathcal{S}(\mathcal{X}, \mathcal{Y}), \quad (7.18)$$

with  $\mathcal{S}(\mathcal{X}, \mathcal{Y}) = \mathcal{S}(\mathcal{Y}, \mathcal{X})$ .

2.  $\nabla$  preserves the inner product  $\langle \cdot \rangle_{\mathbf{P}}$  if for all  $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{g}_{\mathcal{M}}$

$$\langle \mathbf{B}(\mathcal{X}, \mathcal{Y}) \cdot \mathcal{Z} \rangle_1 + \langle \mathcal{Y} \cdot \mathbf{B}(\mathcal{X}, \mathcal{Z}) \rangle_1 = 0, \quad (7.19)$$

*This happens to be if, for any choice  $\mathcal{K}, \mathcal{L} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ ,*

$$\mathbf{B}(\mathcal{X}, \mathcal{Y}) = \mathcal{J}^{-1} \mathcal{P}_{\mathfrak{g}} \mathcal{K}^* [\mathcal{L}\mathcal{X}, (\mathcal{K}\mathcal{Y})^\top], \quad (7.20)$$

*or, with any  $\mathcal{A} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}_-^{n \times n}$  (= skew-symmetric matrices),*

$$\mathbf{B}(\mathcal{X}, \mathcal{Y}) = \mathcal{J}^{-1} \mathcal{P}_{\mathfrak{g}} \mathcal{K}^* ((\mathcal{A}\mathcal{X})(\mathcal{K}\mathcal{Y})). \quad (7.21)$$

*Also linear combination of such expressions will do.*

3.  $\nabla$  is a Riemannian Connexion if  $\mathbf{B}(\mathcal{X}, \mathcal{Y}) = \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) + \mathbf{S}(\mathcal{X}, \mathcal{Y})$ , where for all  $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{g}_{\mathcal{M}}$ ,  $\mathbf{S}$  satisfies

$$\begin{aligned} \left\langle \mathbf{S}(\mathcal{X}, \mathcal{Y}) \cdot \mathcal{Z} \right\rangle_1 &= -\frac{1}{2} \left\{ \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y} + \mathcal{Y}\mathcal{X}) \cdot \mathcal{Z} \right\rangle_1 + \right. \\ &\quad \left. + \left\langle (\mathcal{X}\mathcal{Z} - \mathcal{Z}\mathcal{X}) \cdot \mathcal{Y} \right\rangle_1 + \left\langle (\mathcal{Y}\mathcal{Z} - \mathcal{Z}\mathcal{Y}) \cdot \mathcal{X} \right\rangle_1 \right\}. \end{aligned} \quad (7.22)$$

This amounts to

$$\begin{aligned} \mathbf{B}(\mathcal{X}, \mathcal{Y}) &= \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) - \frac{1}{2} \mathcal{J}^{-1} \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y} + \mathcal{Y}\mathcal{X}) - \frac{1}{2} \text{ad}_{\mathcal{X}}^* \mathcal{Y} - \frac{1}{2} \text{ad}_{\mathcal{Y}}^* \mathcal{X} \\ &= \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) - \frac{1}{2} \mathcal{J}^{-1} \mathcal{P}_{\mathfrak{g}} \{ (\mathcal{X}\mathcal{Y} + \mathcal{Y}\mathcal{X}) + [\mathcal{X}^\top, \mathcal{J}\mathcal{Y}] + [\mathcal{Y}^\top, \mathcal{J}\mathcal{X}] \}. \end{aligned} \quad (7.23)$$

### Proof

**ad 1.** Calculate  $\nabla_{\mathfrak{F}} \mathfrak{G} - \nabla_{\mathfrak{G}} \mathfrak{F}$  as defined by (7.12) and compare with the 2nd line of (6.5). Note that  $\mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) - \mathcal{P}_{\mathfrak{g}}(\mathcal{Y}\mathcal{X}) = \mathcal{X}\mathcal{Y} - \mathcal{Y}\mathcal{X}$ . The contributions of  $\mathbf{S}$  cancel out because of its symmetry.

**ad 2.** Compare

$$\begin{aligned} \mathfrak{F} \left\langle \mathfrak{G} \cdot \mathfrak{H} \right\rangle_{\mathfrak{P}} &= \frac{d}{dt} \left\langle \mathcal{G}(\text{P}e^{t\mathcal{F}(\text{P})}) \cdot \mathcal{H}(\text{P}e^{t\mathcal{F}(\text{P})}) \right\rangle_1 \Big|_{t=0} = \\ &= \left\langle \frac{d}{dt} \mathcal{G}(\text{P}e^{t\mathcal{F}(\text{P})}) \Big|_{t=0} \cdot \mathcal{H}(\text{P}) \right\rangle_1 + \left\langle \mathcal{G}(\text{P}) \cdot \frac{d}{dt} \mathcal{H}(\text{P}e^{t\mathcal{F}(\text{P})}) \Big|_{t=0} \right\rangle_1, \end{aligned}$$

with  $\left\langle \nabla_{\mathfrak{F}} \mathfrak{G}(\text{P}) \cdot \mathfrak{H}(\text{P}) \right\rangle_{\mathfrak{P}} + \left\langle \mathfrak{G}(\text{P}) \cdot \nabla_{\mathfrak{F}} \mathfrak{H}(\text{P}) \right\rangle_{\mathfrak{P}}$ . It is clear that equality occurs iff

$$\left\langle \mathbf{B}(\mathcal{F}(\text{P}), \mathcal{G}(\text{P})) \cdot \mathcal{H}(\text{P}) \right\rangle_1 + \left\langle \mathcal{G}(\text{P}) \cdot \mathbf{B}(\mathcal{F}(\text{P}), \mathcal{H}(\text{P})) \right\rangle_1 = 0.$$

In order to check (7.20), calculate with  $\mathcal{J} \mathcal{P}_{\mathfrak{g}} \mathcal{J}^{-1} \mathcal{P}_{\mathfrak{g}} = \mathcal{P}_{\mathfrak{g}}$ ,

$$\begin{aligned} \left\langle \mathbf{B}(\mathcal{X}, \mathcal{Y}) \cdot \mathcal{Z} \right\rangle_1 &= \text{Trace} [ (\mathcal{J} \mathbf{B}(\mathcal{X}, \mathcal{Y}))^\top \mathcal{Z} ] = \text{Trace} [ (\mathcal{P}_{\mathfrak{g}} \mathcal{K}^* [\mathcal{L}\mathcal{X}, (\mathcal{K}\mathcal{Y})^\top])^\top \mathcal{Z} ] = \\ &= \text{Trace} [ (\mathcal{K}\mathcal{Y})(\mathcal{L}\mathcal{X})^\top (\mathcal{K}\mathcal{Z}) - (\mathcal{L}\mathcal{X})^\top (\mathcal{K}\mathcal{Y})(\mathcal{K}\mathcal{Z}) ]. \end{aligned}$$

We find  $\left\langle \mathcal{Y} \cdot \mathbf{B}(\mathcal{X}, \mathcal{Z}) \right\rangle_1$  by exchanging  $\mathcal{Y}, \mathcal{Z}$  in the previous. Add the two expressions and apply the rule  $\text{Trace}[AB] = \text{Trace}[BA]$  to arrive at the result.

**ad 3.** We have to find *symmetric*  $\mathbf{S}$  such that (7.18) satisfies (7.19). That means we have to solve the tensor  $\mathbf{S}$  from

$$\left\langle \mathbf{S}(\mathcal{X}, \mathcal{Y}) \cdot \mathcal{Z} \right\rangle_1 + \left\langle \mathcal{Y} \cdot \mathbf{S}(\mathcal{X}, \mathcal{Z}) \right\rangle_1 = -\left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) \cdot \mathcal{Z} \right\rangle_1 - \left\langle \mathcal{Y} \cdot \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Z}) \right\rangle_1.$$

Considering the even permutations of  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$  together with the symmetry of  $\mathbf{S}$  and the symmetry of the inner product, we find

$$2 \left\langle \mathbf{S}(\mathcal{X}, \mathcal{Y}) \cdot \mathcal{Z} \right\rangle_1 = -\left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y}) \cdot \mathcal{Z} \right\rangle_1 - \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{Y}\mathcal{X}) \cdot \mathcal{Z} \right\rangle_1 +$$

$$\begin{aligned}
& + \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{Z}\mathcal{X}) \cdot \mathcal{Y} \right\rangle_1 - \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Z}) \cdot \mathcal{Y} \right\rangle_1 + \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{Z}\mathcal{Y}) \cdot \mathcal{X} \right\rangle_1 - \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{Y}\mathcal{Z}) \cdot \mathcal{X} \right\rangle_1 = \\
& = - \left\langle \mathcal{P}_{\mathfrak{g}}(\mathcal{X}\mathcal{Y} + \mathcal{Y}\mathcal{X}) \cdot \mathcal{Z} \right\rangle_1 - \left\langle \text{ad}_{\mathcal{X}}\mathcal{Z} \cdot \mathcal{Y} \right\rangle_1 - \left\langle \text{ad}_{\mathcal{Y}}\mathcal{Z} \cdot \mathcal{X} \right\rangle_1.
\end{aligned}$$

With Lemma 5.3, which says  $\text{ad}_{\mathcal{A}}^*\mathcal{B} = \mathcal{J}^{-1}[\mathcal{A}^\top, \mathcal{J}\mathcal{B}]$ , the result follows.  $\blacksquare$

### Example 7.9

- If  $\mathsf{X} \mapsto \mathfrak{A}(\mathsf{X}) = \mathsf{X}\mathcal{A}$ ,  $\mathsf{X} \mapsto \mathfrak{B}(\mathsf{X}) = \mathsf{X}\mathcal{B}$  are left invariant vector fields then

$$\nabla_{\mathfrak{A}}\mathfrak{B}(\mathsf{X}) = \mathsf{X}\mathcal{B}(\mathcal{A}, \mathcal{B}),$$

which is another left invariant vector field.

- Let  $\{\mathcal{A}_k\}_{k=1}^{\mathfrak{d}} \subset \mathfrak{g}_{\mathcal{M}}$  be a basis. Let  $\mathsf{X} \mapsto \mathfrak{A}_k(\mathsf{X}) = \mathsf{X}\mathcal{A}_k$ ,  $1 \leq k \leq \mathfrak{d}$  be the corresponding left invariant vector fields. Expand  $\mathfrak{F} = f^k\mathfrak{A}_k$  and  $\mathfrak{G} = g^\ell\mathfrak{A}_\ell$ . Then

$$\nabla_{\mathfrak{F}}\mathfrak{G}(\mathsf{X}) = f^k(\mathsf{X})(\mathfrak{A}_k g^\ell)(\mathsf{X})\mathfrak{A}_\ell(\mathsf{X}) + f^k(\mathsf{X})g^\ell(\mathsf{X})\mathcal{B}(\mathcal{A}_k, \mathcal{A}_\ell). \quad (7.24)$$

Here  $(\mathfrak{A}_k g^\ell)(\mathsf{X}) = \left. \frac{d}{dt} g^\ell(\mathsf{X}e^{t\mathcal{A}_k}) \right|_{t=0}$ , as before.



## A Appendix. Note on the symmetric top

The co-moving coordinate system of the mass constellation in section 1 can be chosen such that the center of mass lies at  $h\underline{e}_3 = \text{column}[0, 0, h]$  and  $J = \text{diag}[J_1, J_2, J_3]$ .

The torque, because of gravity, is

$$\underline{\tau} = -mgh\{\mathbf{R}(t)\underline{e}_3 \times \underline{e}_3\} \quad (\text{A.1})$$

### The Euler angles

Rotations around the  $z$ -axis and  $x$ -axis are denoted, respectively, by

$$\mathbf{Z}(\varphi) = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{X}(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}.$$

The representation of a dense open set of orthogonal matrices in terms of *Euler angles* is now given by

$$\begin{aligned} \mathbf{R}(\varphi, \theta, \psi) &= \mathbf{Z}(\varphi)\mathbf{X}(\theta)\mathbf{Z}(\psi) = \\ &= \begin{bmatrix} \cos \varphi \cos \psi - \sin \varphi \cos \theta \sin \psi & -\cos \varphi \sin \psi - \sin \varphi \cos \theta \cos \psi & \sin \varphi \sin \theta \\ \sin \varphi \cos \psi + \cos \varphi \cos \theta \sin \psi & -\sin \varphi \sin \psi + \cos \varphi \cos \theta \cos \psi & -\cos \varphi \sin \theta \\ \sin \theta \sin \psi & \sin \theta \cos \psi & \cos \theta \end{bmatrix} \end{aligned} \quad (\text{A.2})$$

with  $0 < \theta < \pi$ ,  $0 < \varphi < 2\pi$ ,  $0 < \psi < 2\pi$ . For any given rotation matrix the values of  $\theta, \varphi, \psi$  can be easily obtained, successively. If it happens that  $0 < \theta < \pi$ , then  $\varphi$  and  $\psi$  can be uniquely read off from the 3rd row and 3rd column of the given orthogonal matrix. Note that the Euler angles do NOT provide a local coordinate system of  $\mathbf{SO}(3)$  near the unit element  $\mathbf{I}$ .

A straightforward calculation shows that, in case the angles depend on a real variable  $t$ ,

$$\begin{aligned} \mathcal{A}(t) &= \mathbf{R}^{-1}(\varphi(t), \theta(t), \psi(t)) \frac{d}{dt} \mathbf{R}(\varphi(t), \theta(t), \psi(t)) = \\ &= \begin{bmatrix} 0 & -\cos \theta & \sin \theta \cos \psi \\ \cos \theta & 0 & -\sin \theta \sin \psi \\ -\sin \theta \cos \psi & \sin \theta \sin \psi & 0 \end{bmatrix} \frac{d\varphi}{dt} + \\ &\quad + \begin{bmatrix} 0 & 0 & -\sin \psi \\ 0 & 0 & -\cos \psi \\ \sin \psi & \cos \psi & 0 \end{bmatrix} \frac{d\theta}{dt} + \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{d\psi}{dt}. \end{aligned} \quad (\text{A.3})$$

This corresponds, cf. (1.4), to the relative angular velocity

$$\underline{\alpha} = \begin{bmatrix} \sin \theta \sin \psi \\ \sin \theta \cos \psi \\ \cos \theta \end{bmatrix} \frac{d\varphi}{dt} + \begin{bmatrix} \cos \psi \\ -\sin \psi \\ 0 \end{bmatrix} \frac{d\theta}{dt} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \frac{d\psi}{dt}. \quad (\text{A.4})$$

For possible later reference we mention

$$\begin{aligned} \frac{d\alpha}{dt} = & \begin{bmatrix} \sin \theta \sin \psi \\ \sin \theta \cos \psi \\ \cos \theta \end{bmatrix} \frac{d^2\varphi}{dt^2} + \begin{bmatrix} \cos \psi \\ -\sin \psi \\ 0 \end{bmatrix} \frac{d^2\theta}{dt^2} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \frac{d^2\psi}{dt^2} + \\ & + \begin{bmatrix} \cos \theta \sin \psi \\ \cos \theta \cos \psi \\ -\sin \theta \end{bmatrix} \frac{d\theta}{dt} \frac{d\varphi}{dt} + \begin{bmatrix} \sin \theta \cos \psi \\ -\sin \theta \sin \psi \\ 0 \end{bmatrix} \frac{d\psi}{dt} \frac{d\varphi}{dt} + \begin{bmatrix} -\sin \psi \\ -\cos \psi \\ 0 \end{bmatrix} \frac{d\psi}{dt} \frac{d\theta}{dt}. \end{aligned} \quad (\text{A.5})$$

With the torque (A.1),  $\kappa = mgh$ , and  $\mathbf{F} = \mathbf{R}$  equation (1.6) reads

$$\underline{\alpha}(t) \times J\underline{\alpha}(t) + J\dot{\underline{\alpha}}(t) = -\kappa\{\underline{e}_3 \times \mathbf{R}^{-1}(t)\underline{e}_3\}. \quad (\text{A.6})$$

Inner multiplication with, respectively,  $\mathbf{R}^{-1}\underline{e}_3$ ,  $\underline{\alpha}$  and  $\underline{e}_3$ , leads to

$$\begin{cases} \frac{d}{dt}[(\mathbf{R}J\underline{\alpha} \cdot \underline{e}_3)] = 0, \\ \frac{d}{dt}[\frac{1}{2}(J\underline{\alpha} \cdot \underline{\alpha}) + \kappa \cos \theta] = 0, \\ \frac{d}{dt}[(J\underline{\alpha} \cdot \underline{e}_3)] + (\underline{\alpha} \times J\underline{\alpha} \cdot \underline{e}_3) = 0. \end{cases} \quad (\text{A.7})$$

Indeed, with (1.1)-(1.3),

- $\frac{d}{dt}(\mathbf{R}J\underline{\alpha} \cdot \underline{e}_3) = (\dot{\mathbf{R}}J\underline{\alpha} \cdot \underline{e}_3) + (\mathbf{R}J\dot{\underline{\alpha}} \cdot \underline{e}_3) = (\mathbf{R}(\mathbf{R}^{-1}\dot{\mathbf{R}})J\underline{\alpha} \cdot \underline{e}_3) + (\mathbf{R}J\dot{\underline{\alpha}} \cdot \underline{e}_3) =$   
 $= ((\underline{\alpha} \times J\underline{\alpha} + J\dot{\underline{\alpha}}) \cdot \mathbf{R}^{-1}\underline{e}_3).$
- $(\underline{e}_3 \times \mathbf{R}^{-1}\underline{e}_3 \cdot \underline{\alpha}) = (\underline{\alpha} \times \underline{e}_3 \cdot \mathbf{R}^{-1}\underline{e}_3) = (\mathbf{R}^{-1}\dot{\mathbf{R}}\underline{e}_3 \cdot \mathbf{R}^{-1}\underline{e}_3) = (\dot{\mathbf{R}}\underline{e}_3 \cdot \underline{e}_3) = \frac{d}{dt}(\mathbf{R}\underline{e}_3 \cdot \underline{e}_3) = \frac{d}{dt} \cos \theta.$

The 3rd equation in (A.7) becomes a conservation law only if  $(\underline{\alpha} \times J\underline{\alpha} \cdot \underline{e}_3) = 0$  for all  $\underline{\alpha}$ . This happens precisely if  $J_1 = J_2$ . Then the 3rd component of the 'internal angular momentum' is conserved. We assume  $J_1 = J_2$ . From (A.7) and  $\underline{\alpha}$  expressed in Euler angles, (A.4), we find

$$\begin{cases} (\mathbf{R}J\underline{\alpha} \cdot \underline{e}_3) = (J_1 \sin^2 \theta + J_3 \cos^2 \theta)\dot{\varphi} + (J_3 \cos \theta)\dot{\psi} = M_z = \text{constant}, \\ \frac{1}{2}(J\underline{\alpha} \cdot \underline{\alpha}) + \kappa \cos \theta = \frac{1}{2}J_1(\dot{\varphi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2}J_3(\dot{\varphi} \cos \theta + \dot{\psi})^2 + \kappa \cos \theta = E = \text{constant}, \\ (J\underline{\alpha} \cdot \underline{e}_3) = (J_3 \cos \theta)\dot{\varphi} + J_3\dot{\psi} = M_3 = \text{constant}. \end{cases} \quad (\text{A.8})$$

Here  $M_z$  is the 3rd component of the objective angular momentum,  $M_3$  is the 3rd component of the relative angular momentum,  $E$  is the sum of the kinetic energy and the potential energy of the top. Those constants can be calculated from the initial conditions.

From the 1st and the 3rd identity in (A.8) we solve

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\psi} \end{bmatrix} = \frac{1}{J_1 J_3 \sin^2 \theta} \begin{bmatrix} J_3 & -J_3 \cos \theta \\ -J_3 \cos \theta & J_1 \sin^2 \theta + J_3 \cos^2 \theta \end{bmatrix} \begin{bmatrix} M_z \\ M_3 \end{bmatrix}. \quad (\text{A.9})$$

It follows

$$\dot{\varphi} = \frac{M_z - M_3 \cos \theta}{J_1 \sin^2 \theta}, \quad \dot{\psi} = \frac{M_3}{J_3} - \dot{\varphi} \cos \theta. \quad (\text{A.10})$$

After substitution of this result in the 2nd identity of (A.8), we are left with a 1st-order ordinary differential equation for  $t \mapsto \theta(t)$  which resembles the mathematical description of 'a particle in a 1-dimensional potential well',

$$\frac{d\theta}{dt} = \pm \sqrt{\frac{2}{J_1} \left[ E - \frac{M_3^2}{2J_3} - \left\{ \frac{(M_z - M_3 \cos \theta)^2}{2J_1 \sin^2 \theta} + \kappa \cos \theta \right\} \right]}. \quad (\text{A.11})$$

After separation this leads to an elliptic integral. The *nutation angle*  $t \mapsto \theta(t)$  is a periodic function. Next, by means of (A.10) the functions  $t \mapsto \varphi(t)$  and  $t \mapsto \psi(t)$  can be found by integration. For further qualitative considerations and lots of beautiful worked out examples I refer to **[LL]** Chapter VI. Note however that the 'theoretical' considerations in that chapter VI are, as in many physics books, more ritual than logical and therefore difficult to follow by mathematicians.

That chapter in that famous book is the *raison d'être* for section 1 and this Appendix.

## Acknowledgement

The author thanks Dr. Remco Duits for helpful discussions.

## References

- [G] Gantmacher, F.R.: Matrizenrechnung. DVdW Berlin 1965-1966.
- [dG] Graaf, J. de: Tensorrekening en Differentiaalmeetkunde. 2F800. Eindhoven University of Technology. 2004.
- [H] Noel J. Hicks: Notes on Differential Geometry. ISBN 0 442 034059. Van Nostrand. London 1971.
- [LL] Landau, L.D. and E.M. Lifshitz: Mechanics. Pergamon. Oxford 1969.

Eindhoven, 31 January 2013, 30 April 2015.

## PREVIOUS PUBLICATIONS IN THIS SERIES:

Number	Author(s)	Title	Month
15-17	C.J. van Duijn X. Cao I.S. Pop	Two-phase flow in porous media: dynamic capillarity and heterogeneous media	Apr. '15
15-18	S.P. Korzilius W.H.A. Schilders M.J.H. Anthonissen	An improved CSPM approximation for multi-dimensional second-order derivatives	Apr. '15
15-19	M.E. Hochstenbach A. Muhič B. Plestenjak	Jacobi-Davidson methods for polynomial two-parameter eigenvalue problems	Apr. '15
15-20	I.S. Pop J. Bogers K. Kumar	Analysis and upscaling of a reactive transport model in fractured porous media involving nonlinear a transmission condition	May '15
15-21	J. de Graaf	Geodesics and connexions on matrix Lie groups, 2015 ed.	May '15