

# Geometric Mechanics

## Part II: Rotating, Translating and Rolling

Darryl D Holm  
Mathematics Department  
Imperial College London

2nd Edition  
23 Sept 2011







To Justine, for her love, kindness and patience.  
Thanks for letting me think this was important.



# Contents

---

<b>Preface</b>	<b>xv</b>
<b>1 Galileo</b>	<b>1</b>
1.1 Principle of Galilean relativity	2
1.2 Galilean transformations	3
1.2.1 Admissible force laws for an $N$ -particle system	6
1.3 Subgroups of the Galilean transformations	8
1.3.1 Matrix representation of $SE(3)$	9
1.4 Lie group actions of $SE(3)$	11
1.5 Lie group actions of $G(3)$	12
1.5.1 Matrix representation of $G(3)$	14
1.6 Lie algebra of $SE(3)$	15
1.7 Lie algebra of $G(3)$	17
<b>2 Newton, Lagrange, Hamilton and the rigid body</b>	<b>19</b>
2.1 Newton	21
2.1.1 Newtonian form of free rigid rotation	21
2.1.2 Newtonian form of rigid-body motion	30
2.2 Lagrange	36
2.2.1 The principle of stationary action	36
2.3 Noether's theorem	39
2.3.1 Lie symmetries and conservation laws	39

## viii CONTENTS

2.3.2	Infinitesimal transformations of a Lie group	40
2.4	Lagrangian form of rigid-body motion	46
2.4.1	Hamilton–Pontryagin constrained variations	50
2.4.2	Manakov’s formulation of the $SO(n)$ rigid body	54
2.4.3	Matrix Euler–Poincaré equations	55
2.4.4	An isospectral eigenvalue problem for the $SO(n)$ rigid body	56
2.4.5	Manakov’s integration of the $SO(n)$ rigid body	58
2.5	Hamilton	60
2.5.1	Hamiltonian form of rigid-body motion	62
2.5.2	Lie–Poisson Hamiltonian rigid-body dynamics	63
2.5.3	Lie–Poisson bracket	64
2.5.4	Nambu’s $\mathbb{R}^3$ Poisson bracket	65
2.5.5	Clebsch variational principle for the rigid body	69
2.5.6	Rotating motion with potential energy	72
<b>3</b>	<b>Quaternions</b>	<b>77</b>
3.1	Operating with quaternions	78
3.1.1	Multiplying quaternions using Pauli matrices	79
3.1.2	Quaternionic conjugate	82
3.1.3	Decomposition of three-vectors	85
3.1.4	Alignment dynamics for Newton’s second law	86
3.1.5	Quaternionic dynamics of Kepler’s problem	90
3.2	Quaternionic conjugation	93
3.2.1	Cayley–Klein parameters	93
3.2.2	Pure quaternions, Pauli matrices and $SU(2)$	99
3.2.3	Tilde map: $\mathbb{R}^3 \simeq su(2) \simeq so(3)$	102
3.2.4	Dual of the tilde map: $\mathbb{R}^{3*} \simeq su(2)^* \simeq so(3)^*$	103
3.2.5	Pauli matrices and Poincaré’s sphere $\mathbb{C}^2 \rightarrow S^2$	103
3.2.6	Poincaré’s sphere and Hopf’s fibration	105



3.2.7	Coquaternions	108
<b>4</b>	<b>Adjoint and coadjoint actions</b>	<b>111</b>
4.1	Cayley–Klein dynamics for the rigid body	112
4.1.1	Cayley–Klein parameters, rigid-body dynamics	112
4.1.2	Body angular frequency	113
4.1.3	Cayley–Klein parameters	115
4.2	Actions of quaternions, Lie groups and Lie algebras	116
4.2.1	$AD$ , $Ad$ , $ad$ , $Ad^*$ and $ad^*$ actions of quaternions	117
4.2.2	$AD$ , $Ad$ , and $ad$ for Lie algebras and groups	118
4.3	Example: The Heisenberg Lie group	124
4.3.1	Definitions for the Heisenberg group	124
4.3.2	Adjoint actions: $AD$ , $Ad$ and $ad$	126
4.3.3	Coadjoint actions: $Ad^*$ and $ad^*$	127
4.3.4	Coadjoint motion and harmonic oscillations	129
<b>5</b>	<b>The special orthogonal group <math>SO(3)</math></b>	<b>131</b>
5.1	Adjoint and coadjoint actions of $SO(3)$	132
5.1.1	$Ad$ and $ad$ operations for the hat map	132
5.1.2	$AD$ , $Ad$ and $ad$ actions of $SO(3)$	133
5.1.3	Dual Lie algebra isomorphism	135
<b>6</b>	<b>Adjoint and coadjoint semidirect-product group actions</b>	<b>141</b>
6.1	Special Euclidean group $SE(3)$	142
6.2	Adjoint operations for $SE(3)$	144
6.3	Adjoint actions of $SE(3)$ 's Lie algebra	148
6.3.1	The $ad$ action of $se(3)$ on itself	148
6.3.2	The $ad^*$ action of $se(3)$ on its dual $se(3)^*$	149

## X CONTENTS

6.3.3	Left versus right	151
6.4	Special Euclidean group $SE(2)$	153
6.5	Semidirect-product group $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$	156
6.5.1	Definitions for $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$	156
6.5.2	$AD$ , $Ad$ , and $ad$ actions	158
6.5.3	$Ad^*$ and $ad^*$ actions	160
6.5.4	Coadjoint motion relation	162
6.6	Galilean group	164
6.6.1	Definitions for $G(3)$	164
6.6.2	$AD$ , $Ad$ , and $ad$ actions of $G(3)$	165
6.7	Iterated semidirect products	167
<b>7 Euler–Poincaré and Lie–Poisson equations on <math>SE(3)</math></b>		<b>169</b>
7.1	Euler–Poincaré equations for left-invariant Lagrangians under $SE(3)$	170
7.1.1	Legendre transform from $se(3)$ to $se(3)^*$	172
7.1.2	Lie–Poisson bracket on $se(3)^*$	172
7.1.3	Coadjoint motion on $se(3)^*$	173
7.2	Kirchhoff equations on $se(3)^*$	176
7.2.1	Looks can be deceiving: The heavy top	178
<b>8 Heavy-top equations</b>		<b>181</b>
8.1	Introduction and definitions	182
8.2	Heavy-top action principle	183
8.3	Lie–Poisson brackets	184
8.3.1	Lie–Poisson brackets and momentum maps	185
8.3.2	Lie–Poisson brackets for the heavy top	186
8.4	Clebsch action principle	187
8.5	Kaluza–Klein construction	188

<b>9 The Euler–Poincaré theorem</b>	<b>193</b>
9.1 Action principles on Lie algebras	194
9.2 Hamilton–Pontryagin principle	198
9.3 Clebsch approach to Euler–Poincaré	199
9.3.1 Defining the Lie derivative	201
9.3.2 Clebsch Euler–Poincaré principle	202
9.4 Lie–Poisson Hamiltonian formulation	206
9.4.1 Cotangent-lift momentum maps	207
<b>10 Lie–Poisson Hamiltonian form of a continuum spin chain</b>	<b>209</b>
10.1 Formulating continuum spin chain equations	210
10.2 Euler–Poincaré equations	212
10.3 Hamiltonian formulation	213
<b>11 Momentum maps</b>	<b>221</b>
11.1 The momentum map	222
11.2 Cotangent lift	224
11.3 Examples of momentum maps	226
11.3.1 The Poincaré sphere $S^2 \in S^3$	237
11.3.2 Overview	242
<b>12 Round, rolling rigid bodies</b>	<b>245</b>
12.1 Introduction	246
12.1.1 Holonomic versus nonholonomic	246
12.1.2 The Chaplygin ball	248
12.2 Nonholonomic Hamilton–Pontryagin variational principle	252
12.2.1 HP principle for the Chaplygin ball	256

## xii CONTENTS

12.2.2	Circular disk rocking in a vertical plane	265
12.2.3	Euler's rolling and spinning disk	268
12.3	Nonholonomic Euler–Poincaré reduction	275
12.3.1	Semidirect-product structure	276
12.3.2	Euler–Poincaré theorem	278
12.3.3	Constrained reduced Lagrangian	282
<b>A Geometrical structure of classical mechanics</b>		<b>287</b>
A.1	Manifolds	288
A.2	Motion: Tangent vectors and flows	296
A.2.1	Vector fields, integral curves and flows	297
A.2.2	Differentials of functions: The cotangent bundle	299
A.3	Tangent and cotangent lifts	300
A.3.1	Summary of derivatives on manifolds	301
<b>B Lie groups and Lie algebras</b>		<b>305</b>
B.1	Matrix Lie groups	306
B.2	Defining matrix Lie algebras	310
B.3	Examples of matrix Lie groups	312
B.4	Lie group actions	314
B.4.1	Left and right translations on a Lie group	316
B.5	Tangent and cotangent lift actions	317
B.6	Jacobi–Lie bracket	320
B.7	Lie derivative and Jacobi–Lie bracket	323
B.7.1	Lie derivative of a vector field	323
B.7.2	Vector fields in ideal fluid dynamics	325
<b>C Enhanced coursework</b>		<b>327</b>
C.1	Variations on rigid-body dynamics	328

CONTENTS **xiii**

<b>C.1.1 Two times</b>	<b>328</b>
<b>C.1.2 Rotations in complex space</b>	<b>334</b>
<b>C.1.3 Rotations in four dimensions: <math>SO(4)</math></b>	<b>338</b>
<b>C.2 <math>\mathbb{C}^3</math> oscillators</b>	<b>343</b>
<b>C.3 Momentum maps for <math>GL(n, \mathbb{R})</math></b>	<b>348</b>
<b>C.4 Motion on the symplectic Lie group <math>Sp(2)</math></b>	<b>354</b>
<b>C.5 Two coupled rigid bodies</b>	<b>359</b>
<b>D Poincaré's 1901 paper</b>	<b>363</b>
<b><i>Bibliography</i></b>	<b>367</b>
<b><i>Index</i></b>	<b>385</b>



# Preface

---

## Introduction

This is the 2nd edition of the text for a course of 33 lectures in geometric mechanics, taught annually by the author to fourth-year undergraduates in their last term in applied mathematics at Imperial College London. The text mimics the lectures, which attempt to provide an air of immediacy and flexibility in which students may achieve insight and proficiency in using one of the fundamental approaches for solving a variety of problems in geometric mechanics. This is the Euler–Poincaré approach, which uses the Lie group invariance of Hamilton’s principle to produce symmetry-reduced motion equations and reveal their geometrical meaning. It has been taught to students with various academic backgrounds from mathematics, physics and engineering.

Each chapter of the text is presented as a line of inquiry, often by asking sequences of related questions such as “What is angular velocity?”, “What is kinetic energy?”, “What is angular momentum?”, and so forth. In adopting such an inquiry-based approach, one focuses on a sequence of exemplary problems, each of whose solution facilitates taking the next step. The present text takes those steps, forgoing any attempt at mathematical rigour. Readers interested in a more rigorous approach are invited to consult some of the many works cited in the bibliography which treat the subject in that style. This book is meant to be an intermediate introduction to geometric mechanics that bridges standard textbooks and more advanced study.

**xvi** PREFACE**Prerequisites**

The prerequisites are standard for an advanced undergraduate student. Namely, the student should be familiar with the linear algebra of vectors and matrices, ordinary differential equations, multivariable calculus and have some familiarity with variational principles and canonical Poisson brackets in classical mechanics at the level of a second- or third-year undergraduate in mathematics, physics or engineering. An undergraduate background in physics is particularly helpful, because all the examples of rotating, spinning and rolling rigid bodies treated here from a geometric viewpoint would be familiar from undergraduate physics classes.

**How to read this book**

Most of the book is meant to be read in sequential order from front to back. The 120 Exercises are an important aspect of the text. These are shaded, indented and marked with ★.

**Exercise.**



Their 55 Worked Answers are indented and marked with ▲.

Key theorems, results and remarks are placed in frames.

The three appendices provide supplementary material, such as condensed summaries of the essentials of manifolds (Appendix A) and Lie groups (Appendix B) for students who may wish to acquire a bit more mathematical background. In addition, the appendices provide material for supplementary lectures that extend the course material. Examples include variants of rotating motion that depend on more than one time variable, as well as rotations in complex space



and in higher dimensions in Appendix C. The appendices also contain ideas for additional homework and exam problems that go beyond the many exercises and examples sprinkled throughout the text.

## Description of contents

Galilean relativity and the idea of a uniformly moving reference frame are explained in Chapter 1. Freely rotating motion is then treated in Chapters 2, 3 and 4, first by reviewing Newton's and Lagrange's approaches, then by following Hamilton's approach via quaternions and Cayley–Klein parameters, not Euler angles.

Hamilton's rules for multiplication of quaternions introduced the adjoint and coadjoint actions that lie at the heart of geometric mechanics. For the rotations and translations in  $\mathbb{R}^3$  studied in Chapters 5 and 6, the adjoint and coadjoint actions are both equivalent to the vector cross product. Poincaré [Po1901] opened the field of geometric mechanics by noticing that these actions define the motion generated by any Lie group.

When applied to Hamilton's principle defined on the tangent space of an arbitrary Lie group, the adjoint and coadjoint actions studied in Chapter 6 result in the Euler–Poincaré equations derived in Chapter 7. Legendre-transforming the Lagrangian in Hamilton's principle summons the Lie–Poisson Hamiltonian formulation of dynamics on a Lie group. The Euler–Poincaré equations provide the framework for all of the applications treated in this text. These applications include finite-dimensional dynamics of three-dimensional rotations and translations in the special Euclidean group  $SE(3)$ . The Euler–Poincaré problem on  $SE(3)$  recovers Kirchhoff's classic treatment in modern form of the dynamics of an ellipsoidal body moving in an incompressible fluid flow without vorticity.

The Euler–Poincaré formulation of Kirchhoff's problem on  $SE(3)$  in Chapter 7 couples rotations and translations, but it does not yet introduce potential energy. The semidirect-product struc-

## xviii PREFACE

ture of  $SE(3)$ , however, introduces the key idea for incorporating potential energy. Namely, the same semidirect-product structure is also invoked in passing from rotations of a free rigid body to rotations of a heavy top with a *fixed* point of support under gravity. Thereby, semidirect-product reduction becomes a central focus of the text.

The heavy top treated in Chapter 8 is a key example, because it introduces the dual representation of the action of a Lie algebra on a vector space. This is the diamond operation ( $\diamond$ ), by which the forces and torques produced by potential energy gradients are represented in the Euler–Poincaré framework in Chapters 9 and 10. The diamond operation is then found in Chapter 11 to lie at the heart of the standard (cotangent-lift) momentum map.

This observation reveals the relation between the results of reduction by Lie symmetry on the Lagrangian and Hamiltonian sides. Namely,

Lie symmetry reduction on the Lagrangian side produces the Euler–Poincaré equation, whose formulation on the Hamiltonian side as a Lie–Poisson equation governs the dynamics of the momentum map associated with the cotangent lift of the Lie algebra action of that Lie symmetry on the configuration manifold.

The chief purpose of this book is to explain that statement, so that it may be understood by undergraduate students in mathematics, physics and engineering.

In the Euler–Poincaré framework, the adjoint and coadjoint actions combine with the diamond operation to provide a powerful tool for investigating other applications of geometric mechanics, including nonholonomic constraints discussed in Chapter 12. In the same chapter, nonholonomic mechanics is discussed in the context of two classic problems, known as Chaplygin’s ball (an unbalanced rolling ball) and Euler’s disk (a spinning, falling, rolling flat coin). In these classic examples, the semidirect-product structure couples rotations, translations and potential energy together with the rolling constraint.

## What is new in the 2nd edition?

The organisation of the 1st edition has been preserved in the 2nd edition. However, the substance of the text has been rewritten throughout to improve the flow and enrich the development of the material. Some examples of the new improvements include the following:

- The Galilean group and the implications of Galilean relativity for Noether's theorem have been developed further.
- The role of Noether's theorem has been given added emphasis throughout, with various applications for Euler–Poincaré systems, Lie–Poisson systems and nonholonomically constrained systems, such as the rolling bodies treated in Chapter 12.
- Additional examples of adjoint and coadjoint actions of Lie groups have been worked out. These include the Heisenberg group, and the semidirect-product Lie group  $SL(2, R) \ltimes \mathbb{R}^2$ , which is the group of motions of ellipses that translate, rotate and dilate in the plane, while preserving their area.
- Manakov's approach of regarding the rigid body as an isospectral eigenvalue problem has been developed further and additional examples of its application have been given, particularly for the Euler top and the Lagrange top.
- A section has been added about coquaternions. The coquaternions comprise a representation of  $SP(2)$ , while the more well-known quaternions form a representation of  $SU(2)$ .
- Momentum maps are discussed in more depth in the 2nd edition. For example, dual pairs of momentum maps are discussed in Chapter 11 in the context of the Hopf fibration, under the reduction of  $\mathbb{C}^2$  by  $S^1$  from the right and by  $SU(2)$  from the left.
- Additional enhanced coursework has been provided, including treatments of Euler–Poincaré equations for (i) the articulated motion of two coupled rigid bodies; and (ii) geodesic

**XX** PREFACE

motion on the symplectic group  $SP(2)$ , which leads to a *Bloch–Iserles equation* [BII2006].

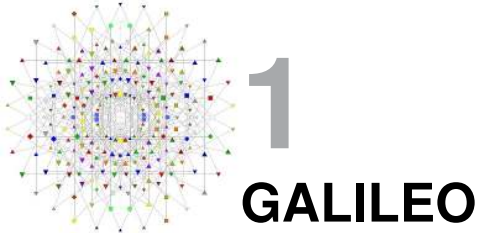
- A new English translation is provided in Appendix D of Poincaré’s famous two-page paper [Po1901]. There, Poincaré derived the Euler–Poincaré equations by using Lie’s theory of infinitesimal transformations for variations in Hamilton’s principle. Thus began modern geometric mechanics.

**Acknowledgements**

I am enormously grateful to my students, who have been partners in this endeavour, for their comments, questions and interest. I am also grateful to my friends and collaborators for their camaraderie in first forging many of these tools. We are in this together. Thanks for all the fun we’ve had. I appreciate the many helpful remarks and recommendations regarding the text from Poul Hjorth (who also drew most of the figures), David Ellis, Rossen Ivanov, Byung-soo Kim, Vakhtang Putkaradze, Peter Olver, Tudor Ratiu, Cristina Stoica, Tanya Schmah and Cesare Tronci.

In addition, I am grateful to the reviewers of the 1st edition, for their uniformly encouraging comments and critiques. Of course, the errors that still remain are all my own. I am also grateful to Tasha D’Cruz, Amy Hendrickson, and Kwong Lai Fun for their diligent cooperation with me in the process of editing and type-setting this 2nd edition.

Finally, I am grateful for three decades of camaraderie with my late friend Jerry Marsden, who inspired me in many ways.



---

### Contents

1.1	Principle of Galilean relativity	2
1.2	Galilean transformations	3
1.2.1	Admissible force laws for an $N$ -particle system	6
1.3	Subgroups of the Galilean transformations	8
1.3.1	Matrix representation of $SE(3)$	9
1.4	Lie group actions of $SE(3)$	11
1.5	Lie group actions of $G(3)$	12
1.5.1	Matrix representation of $G(3)$	14
1.6	Lie algebra of $SE(3)$	15
1.7	Lie algebra of $G(3)$	17

---

## 2 1 : GALILEO

---

## 1.1 Principle of Galilean relativity



Galileo Galilei

Principles of relativity address the problem of how events that occur in one place or state of motion are observed from another. And if events occurring in one place or state of motion look different from those in another, how should one determine the laws of motion?

Galileo approached this problem via a thought experiment which imagined observations of motion made inside a ship by people who could not see outside. He showed that the people isolated

inside a uniformly moving ship would be *unable to determine by measurements made inside it whether they were moving!*

...have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still.

– Galileo, *Dialogue Concerning the Two Chief World Systems* [Ga1632]

Galileo's thought experiment showed that a man who is below decks on a ship cannot tell whether the ship is docked or is moving uniformly through the water at constant velocity. He may observe water dripping from a bottle, fish swimming in a tank, butterflies flying, etc. Their behaviour will be just the same, whether the ship is moving or not.

**Definition 1.1.1 (Galilean transformations)** *Transformations of reference location, time, orientation or state of uniform translation at constant velocity are called Galilean transformations.*

**Definition 1.1.2 (Uniform rectilinear motion)** *Coordinate systems related by Galilean transformations are said to be in **uniform rectilinear motion** relative to each other.*

Galileo's thought experiment led him to the following principle.

**Definition 1.1.3 (Principle of Galilean relativity)** *The laws of motion are independent of reference location, time, orientation or state of uniform translation at constant velocity. Hence, these laws are invariant (i.e., they do not change their forms) under Galilean transformations.*

**Remark 1.1.1 (Two tenets of Galilean relativity)** Galilean relativity sets out two important tenets:

- It is impossible to determine who is actually at rest.
- Objects continue in uniform motion unless acted upon.

The second tenet is known as *Galileo's law of inertia*. It is also the basis for *Newton's first law of motion*.  $\square$

---

## 1.2 Galilean transformations

**Definition 1.2.1 (Galilean transformations)** *Galilean transformations of a coordinate frame consist of space-time translations, rotations and reflections of spatial coordinates, as well as Galilean "boosts" into uniform rectilinear motion.*

*In three dimensions, the Galilean transformations depend smoothly on ten real parameters, as follows:*

- *Space-time translations,*

$$g_1(\mathbf{r}, t) = (\mathbf{r} + \mathbf{r}_0, t + t_0).$$

*These possess four real parameters:  $(\mathbf{r}_0, t_0) \in \mathbb{R}^3 \times \mathbb{R}$ , for the three dimensions of space, plus time.*

#### 4 1 : GALILEO

- *Spatial rotations and reflections,*

$$g_2(\mathbf{r}, t) = (O\mathbf{r}, t),$$

for any linear **orthogonal transformation**  $O : \mathbb{R}^3 \mapsto \mathbb{R}^3$  with  $O^T = O^{-1}$ . These have three real parameters, for the three axes of rotation and reflection. Because the inverse of an orthogonal transformation is its transpose ( $O^{-1} = O^T$ ) they preserve both the lengths and relative orientations of vectors. It has two connected components corresponding to the positive and negative values of the determinant,  $\det O = \pm 1$ , which changes sign under reflections.

- *Galilean boosts into uniform rectilinear motion,*

$$g_3(\mathbf{r}, t) = (\mathbf{r} + \mathbf{v}_0 t, t).$$

These have three real parameters:  $\mathbf{v}_0 \in \mathbb{R}^3$ , for the three components of the velocity boost vector.

**Definition 1.2.2 (Group)** A **group**  $G$  is a set of elements that possesses a binary product (multiplication),  $G \times G \rightarrow G$ , such that the following properties hold:

- The product  $gh$  of  $g$  and  $h$  is associative, that is,  $(gh)k = g(hk)$ .
- An identity element exists,  $e : eg = g$  and  $ge = g$ , for all  $g \in G$ .
- The inverse operation exists,  $G \rightarrow G$ , so that  $gg^{-1} = g^{-1}g = e$ .

**Definition 1.2.3 (Lie group)** A **Lie group** is a group that depends smoothly on a set of parameters. That is, a Lie group is both a group and a smooth manifold, for which the group operation is by composition of smooth invertible functions.

**Proposition 1.2.1 (Lie group property)** Galilean transformations form a Lie group, modulo reflections.

**Proof.** Any Galilean transformation

$$g \in G(3) : \mathbb{R}^3 \times \mathbb{R} \mapsto \mathbb{R}^3 \times \mathbb{R}$$



## 1.2 GALILEAN TRANSFORMATIONS 5

may be expressed uniquely as a composition of the three basic transformations  $\{g_1, g_2, g_3\} \in G(3)$ . Consequently, the set of elements comprising the transformations  $\{g_1, g_2, g_3\} \in G(3)$  closes under the binary operation of composition. The Galilean transformations also possess an identity element  $e : eg_i = g_i = g_i e, i = 1, 2, 3$ , and each element  $g$  possesses a unique inverse  $g^{-1}$ , so that  $gg^{-1} = e = g^{-1}g$ .

These properties, plus associativity, define a group. The smooth dependence of the group of Galilean transformations on its ten parameters means that the *Galilean group*  $G(3)$  is a Lie group (except for the reflections, which are discrete, not smooth). ■

**Remark 1.2.1** Compositions of Galilean boosts and translations commute. That is,

$$g_1 g_3 = g_3 g_1 .$$

However, the order of composition does matter in Galilean transformations when rotations and reflections are involved. For example, the action of the Galilean group composition  $g_1 g_3 g_2$  on  $(\mathbf{r}, t)$  from the left is given by

$$g(\mathbf{r}, t) = (O\mathbf{r} + t\mathbf{v}_0 + \mathbf{r}_0, t + t_0) ,$$

for

$$g = g_1(\mathbf{r}_0, t_0)g_3(\mathbf{v}_0)g_2(O) =: g_1 g_3 g_2 .$$

However, the result for another composition, say  $g_1 g_2 g_3$ , would in general be different. □

**Exercise.** Write the corresponding transformations for  $g_1 g_2 g_3$ ,  $g_1 g_3 g_2$ ,  $g_2 g_1 g_3$  and  $g_3 g_2 g_1$ , showing how they depend on the order in which the rotations, boosts and translations are composed. Write the inverse transformation for each of these compositions of left actions. ★

## 6 1 : GALILEO

**Answer.** The various compositions of translations  $g_1(\mathbf{r}_0, t_0)$ , rotations  $g_2(O)$  and boosts  $g_3(\mathbf{v}_0)$  in general yield different results, as

$$g_1 g_2 g_3(\mathbf{r}, t) = \left( O(\mathbf{r} + t\mathbf{v}_0) + \mathbf{r}_0, t + t_0 \right),$$

$$g_1 g_3 g_2(\mathbf{r}, t) = \left( O\mathbf{r} + t\mathbf{v}_0 + \mathbf{r}_0, t + t_0 \right),$$

$$g_2 g_1 g_3(\mathbf{r}, t) = \left( O(\mathbf{r} + t\mathbf{v}_0 + \mathbf{r}_0), t + t_0 \right),$$

$$g_3 g_2 g_1(\mathbf{r}, t) = \left( O(\mathbf{r} + \mathbf{r}_0) + t\mathbf{v}_0, t + t_0 \right).$$

The inverses are  $(g_1 g_2 g_3)^{-1} = g_3^{-1} g_2^{-1} g_1^{-1}$ , etc. ▲

**Remark 1.2.2 (Decomposition of the Galilean group)** Because the rotations take vectors into vectors, any element of the transformations  $g_1 g_2 g_3$ ,  $g_2 g_1 g_3$  and  $g_3 g_2 g_1$  in the Galilean group may be written uniquely in the simplest form, as  $g_1 g_3 g_2$ .

Thus, any element of the Galilean group may be written uniquely as a rotation, followed by a space translation, a Galilean boost and a time translation. The latter three may be composed in any order, because they commute with each other. □

**Exercise.** What properties are preserved by the Galilean group? ★

**Answer.** The Galilean group  $G(3)$  preserves the results of measuring length and time intervals, and relative orientation in different frames of motion related to each other by Galilean transformations. ▲

### 1.2.1 Admissible force laws for an $N$ -particle system

For a system of  $N$  interacting particles, Newton's second law of motion (the law of acceleration) determines the motion resulting from

## 1.2 GALILEAN TRANSFORMATIONS 7

the force  $\mathbf{F}_j$  exerted on the  $j$ th particle by the other  $N - 1$  particles as

$$m_j \ddot{\mathbf{r}}_j = \mathbf{F}_j(\mathbf{r}_k - \mathbf{r}_l, \dot{\mathbf{r}}_k - \dot{\mathbf{r}}_l), \quad \text{with } j, k, l = 1, 2, \dots, N \text{ (no sum)}.$$

This force law is independent of reference location, time or state of uniform translation at constant velocity. It will also be independent of reference orientation and thus will be *Galilean-invariant*, provided the forces  $\mathbf{F}_j$  transform under rotations and parity reflections as *vectors*

$$m_j O \ddot{\mathbf{r}}_j = O \mathbf{F}_j = \mathbf{F}_j(O(\mathbf{r}_k - \mathbf{r}_l), O(\dot{\mathbf{r}}_k - \dot{\mathbf{r}}_l)), \quad (1.2.1)$$

for any orthogonal transformation  $O$ .

This requirement for Galilean invariance that the force in Newton's law of acceleration transforms as a vector is the reason that vectors are so important in classical mechanics.

For example, Newton's law of gravitational motion is given by

$$m_j \ddot{\mathbf{r}}_j = \sum_{k \neq j} \mathbf{F}_{jk}, \quad (1.2.2)$$

in which the gravitational forces  $\mathbf{F}_{jk}$  between  $(j, k)$  particle pairs are given by

$$\mathbf{F}_{jk} = \frac{\gamma m_j m_k}{|\mathbf{r}_{jk}|^3} \mathbf{r}_{jk}, \quad \text{with } \mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k, \quad (1.2.3)$$

and  $\gamma$  is the gravitational constant.

**Exercise.** Prove that Newton's law (1.2.2) for gravitational forces (1.2.3) is Galilean-invariant. That is, prove that Newton's law of gravitational motion takes the same form in any Galilean reference frame. ★

### 1.3 Subgroups of the Galilean transformations

**Definition 1.3.1 (Subgroup)** A *subgroup* is a subset of a group whose elements also satisfy the defining properties of a group.

**Exercise.** List the subgroups of the Galilean group that do not involve time. ★

**Answer.** The subgroups of the Galilean group that are independent of time consist of

- Spatial translations  $g_1(\mathbf{r}_0)$  acting on  $\mathbf{r}$  as  $g_1(\mathbf{r}_0)\mathbf{r} = \mathbf{r} + \mathbf{r}_0$ .
- Proper rotations  $g_2(O)$  with  $g_2(O)\mathbf{r} = O\mathbf{r}$  where  $O^T = O^{-1}$  and  $\det O = +1$ . This subgroup is called  $SO(3)$ , the *special orthogonal group* in three dimensions.
- Rotations and reflections  $g_2(O)$  with  $O^T = O^{-1}$  and  $\det O = \pm 1$ . This subgroup is called  $O(3)$ , the *orthogonal group* in three dimensions.
- Spatial translations  $g_1(\mathbf{r}_0)$  with  $\mathbf{r}_0 \in \mathbb{R}^3$  compose with proper rotations  $g_2(O) \in SO(3)$  acting on a vector  $\mathbf{r} \in \mathbb{R}^3$  as

$$E(O, \mathbf{r}_0)\mathbf{r} = g_1(\mathbf{r}_0)g_2(O)\mathbf{r} = O\mathbf{r} + \mathbf{r}_0,$$

where  $O^T = O^{-1}$  and  $\det O = +1$ . This subgroup is called  $SE(3)$ , the *special Euclidean group* in three dimensions. Its *action* on  $\mathbb{R}^3$  is written abstractly as  $SE(3) \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ .

- Spatial translations  $g_1(\mathbf{r}_0)$  compose with proper rotations and reflections  $g_2(O)$ , as  $g_1(\mathbf{r}_0)g_2(O)$  acting on  $\mathbf{r}$ . This subgroup is called  $E(3)$ , the *Euclidean group* in three dimensions. ▲

**Remark 1.3.1** Spatial translations and rotations do not commute in general. That is,  $g_1 g_2 \neq g_2 g_1$ , unless the direction of translation and axis of rotation are collinear.  $\square$

### 1.3.1 Matrix representation of $SE(3)$

As we have seen, the special Euclidean group in three dimensions  $SE(3)$  acts on a position vector  $\mathbf{r} \in \mathbb{R}^3$  by

$$E(O, \mathbf{r}_0)\mathbf{r} = O\mathbf{r} + \mathbf{r}_0.$$

A  $4 \times 4$  *matrix representation* of this action may be found by noticing that its right-hand side arises in multiplying the matrix times the *extended* vector  $(\mathbf{r}, 1)^T$  as

$$\begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ 1 \end{pmatrix} = \begin{pmatrix} O\mathbf{r} + \mathbf{r}_0 \\ 1 \end{pmatrix}.$$

Therefore we may identify a group element of  $SE(3)$  with a  $4 \times 4$  matrix,

$$E(O, \mathbf{r}_0) = \begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix}.$$

The group  $SE(3)$  has six parameters. These are the angles of rotation about each of the three spatial axes by the orthogonal matrix  $O \in SO(3)$  with  $O^T = O^{-1}$  and the three components of the vector of translations  $\mathbf{r}_0 \in \mathbb{R}^3$ .

The *group composition law* for  $SE(3)$  is expressed as

$$\begin{aligned} E(\tilde{O}, \tilde{\mathbf{r}}_0)E(O, \mathbf{r}_0)\mathbf{r} &= E(\tilde{O}, \tilde{\mathbf{r}}_0)(O\mathbf{r} + \mathbf{r}_0) \\ &= \tilde{O}(O\mathbf{r} + \mathbf{r}_0) + \tilde{\mathbf{r}}_0, \end{aligned}$$

with  $(O, \tilde{O}) \in SO(3)$  and  $(\mathbf{r}, \tilde{\mathbf{r}}_0) \in \mathbb{R}^3$ . This formula for group composition may be represented by matrix multiplication from the

## 10 1 : GALILEO

left as

$$\begin{aligned} E(\tilde{O}, \tilde{\mathbf{r}}_0)E(O, \mathbf{r}_0) &= \begin{pmatrix} \tilde{O} & \tilde{\mathbf{r}}_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \tilde{O}O & \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{r}}_0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

which may also be expressed by simply writing the top row,

$$(\tilde{O}, \tilde{\mathbf{r}}_0)(O, \mathbf{r}_0) = (\tilde{O}O, \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{r}}_0).$$

The identity element ( $e$ ) of  $SE(3)$  is represented by

$$e = E(I, \mathbf{0}) = \begin{pmatrix} I & \mathbf{0} \\ 0 & 1 \end{pmatrix},$$

or simply  $e = (I, \mathbf{0})$ . The inverse element is represented by the matrix inverse

$$E(O, \mathbf{r}_0)^{-1} = \begin{pmatrix} O^{-1} & -O^{-1}\mathbf{r}_0 \\ 0 & 1 \end{pmatrix}.$$

In this matrix representation of  $SE(3)$ , one checks directly that

$$\begin{aligned} E(O, \mathbf{r}_0)^{-1}E(O, \mathbf{r}_0) &= \begin{pmatrix} O^{-1} & -O^{-1}\mathbf{r}_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} O & \mathbf{r}_0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} I & \mathbf{0} \\ 0 & 1 \end{pmatrix} = (I, \mathbf{0}) = e. \end{aligned}$$

In the shorter notation, the inverse may be written as

$$(O, \mathbf{r}_0)^{-1} = (O^{-1}, -O^{-1}\mathbf{r}_0)$$

and  $O^{-1} = O^T$  since the  $3 \times 3$  matrix  $O \in SO(3)$  is orthogonal.

**Remark 1.3.2** The inverse operation of  $SE(3)$  involves composition of the inverse for rotations with the inverse for translations. This entwining means that the group structure of  $SE(3)$  is not simply a direct product of its two subgroups  $\mathbb{R}^3$  and  $SO(3)$ .  $\square$

## 1.4 Lie group actions of $SE(3)$

Group multiplication in  $SE(3)$  is denoted as

$$(\tilde{O}, \tilde{\mathbf{r}}_0)(O, \mathbf{r}_0) = (\tilde{O}O, \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{r}}_0). \quad (1.4.1)$$

This notation demonstrates the following group properties of  $SE(3)$ :

- Translations in the subgroup  $\mathbb{R}^3 \subset SE(3)$  act on each other by vector addition,

$$\mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3 : (I, \tilde{\mathbf{r}}_0)(I, \mathbf{r}_0) = (I, \mathbf{r}_0 + \tilde{\mathbf{r}}_0).$$

- Rotations in the subgroup  $SO(3) \subset SE(3)$  act on each other by composition,

$$SO(3) \times SO(3) \mapsto SO(3) : (\tilde{O}, \mathbf{0})(O, \mathbf{0}) = (\tilde{O}O, \mathbf{0}).$$

- Rotations in the subgroup  $SO(3) \subset SE(3)$  act *homogeneously* on the vector space of translations in the subgroup  $\mathbb{R}^3 \subset SE(3)$ ,

$$SO(3) \times \mathbb{R}^3 \mapsto \mathbb{R}^3 : (\tilde{O}, \mathbf{0})(I, \mathbf{r}_0) = (\tilde{O}, \tilde{O}\mathbf{r}_0).$$

That is, the action of the subgroup  $SO(3) \subset SE(3)$  on the subgroup  $\mathbb{R}^3 \subset SE(3)$  maps  $\mathbb{R}^3$  into itself. The translations  $\mathbb{R}^3 \subset SE(3)$  are thus said to form a *normal*, or *invariant subgroup* of the group  $SE(3)$ .

- Every element of  $(O, \mathbf{r}_0)$  of  $SE(3)$  may be represented uniquely by composing a translation acting from the left on a rotation. That is, each element may be decomposed into

$$(O, \mathbf{r}_0) = (I, \mathbf{r}_0)(O, \mathbf{0}),$$

for a *unique*  $\mathbf{r}_0 \in \mathbb{R}^3$  and  $O \in SO(3)$ . Conversely, one may uniquely represent

$$(O, \mathbf{r}_0) = (O, \mathbf{0})(I, O^{-1}\mathbf{r}_0),$$

## 12 1 : GALILEO

by composing a rotation acting from the left on a translation. This equivalence endows the Lie group  $SE(3)$  with a semidirect-product structure,

$$SE(3) = SO(3) \ltimes \mathbb{R}^3. \quad (1.4.2)$$

**Definition 1.4.1 (Semidirect-product Lie group)** *A Lie group  $G$  that may be decomposed uniquely into a normal subgroup  $N$  and a subgroup  $H$  such that every group element may be written as*

$$g = nh \quad \text{or} \quad g = hn \quad (\text{in either order}), \quad (1.4.3)$$

*for unique choices of  $n \in N$  and  $h \in H$ , is called a **semidirect product** of  $H$  and  $N$ , denoted here by  $\ltimes$ , as in*

$$G = H \ltimes N.$$

When the normal subgroup  $N$  is a vector space, the action of a semidirect-product group on itself is given as in formula (1.4.1) for  $SE(3)$ . If the normal subgroup  $N$  is not a vector space, then the operation of addition in formula (1.4.1) is replaced by the composition law for  $N$ .

## 1.5 Lie group actions of $G(3)$

The Galilean group in three dimensions  $G(3)$  has ten parameters ( $O \in SO(3)$ ,  $\mathbf{r}_0 \in \mathbb{R}^3$ ,  $\mathbf{v}_0 \in \mathbb{R}^3$ ,  $t_0 \in \mathbb{R}$ ). The Galilean group is *also* a semidirect-product Lie group, which may be written as

$$G(3) = SE(3) \ltimes \mathbb{R}^4 = \left( SO(3) \ltimes \mathbb{R}^3 \right) \ltimes \mathbb{R}^4. \quad (1.5.1)$$

That is, the subgroup of Euclidean motions consisting of rotations and Galilean velocity boosts  $(O, \mathbf{v}_0) \in SE(3)$  acts homogeneously on the subgroups of space and time translations  $(\mathbf{r}_0, t_0) \in \mathbb{R}^4$  which commute with each other.



**Exercise.** Compute explicitly the *inverse* of the Galilean group element  $g = g_1 g_3 g_2$  obtained by representing the action of the Galilean group as matrix multiplication  $G(3) \times \mathbb{R}^4 \rightarrow \mathbb{R}^4$  on the extended vector  $(\mathbf{r}, t, 1)^T \in \mathbb{R}^4$ ,

$$\begin{aligned} g_1 g_3 g_2 \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} &= \begin{pmatrix} O & \mathbf{v}_0 & \mathbf{r}_0 \\ \mathbf{0} & 1 & t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} \quad (1.5.2) \\ &= \begin{pmatrix} O\mathbf{r} + t\mathbf{v}_0 + \mathbf{r}_0 \\ t + t_0 \\ 1 \end{pmatrix}. \end{aligned}$$

★

**Answer.** Write the product  $g = g_1 g_3 g_2$  as

$$g = g_1 g_3 g_2 = \begin{pmatrix} I & \mathbf{0} & \mathbf{r}_0 \\ \mathbf{0} & 1 & t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} I & \mathbf{v}_0 & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} O & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{pmatrix}.$$

Then, the product  $g^{-1} = (g_1 g_3 g_2)^{-1} = g_2^{-1} g_3^{-1} g_1^{-1}$  appears in matrix form as

$$\begin{aligned} g^{-1} &= \begin{pmatrix} O^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} I & -\mathbf{v}_0 & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} I & \mathbf{0} & -\mathbf{r}_0 \\ \mathbf{0} & 1 & -t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} O^{-1} & -O^{-1}\mathbf{v}_0 & -O^{-1}(\mathbf{r}_0 - t\mathbf{v}_0) \\ \mathbf{0} & 1 & -t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix}. \end{aligned}$$

▲

## 14 1 : GALILEO

**Exercise.** Write the corresponding matrices for the Galilean transformations for  $g_1g_2g_3$ ,  $g_2g_1g_3$  and  $g_3g_2g_1$ . ★

1.5.1 Matrix representation of  $G(3)$ 

The formula for group composition  $G(3) \times G(3) \rightarrow G(3)$  may be represented by matrix multiplication from the left as

$$\begin{aligned} \begin{pmatrix} \tilde{O} & \tilde{\mathbf{v}}_0 & \tilde{\mathbf{r}}_0 \\ \mathbf{0} & 1 & \tilde{t}_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} O & \mathbf{v}_0 & \mathbf{r}_0 \\ \mathbf{0} & 1 & t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} & \quad (1.5.3) \\ = \begin{pmatrix} \tilde{O}O & \tilde{O}\mathbf{v}_0 + \tilde{\mathbf{v}}_0 & \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{v}}_0t_0 + \tilde{\mathbf{r}}_0 \\ \mathbf{0} & 1 & \tilde{t}_0 + t_0 \\ \mathbf{0} & 0 & 1 \end{pmatrix}, \end{aligned}$$

which may be expressed more succinctly as

$$\begin{aligned} (\tilde{O}, \tilde{\mathbf{v}}_0, \tilde{\mathbf{r}}_0, \tilde{t}_0)(O, \mathbf{v}_0, \mathbf{r}_0, t_0) & \quad (1.5.4) \\ = (\tilde{O}O, \tilde{O}\mathbf{v}_0 + \tilde{\mathbf{v}}_0, \tilde{O}\mathbf{r}_0 + \tilde{\mathbf{v}}_0t_0 + \tilde{\mathbf{r}}_0, \tilde{t}_0 + t_0). \end{aligned}$$

**Exercise.** Check the semidirect-product structure (1.5.1) for the Lie group  $G(3) = SE(3) \ltimes \mathbb{R}^4$ , by writing explicit matrix expressions for  $g = nh$  and  $g = hn$  with  $h = SE(3)$  and  $n = \mathbb{R}^4$ . ★

**Answer.** In verifying the semidirect-product structure condition (1.4.3) that  $g = nh$  or  $g = hn$  in either order, we write explicitly

$$\begin{aligned} (O, \mathbf{v}_0, \mathbf{r}_0, t_0) &= (I, 0, \mathbf{r}_0, t_0)(O, \mathbf{v}_0, 0, 0) & (1.5.5) \\ &= (O, \mathbf{v}_0, 0, 0)(I, 0, O^{-1}(\mathbf{r}_0 - \mathbf{v}_0t_0), t_0). \end{aligned}$$

▲

## 1.6 Lie algebra of $SE(3)$

A  $4 \times 4$  matrix representation of tangent vectors for  $SE(3)$  at the identity may be found by first computing the derivative of a general group element  $(O(s), \mathbf{r}_0(s))$  along the group path with parameter  $s$  and bringing the result back to the identity at  $s = 0$ ,

$$\begin{aligned} & \left[ \begin{pmatrix} O(s) & \mathbf{r}_0(s) \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} O'(s) & \mathbf{r}'_0(s) \\ 0 & 0 \end{pmatrix} \right]_{s=0} \\ &= \begin{pmatrix} O^{-1}(0)O'(0) & O^{-1}(0)\mathbf{r}'_0(0) \\ 0 & 0 \end{pmatrix} =: \begin{pmatrix} \hat{\Xi} & \mathbf{r}_0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

where in the last step we have dropped the unnecessary superscript prime ( $'$ ). The quantity  $\hat{\Xi} = O^{-1}(s)O'(s)|_{s=0}$  is a  $3 \times 3$  skew-symmetric matrix, since  $O$  is a  $3 \times 3$  orthogonal matrix. Thus,  $\hat{\Xi}$  may be written using the *hat map*, defined by

$$\hat{\Xi} = \begin{pmatrix} 0 & -\Xi_3 & \Xi_2 \\ \Xi_3 & 0 & -\Xi_1 \\ -\Xi_2 & \Xi_1 & 0 \end{pmatrix}, \quad (1.6.1)$$

in terms of a vector  $\Xi \in \mathbb{R}^3$  with components  $\Xi_i$ , with  $i = 1, 2, 3$ . Infinitesimal rotations are expressed by the vector cross product,

$$\hat{\Xi}\mathbf{r} = \Xi \times \mathbf{r}. \quad (1.6.2)$$

The matrix components of  $\hat{\Xi}$  may also be written in terms of the components of the vector  $\Xi$  as

$$\hat{\Xi}_{jk} = \left( O^{-1} \frac{dO}{ds} \right)_{jk} \Big|_{s=0} = -\Xi_i \epsilon_{ijk},$$

where  $\epsilon_{ijk}$  with  $i, j, k = 1, 2, 3$  is the totally antisymmetric tensor with  $\epsilon_{123} = 1$ ,  $\epsilon_{213} = -1$ , etc. One may compute directly, for a fixed vector  $\mathbf{r}$ ,

$$\frac{d}{ds} e^{s\hat{\Xi}} \mathbf{r} = \hat{\Xi} e^{s\hat{\Xi}} \mathbf{r} = \Xi \times e^{s\hat{\Xi}} \mathbf{r}.$$

## 16 1 : GALILEO

Consequently, one may evaluate, at  $s = 0$ ,

$$\left. \frac{d}{ds} e^{s\hat{\Xi}} \mathbf{r} \right|_{s=0} = \hat{\Xi} \mathbf{r} = \Xi \times \mathbf{r}.$$

This expression recovers the expected result in (1.6.2) in terms of the exponential notation. It means the quantity  $\mathbf{r}(s) = \exp(s\hat{\Xi})\mathbf{r}$  describes a finite, right-handed rotation of the initial vector  $\mathbf{r} = \mathbf{r}(0)$  by the angle  $s|\Xi|$  around the axis pointing in the direction of  $\Xi$ .

**Remark 1.6.1 (Properties of the hat map)** The hat map arises in the infinitesimal rotations

$$\hat{\Xi}_{jk} = (O^{-1}dO/ds)_{jk}|_{s=0} = -\Xi_i \epsilon_{ijk}.$$

The hat map is an isomorphism:

$$(\mathbb{R}^3, \times) \mapsto (\mathfrak{so}(3), [\cdot, \cdot]).$$

That is, the hat map identifies the composition of two vectors in  $\mathbb{R}^3$  using the cross product with the commutator of two skew-symmetric  $3 \times 3$  matrices. Specifically, we write for any two vectors  $\mathbf{Q}, \Xi \in \mathbb{R}^3$ ,

$$-(\mathbf{Q} \times \Xi)_k = \epsilon_{klm} \Xi^l Q^m = \hat{\Xi}_{km} Q^m.$$

That is,

$$\Xi \times \mathbf{Q} = \hat{\Xi} \mathbf{Q} \quad \text{for all } \Xi, \mathbf{Q} \in \mathbb{R}^3.$$

The following formulas may be easily verified for  $\mathbf{P}, \mathbf{Q}, \Xi \in \mathbb{R}^3$ :

$$\begin{aligned} (\mathbf{P} \times \mathbf{Q})^\wedge &= [\hat{\mathbf{P}}, \hat{\mathbf{Q}}], \\ [\hat{\mathbf{P}}, \hat{\mathbf{Q}}] \Xi &= (\mathbf{P} \times \mathbf{Q}) \times \Xi, \\ \mathbf{P} \cdot \mathbf{Q} &= -\frac{1}{2} \text{trace}(\hat{\mathbf{P}} \hat{\mathbf{Q}}). \end{aligned}$$

□

**Remark 1.6.2** The commutator of infinitesimal transformation matrices given by the formula

$$\left[ \begin{pmatrix} \hat{\Xi}_1 & \mathbf{r}_1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \hat{\Xi}_2 & \mathbf{r}_2 \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} \hat{\Xi}_1 \hat{\Xi}_2 - \hat{\Xi}_2 \hat{\Xi}_1 & \hat{\Xi}_1 \mathbf{r}_2 - \hat{\Xi}_2 \mathbf{r}_1 \\ 0 & 0 \end{pmatrix}$$

provides a matrix representation of  $se(3)$ , the Lie algebra of the Lie group  $SE(3)$ . In vector notation, this becomes

$$\left[ \begin{pmatrix} \mathbf{\Xi}_1 \times & \mathbf{r}_1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{\Xi}_2 \times & \mathbf{r}_2 \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} (\mathbf{\Xi}_1 \times \mathbf{\Xi}_2) \times & \mathbf{\Xi}_1 \times \mathbf{r}_2 - \mathbf{\Xi}_2 \times \mathbf{r}_1 \\ 0 & 0 \end{pmatrix}.$$

□

**Remark 1.6.3** The  $se(3)$  matrix commutator yields

$$\begin{aligned} \left[ (\widehat{\Xi}_1, \mathbf{r}_1), (\widehat{\Xi}_2, \mathbf{r}_2) \right] &= \left( \widehat{\Xi}_1 \widehat{\Xi}_2 - \widehat{\Xi}_2 \widehat{\Xi}_1, \widehat{\Xi}_1 \mathbf{r}_2 - \widehat{\Xi}_2 \mathbf{r}_1 \right) \\ &= \left( [\widehat{\Xi}_1, \widehat{\Xi}_2], \widehat{\Xi}_1 \mathbf{r}_2 - \widehat{\Xi}_2 \mathbf{r}_1 \right), \end{aligned}$$

which is the classic expression for the Lie algebra of a semidirect-product Lie group. □

## 1.7 Lie algebra of $G(3)$

A  $5 \times 5$  matrix representation of tangent vectors for  $G(3)$  at the identity may be found by computing the derivative of a general group element  $(O(s), \mathbf{v}_0(s), \mathbf{r}_0(s), t_0(s))$  along the group path with parameter  $s$  and bringing the result back to the identity at  $s = 0$ ,

$$\begin{aligned} & \left[ \begin{pmatrix} O(s) & \mathbf{v}_0(s) & \mathbf{r}_0(s) \\ 0 & 1 & t_0(s) \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} O'(s) & \mathbf{v}'_0(s) & \mathbf{r}'_0(s) \\ 0 & 0 & t'_0(s) \\ 0 & 0 & 0 \end{pmatrix} \right]_{s=0} \\ &= \left( \begin{matrix} O^{-1}(s)O'(s) & O^{-1}(s)\mathbf{v}'_0(s) & O^{-1}(s)(\mathbf{r}'_0(s) - \mathbf{v}'_0(s)t'_0(s)) \\ 0 & 0 & t'_0(s) \\ 0 & 0 & 0 \end{matrix} \right) \Big|_{s=0} \\ &= \begin{pmatrix} \widehat{\Xi} & \mathbf{v}_0 & \mathbf{r}_0 - \mathbf{v}_0 t_0 \\ 0 & 0 & t_0 \\ 0 & 0 & 0 \end{pmatrix} =: (\widehat{\Xi}, \mathbf{v}_0, \mathbf{r}_0, t_0), \end{aligned}$$

in terms of the  $3 \times 3$  skew-symmetric matrix  $\widehat{\Xi} = O^{-1}(s)O'(s)|_{s=0}$ . For notational convenience, the superscript primes that would have

## 18 1 : GALILEO

appeared on the tangents of the Galilean shift parameters  $\mathbf{r}'_0(0)$ ,  $\mathbf{v}'_0(0)$  and  $t'_0(0)$  at the identity  $s = 0$  have been dropped in the last line and replaced by the simpler forms  $\mathbf{r}_0$ ,  $\mathbf{v}_0$ ,  $t_0$ , respectively.

**Exercise. (Galilean Lie algebra commutator)** Verify the commutation relation

$$\begin{aligned} & \left[ (\widehat{\Xi}_1, \mathbf{v}_1, \mathbf{r}_1, t_1), (\widehat{\Xi}_2, \mathbf{v}_2, \mathbf{r}_2, t_2) \right] \\ &= \left( \left[ \widehat{\Xi}_1, \widehat{\Xi}_2 \right], \widehat{\Xi}_1 \mathbf{v}_2 - \widehat{\Xi}_2 \mathbf{v}_1, \widehat{\Xi}_1(\mathbf{r}_2, \mathbf{v}_2, t_2) - \widehat{\Xi}_2(\mathbf{r}_1, \mathbf{v}_1, t_1), 0 \right), \end{aligned}$$

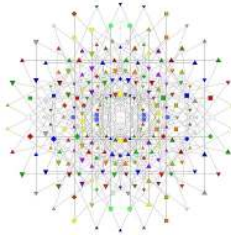
where

$$\begin{aligned} & \widehat{\Xi}_1(\mathbf{r}_2, \mathbf{v}_2, t_2) - \widehat{\Xi}_2(\mathbf{r}_1, \mathbf{v}_1, t_1) \\ &= \left( \widehat{\Xi}_1(\mathbf{r}_2 - \mathbf{v}_2 t_2) + \mathbf{v}_1 t_2 \right) - \left( \widehat{\Xi}_2(\mathbf{r}_1 - \mathbf{v}_1 t_1) + \mathbf{v}_2 t_1 \right). \end{aligned}$$

★

According to the principle of Galilean relativity, the laws of mechanics must take the same form in any uniformly moving reference frame. That is, the expressions of these laws must be invariant in form under Galilean transformations. In this chapter, we have introduced the Galilean transformations, shown that they comprise a Lie group, found its subgroups, endowed them with a matrix representation, and identified their group structure mathematically as a nested semidirect product.

Rigid motion in  $\mathbb{R}^3$  corresponds to a smoothly varying sequence of changes of reference frame along a time-dependent path in the special Euclidean Lie group,  $SE(3)$ . This is the main subject of the text.



# 2

## NEWTON, LAGRANGE, HAMILTON AND THE RIGID BODY

---

### Contents

2.1	Newton	21
2.1.1	Newtonian form of free rigid rotation	21
2.1.2	Newtonian form of rigid-body motion	30
2.2	Lagrange	36
2.2.1	The principle of stationary action	36
2.3	Noether's theorem	39
2.3.1	Lie symmetries and conservation laws	39
2.3.2	Infinitesimal transformations of a Lie group	40
2.4	Lagrangian form of rigid-body motion	46
2.4.1	Hamilton–Pontryagin constrained variations	50
2.4.2	Manakov's formulation of the $SO(n)$ rigid body	54
2.4.3	Matrix Euler–Poincaré equations	55
2.4.4	An isospectral eigenvalue problem for the $SO(n)$ rigid body	56
2.4.5	Manakov's integration of the $SO(n)$ rigid body	58

**20** 2 : NEWTON, LAGRANGE, HAMILTON

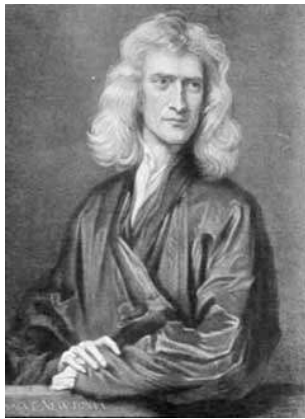
2.5	Hamilton	<b>60</b>
2.5.1	Hamiltonian form of rigid-body motion	<b>62</b>
2.5.2	Lie–Poisson Hamiltonian rigid-body dynamics	<b>63</b>
2.5.3	Lie–Poisson bracket	<b>64</b>
2.5.4	Nambu’s $\mathbb{R}^3$ Poisson bracket	<b>65</b>
2.5.5	Clebsch variational principle for the rigid body	<b>69</b>
2.5.6	Rotating motion with potential energy	<b>72</b>

---



## 2.1 Newton

### 2.1.1 Newtonian form of free rigid rotation



Isaac Newton

**Definition 2.1.1** *In free rigid rotation a body rotates about its centre of mass and the pairwise distances between all points in the body remain fixed.*

**Definition 2.1.2** *A system of coordinates fixed in a body undergoing free rigid rotation is stationary in the rotating orthonormal basis called the **body frame**, introduced by Euler [Eu1758].*

The orientation of the orthonormal frame  $(\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3)$  fixed in the rotating body relative to a basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  fixed in space depends smoothly on time

$t \in \mathbb{R}$ . In the fixed spatial coordinate system, the body frame is seen as the moving frame

$$(O(t)\mathbf{E}_1, O(t)\mathbf{E}_2, O(t)\mathbf{E}_3),$$

where  $O(t) \in SO(3)$  defines the attitude of the body relative to its reference configuration according to the following matrix multiplication on its three unit vectors:

$$\mathbf{e}_a(t) = O(t)\mathbf{E}_a, \quad a = 1, 2, 3. \quad (2.1.1)$$

Here the unit vectors  $\mathbf{e}_a(0) = \mathbf{E}_a$  with  $a = 1, 2, 3$  comprise at initial time  $t = 0$  an orthonormal basis of coordinates and  $O(t)$  is a special ( $\det O(t) = 1$ ) orthogonal ( $O^T(t)O(t) = Id$ )  $3 \times 3$  matrix. That is,  $O(t)$  is a continuous function defined along a curve parameterised by time  $t$  in the special orthogonal matrix group  $SO(3)$ . At the initial time  $t = 0$ , we may take  $O(0) = Id$ , without any loss.

## 22 2 : NEWTON, LAGRANGE, HAMILTON

As the orientation of the body is evolving according to (2.1.1), each basis vector in the set,

$$\mathbf{e}(t) \in \{\mathbf{e}_1(t), \mathbf{e}_2(t), \mathbf{e}_3(t)\},$$

preserves its (unit) length. We prove this and fix notation by writing

$$\begin{aligned} 1 &= |\mathbf{e}(t)|^2 := \mathbf{e}(t) \cdot \mathbf{e}(t) := \mathbf{e}(t)^T \mathbf{e}(t) = (O(t)\mathbf{E})^T O(t)\mathbf{E} \\ &= \mathbf{E}^T O^T(t)O(t)\mathbf{E} = \mathbf{E}^T (Id)\mathbf{E} = |\mathbf{E}|^2, \end{aligned} \quad (2.1.2)$$

which follows because  $O(t)$  is orthogonal; that is,  $O^T(t)O(t) = Id$ .

The basis vectors in the orthonormal frame  $\mathbf{e}_a(0) = \mathbf{E}_a$  define the initial orientation of the set of rotating points with respect to some choice of fixed spatial coordinates at time  $t = 0$ . Each point  $\mathbf{r}(t)$  in the subsequent rigid motion may be represented in either fixed or rotating coordinates as

$$\mathbf{r}(t) = r_0^A(t)\mathbf{e}_A(0) \quad \text{in the fixed basis,} \quad (2.1.3)$$

$$= r^a \mathbf{e}_a(t) \quad \text{in the rotating basis.} \quad (2.1.4)$$

The fixed basis is called the *spatial frame* and the rotating basis is the *body frame*.

The constant components  $r^a$  of a position vector relative to the rotating basis are related to its *initial* spatial position as

$$r^a = O_A^a(0)r_0^A(0).$$

This is simply  $r^a = \delta_A^a r_0^A(0)$  for the choice  $O(0) = Id$  in which the two coordinate bases are initially aligned. The components of any vector  $\mathbf{J}$  in the fixed (spatial) frame are related to those in the moving (body) frame by the mutual rotation of their axes in (2.1.1) at any time. That is,

$$\mathbf{J} = \mathbf{e}_a(0)J_{space}^a(t) = \mathbf{e}_a(t)J_{body}^a = O(t)\mathbf{e}_a(0)J_{body}^a, \quad (2.1.5)$$

or equivalently, as in Equation (2.1.1),

$$\mathbf{J}_{space}(t) = O(t)\mathbf{J}_{body}. \quad (2.1.6)$$

**Lemma 2.1.1** *The velocity  $\dot{\mathbf{r}}(t)$  of a point  $\mathbf{r}(t)$  in free rigid rotation depends linearly on its position relative to the centre of mass.*

**Proof.** In particular,  $\mathbf{r}(t) = r^a O(t) \mathbf{e}_a(0)$  implies

$$\dot{\mathbf{r}}(t) = r^a \dot{\mathbf{e}}_a(t) = r^a \dot{O}(t) \mathbf{e}_a(0) =: r^a \dot{O} O^{-1}(t) \mathbf{e}_a(t) =: \hat{\omega}(t) \mathbf{r}, \quad (2.1.7)$$

which is linear. ■

**Remark 2.1.1 (Wide-hat notation)** The wide-hat notation ( $\hat{\cdot}$ ) denotes a skew-symmetric  $3 \times 3$  matrix. There is no danger of confusing wide-hat notation ( $\hat{\cdot}$ ) with narrow-hat notation ( $\hat{\cdot}$ ), which denotes a unit vector (or, later, a unit quaternion). □

**Lemma 2.1.2 (Skew-symmetry)** *The spatial angular velocity matrix  $\hat{\omega}(t) = \dot{O} O^{-1}(t)$  in (2.1.7) is skew-symmetric, i.e.,*

$$\hat{\omega}^T = -\hat{\omega}.$$

**Proof.** Being orthogonal, the matrix  $O(t)$  satisfies  $OO^T = Id$ . This implies that  $\hat{\omega}$  is skew-symmetric,

$$\begin{aligned} 0 &= (OO^T) \cdot &= \dot{O} O^T + O \dot{O}^T &= \dot{O} O^T + (\dot{O} O^T)^T \\ &= \dot{O} O^{-1} + (\dot{O} O^{-1})^T &= \hat{\omega} + \hat{\omega}^T. \end{aligned}$$

■

**Remark 2.1.2** The skew  $3 \times 3$  real matrices form a closed linear space under addition. □

**Definition 2.1.3 (Commutator product of skew matrices)** *The commutator product of two skew matrices  $\hat{\omega}$  and  $\hat{\xi}$  is defined as the skew matrix product*

$$[\hat{\omega}, \hat{\xi}] := \hat{\omega} \hat{\xi} - \hat{\xi} \hat{\omega}. \quad (2.1.8)$$

**Remark 2.1.3** This commutator product is again a skew  $3 \times 3$  real matrix. □

## 24 2 : NEWTON, LAGRANGE, HAMILTON

**Definition 2.1.4 (Basis set for skew matrices)** Any  $3 \times 3$  antisymmetric matrix  $\widehat{\omega}^T = -\widehat{\omega}$  may be written as a linear combination of the following three linearly independent basis elements for the  $3 \times 3$  skew matrices:

$$\widehat{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \widehat{J}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \widehat{J}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

That is, the element  $\widehat{J}_a$  for this choice of basis has matrix components

$$(\widehat{J}_a)_{bc} = -\epsilon_{abc},$$

where  $\epsilon_{abc}$  is the totally antisymmetric tensor with

$$\epsilon_{123} = +1, \quad \epsilon_{213} = -1, \quad \epsilon_{113} = 0, \quad \text{etc.}$$

**Lemma 2.1.3 (Commutation relations)** The skew matrix basis  $\widehat{J}_a$  with  $a = 1, 2, 3$  satisfies the *commutation relations*,

$$[\widehat{J}_a, \widehat{J}_b] := \widehat{J}_a \widehat{J}_b - \widehat{J}_b \widehat{J}_a = \epsilon_{abc} \widehat{J}_c. \quad (2.1.9)$$

**Proof.** This may be verified by a direct calculation,  $[\widehat{J}_1, \widehat{J}_2] = \widehat{J}_3$ , etc. ■

**Remark 2.1.4** The closure of the basis set of skew-symmetric matrices under the commutator product gives the linear space of skew-symmetric matrices a *Lie algebra structure*. The constants  $\epsilon_{abc}$  in the commutation relations among the skew  $3 \times 3$  matrix basis elements are called the *structure constants* and the corresponding Lie algebra is called  $so(3)$ . This also means the abstract  $so(3)$  Lie algebra may be represented by skew  $3 \times 3$  matrices, which is a great convenience, as we shall see for example in Section 4.2. The Lie algebra  $so(3)$  may also be defined as the tangent space to the Lie group  $SO(3)$  at the identity, as discussed in Appendix B. □

**Theorem 2.1.1 (Hat map)** The components of any  $3 \times 3$  skew matrix  $\widehat{\omega}$  may be identified with the corresponding components of a vector  $\omega \in \mathbb{R}^3$ .

**Proof.** In the basis (2.1.4), one writes the linear invertible relation,

$$\hat{\omega} = \begin{pmatrix} 0 & -\omega^3 & \omega^2 \\ \omega^3 & 0 & -\omega^1 \\ -\omega^2 & \omega^1 & 0 \end{pmatrix} = \omega^a \hat{J}_a =: \boldsymbol{\omega} \cdot \hat{\boldsymbol{J}}, \quad (2.1.10)$$

for  $a = 1, 2, 3$ . This is a one-to-one invertible map, i.e., it is an *isomorphism*, between  $3 \times 3$  skew-symmetric matrices and vectors in  $\mathbb{R}^3$ . ■

**Remark 2.1.5** The superscript hat ( $\hat{\phantom{x}}$ ) applied to a vector identifies that vector in  $\mathbb{R}^3$  with a  $3 \times 3$  skew-symmetric matrix. For example, the unit vectors in the Cartesian basis set,  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , are associated with the basis elements  $\hat{J}_a$ , for  $a = 1, 2, 3$ , in Equation (2.1.4) by  $\hat{J}_a = \hat{\mathbf{e}}_a$ , or in matrix components,

$$(\hat{\mathbf{e}}_a)_{bc} = -\delta_a^d \epsilon_{dbc} = -\epsilon_{abc} = (\mathbf{e}_a \times)_{bc}.$$

□

**Remark 2.1.6** The last equality in the definition of the hat map in Equation (2.1.10) introduces the convenient notation  $\hat{\boldsymbol{J}}$  that denotes the basis for the  $3 \times 3$  skew-symmetric matrices  $\hat{J}_a$ , with  $a = 1, 2, 3$  as a vector of matrices. □

**Definition 2.1.5 (Hat map for angular velocity vector)** *The relation  $\hat{\omega} = \boldsymbol{\omega} \cdot \hat{\boldsymbol{J}}$  in Equation (2.1.10) identifies the skew-symmetric  $3 \times 3$  matrix  $\hat{\omega}(t)$  with the **angular velocity vector**  $\boldsymbol{\omega}(t) \in \mathbb{R}^3$  whose components  $\omega^c(t)$ , with  $c = 1, 2, 3$ , are given by*

$$(\dot{O}O^{-1})_{ab}(t) = \hat{\omega}_{ab}(t) = -\epsilon_{abc} \omega^c(t). \quad (2.1.11)$$

Equation (2.1.11) defines the matrix components of the **hat map for angular velocity**.

**Remark 2.1.7** Equivalently, the hat map in Equation (2.1.10) is defined by the identity

$$\hat{\omega} \boldsymbol{\lambda} = \boldsymbol{\omega} \times \boldsymbol{\lambda} \quad \text{for all } \boldsymbol{\omega}, \boldsymbol{\lambda} \in \mathbb{R}^3.$$

Thus, we may write  $\hat{\omega} = \boldsymbol{\omega} \times$  to identify the vector  $\boldsymbol{\omega} \in \mathbb{R}^3$  with the skew  $3 \times 3$  matrix  $\hat{\omega} \in so(3)$ . □

## 26 2 : NEWTON, LAGRANGE, HAMILTON

**Proposition 2.1.1** *The  $3 \times 3$  skew matrices*

$$\widehat{\omega} = \boldsymbol{\omega} \cdot \widehat{\boldsymbol{J}} \quad \text{and} \quad \widehat{\lambda} = \boldsymbol{\lambda} \cdot \widehat{\boldsymbol{J}}$$

*associated with the vectors  $\boldsymbol{\omega}$  and  $\boldsymbol{\lambda}$  in  $\mathbb{R}^3$  satisfy the commutation relation*

$$[\widehat{\omega}, \widehat{\lambda}] = \boldsymbol{\omega} \times \boldsymbol{\lambda} \cdot \widehat{\boldsymbol{J}} =: (\boldsymbol{\omega} \times \boldsymbol{\lambda})^\wedge, \quad (2.1.12)$$

*where  $\boldsymbol{\omega} \times \boldsymbol{\lambda}$  is the vector product in  $\mathbb{R}^3$ .*

**Proof.** Formula (2.1.9) implies the result, by

$$\begin{aligned} [\widehat{\omega}, \widehat{\lambda}] &= [\boldsymbol{\omega} \cdot \widehat{\boldsymbol{J}}, \boldsymbol{\lambda} \cdot \widehat{\boldsymbol{J}}] = [\omega^a \widehat{J}_a, \lambda^b \widehat{J}_b] \\ &= \omega^a \lambda^b [\widehat{J}_a, \widehat{J}_b] = \omega^a \lambda^b \epsilon_{abc} \widehat{J}_c = \boldsymbol{\omega} \times \boldsymbol{\lambda} \cdot \widehat{\boldsymbol{J}}. \end{aligned}$$

■

**Remark 2.1.8** According to Proposition 2.1.1, the hat map  $\widehat{\cdot} : (\mathbb{R}^3, \times) \mapsto (so(3), [\cdot, \cdot])$  allows the velocity in space (2.1.7) of a point at  $\mathbf{r}$  undergoing rigid-body motion to be expressed equivalently either by a skew-matrix multiplication, or as a vector product. That is,

$$\dot{\mathbf{r}}(t) =: \widehat{\omega}(t)\mathbf{r} =: \boldsymbol{\omega}(t) \times \mathbf{r}. \quad (2.1.13)$$

Hence, free rigid motion of a point displaced by  $\mathbf{r}$  from the centre of mass is a rotation in space of  $\mathbf{r}$  about the time-dependent angular velocity vector  $\boldsymbol{\omega}(t)$ . Accordingly,  $d|\mathbf{r}|^2/dt = 2\mathbf{r} \cdot \dot{\mathbf{r}} = 0$ , and the displacement distance is preserved,  $|\mathbf{r}|(t) = |\mathbf{r}|(0)$ . □

### ***Kinetic energy of free rigid rotation***

The kinetic energy for  $N$  particles of masses  $m_j$   $j = 1, 2, \dots, N$ , mutually undergoing free rigid rotation is computed in terms of the

angular velocity as

$$\begin{aligned}
 K &= \frac{1}{2} \sum_{j=1}^N m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j \\
 &= \frac{1}{2} \sum_{j=1}^N m_j (\boldsymbol{\omega} \times \mathbf{r}_j) \cdot (\boldsymbol{\omega} \times \mathbf{r}_j) \\
 &=: \frac{1}{2} \langle\langle \boldsymbol{\omega}, \boldsymbol{\omega} \rangle\rangle.
 \end{aligned}$$

**Definition 2.1.6 (Symmetric mass-weighted pairing)** *The kinetic energy induces a symmetric mass-weighted pairing*

$$\langle\langle \cdot, \cdot \rangle\rangle : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R},$$

defined for any two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$  as

$$\langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle := \sum_j m_j (\mathbf{a} \times \mathbf{r}_j) \cdot (\mathbf{b} \times \mathbf{r}_j) =: \mathbb{I} \mathbf{a} \cdot \mathbf{b}. \quad (2.1.14)$$

**Definition 2.1.7 (Moment of inertia tensor)** *The mass-weighted pairing, or inner product in (2.1.14)*

$$\langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle = \mathbb{I} \mathbf{a} \cdot \mathbf{b},$$

defines the symmetric **moment of inertia tensor**  $\mathbb{I}$  for the particle system.

**Definition 2.1.8 (Angular momentum of rigid motion)** *The angular momentum is defined as the derivative of the kinetic energy with respect to angular velocity. In the present case with (2.1.14), this produces the linear relation*

$$\begin{aligned}
 \mathbf{J} = \frac{\partial K}{\partial \boldsymbol{\omega}} &= -\frac{1}{2} \sum_{j=1}^N m_j \mathbf{r}_j \times (\mathbf{r}_j \times \boldsymbol{\omega}) \\
 &= \frac{1}{2} \sum_{j=1}^N m_j \left( |\mathbf{r}_j|^2 Id - \mathbf{r}_j \otimes \mathbf{r}_j \right) \boldsymbol{\omega} \\
 &=: \mathbb{I} \boldsymbol{\omega}, \quad (2.1.15)
 \end{aligned}$$

where  $\mathbb{I}$  is the moment of inertia tensor defined by the symmetric pairing in (2.1.14).

**Conservation of angular momentum in free rigid rotation**

In free rigid rotation no external torques are applied, so the angular momentum  $\mathbf{J}$  is conserved. In the fixed basis  $\mathbf{J} = J_0^A(t)\mathbf{e}_A(0)$  and this conservation law is expressed as

$$0 = \frac{d\mathbf{J}}{dt} = \frac{dJ_0^A}{dt}\mathbf{e}_A(0), \quad (2.1.16)$$

so each component  $J_0^A$ , for  $A = 1, 2, 3$  of angular momentum in the spatial frame is separately conserved. In the rotating basis  $\mathbf{J} = J^a(t)\mathbf{e}_a(t)$  and angular momentum conservation becomes

$$\begin{aligned} 0 = \frac{d\mathbf{J}}{dt} &= \frac{dJ^a}{dt}\mathbf{e}_a(t) + J^a \frac{d\mathbf{e}_a(t)}{dt} \\ &= \frac{dJ^a}{dt}\mathbf{e}_a(t) + \boldsymbol{\omega} \times J^a\mathbf{e}_a(t) \\ &= \left( \frac{dJ^a}{dt} + (\boldsymbol{\omega} \times \mathbf{J})^a \right) \mathbf{e}_a(t). \end{aligned} \quad (2.1.17)$$

Consequently, the components  $J^a$ , for  $a = 1, 2, 3$  of angular momentum in the body frame satisfy the quadratically nonlinear *system* of Equations (2.1.17) with  $\mathbf{J} = \mathbb{I}\boldsymbol{\omega}$ .

**Lemma 2.1.4 (Space vs body dynamics)** *Upon denoting  $\mathbf{J}_{space} = (J_0^1, J_0^2, J_0^3)$  in the fixed basis  $\mathbf{e}_A(0)$  and  $\mathbf{J}_{body} = (J^1, J^2, J^3)$  in the time-dependent basis  $\mathbf{e}_a(t)$ , with*

$$\mathbf{J}_{space} = O(t)\mathbf{J}_{body}, \quad (2.1.18)$$

*one may summarise the two equivalent sets of Equations (2.1.16) and (2.1.17) as*

$$\frac{d\mathbf{J}_{space}}{dt} = 0 \quad \text{and} \quad \frac{d\mathbf{J}_{body}}{dt} + (\mathbb{I}^{-1}\mathbf{J}_{body}) \times \mathbf{J}_{body} = 0. \quad (2.1.19)$$



**Proof.** The time derivative of relation (2.1.18) gives

$$\begin{aligned}
 \frac{d\mathbf{J}_{body}}{dt} &= \frac{d}{dt} \left( O^{-1}(t) \mathbf{J}_{space}(t) \right) \\
 &= -O^{-1} \dot{O} \left( O^{-1}(t) \mathbf{J}_{space}(t) \right) + O^{-1} \underbrace{\frac{d\mathbf{J}_{space}}{dt}}_{\text{vanishes}} \\
 &= -\widehat{\omega}_{body} \mathbf{J}_{body} = -\boldsymbol{\omega}_{body} \times \mathbf{J}_{body}, \quad (2.1.20)
 \end{aligned}$$

with  $\widehat{\omega}_{body} := O^{-1} \dot{O} = \boldsymbol{\omega}_{body} \times$  and  $\boldsymbol{\omega}_{body} = \mathbb{I}^{-1} \mathbf{J}_{body}$ , which defines the *body angular velocity*. This is the usual heuristic derivation of the dynamics for body angular momentum. ■

**Remark 2.1.9 (Darwin, Coriolis and centrifugal forces)** Many elementary mechanics texts make the following points about the various *noninertial forces* that arise in a rotating frame. For any vector  $\mathbf{r}(t) = r^a(t) \mathbf{e}_a(t)$  the body and space time derivatives satisfy the first time-derivative relation, as in (2.1.17),

$$\begin{aligned}
 \dot{\mathbf{r}}(t) &= \dot{r}^a \mathbf{e}_a(t) + r^a \dot{\mathbf{e}}_a(t) \\
 &= \dot{r}^a \mathbf{e}_a(t) + \boldsymbol{\omega} \times r^a \mathbf{e}_a(t) \\
 &= (\dot{r}^a + \epsilon_{bc}^a \omega^b r^c) \mathbf{e}_a(t) \\
 &=: (\dot{r}^a + (\boldsymbol{\omega} \times \mathbf{r})^a) \mathbf{e}_a(t). \quad (2.1.21)
 \end{aligned}$$

Taking a second time derivative in this notation yields

$$\begin{aligned}
 \ddot{\mathbf{r}}(t) &= (\ddot{r}^a + (\dot{\boldsymbol{\omega}} \times \mathbf{r})^a + (\boldsymbol{\omega} \times \dot{\mathbf{r}})^a) \mathbf{e}_a(t) + (\dot{r}^a + (\boldsymbol{\omega} \times \mathbf{r})^a) \dot{\mathbf{e}}_a(t) \\
 &= (\ddot{r}^a + (\dot{\boldsymbol{\omega}} \times \mathbf{r})^a + (\boldsymbol{\omega} \times \dot{\mathbf{r}})^a) \mathbf{e}_a(t) + (\boldsymbol{\omega} \times (\dot{\mathbf{r}} + \boldsymbol{\omega} \times \mathbf{r}))^a \mathbf{e}_a(t) \\
 &= (\ddot{r}^a + (\dot{\boldsymbol{\omega}} \times \mathbf{r})^a + 2(\boldsymbol{\omega} \times \dot{\mathbf{r}})^a + (\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r})^a) \mathbf{e}_a(t).
 \end{aligned}$$

Newton's second law for the evolution of the position vector  $\mathbf{r}(t)$  of a particle of mass  $m$  in a frame rotating with time-dependent angular velocity  $\boldsymbol{\omega}(t)$  becomes

$$\mathbf{F}(\mathbf{r}) = m \left( \ddot{\mathbf{r}} + \underbrace{\dot{\boldsymbol{\omega}} \times \mathbf{r}}_{\text{Darwin}} + \underbrace{2(\boldsymbol{\omega} \times \dot{\mathbf{r}})}_{\text{Coriolis}} + \underbrace{\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})}_{\text{centrifugal}} \right). \quad (2.1.22)$$

## 30 2 : NEWTON, LAGRANGE, HAMILTON

- The Darwin force is usually small; so it is often neglected.
- Only the Coriolis force depends on the velocity in the moving frame. The Coriolis force is very important in large-scale motions on Earth. For example, pressure balance with the Coriolis force dominates the (geostrophic) motion of weather systems that comprise the climate.
- The centrifugal force is important, for example, in obtaining orbital equilibria in gravitationally attracting systems. □

**Remark 2.1.10** The space and body angular velocities differ by

$$\widehat{\omega}_{body} := O^{-1}\dot{O} \quad \text{versus} \quad \widehat{\omega}_{space} := \dot{O}O^{-1} = O\widehat{\omega}_{body}O^{-1}.$$

Namely,  $\widehat{\omega}_{body}$  is left-invariant under  $O \rightarrow RO$  and  $\widehat{\omega}_{space}$  is right-invariant under  $O \rightarrow OR$ , for any choice of matrix  $R \in SO(3)$ . This means that neither angular velocity depends on the initial orientation. □

**Remark 2.1.11** The angular velocities  $\widehat{\omega}_{body} = O^{-1}\dot{O}$  and  $\widehat{\omega}_{space} = \dot{O}O^{-1}$  are respectively the left and right translations to the identity of the tangent matrix  $\dot{O}(t)$  at  $O(t)$ . These are called the left and right tangent spaces of  $SO(3)$  at its identity. □

**Remark 2.1.12** Equations (2.1.19) for free rigid rotations of particle systems are prototypes of Euler's equations for the motion of a rigid body. □

### 2.1.2 Newtonian form of rigid-body motion

In describing rotations of a rigid body, for example a solid object occupying a spatial domain  $\mathcal{B} \subset \mathbb{R}^3$ , one replaces the mass-weighted sums over points in space in the previous definitions of dynamical

quantities for free rigid rotation, with volume integrals weighted with a mass density as a function of position in the body. That is,

$$\sum_j m_j \rightarrow \int_{\mathcal{B}} d^3\mathbf{X} \rho(\mathbf{X}),$$

where  $\rho(\mathbf{X})$  is the mass density at a point  $\mathbf{X} \in \mathcal{B}$  fixed inside the body, as measured in coordinates whose origin is at the centre of mass.

**Example 2.1.1 (Kinetic energy of a rotating solid body)** *The kinetic energy of a solid body rotating about its centre of mass is given by*

$$K = \frac{1}{2} \int_{\mathcal{B}} \rho(\mathbf{X}) |\dot{\mathbf{x}}(\mathbf{X}, t)|^2 d^3\mathbf{X}, \quad (2.1.23)$$

where the *spatial path* in  $\mathbb{R}^3$  of a point  $\mathbf{X} \in \mathcal{B}$  in the rotating body is given by

$$\mathbf{x}(\mathbf{X}, t) = O(t)\mathbf{X} \in \mathbb{R}^3 \quad \text{with} \quad O(t) \in SO(3).$$

The time derivative of this rotating motion yields the *spatial velocity*

$$\dot{\mathbf{x}}(\mathbf{X}, t) = \dot{O}(t)\mathbf{X} = \dot{O}O^{-1}(t)\mathbf{x} =: \hat{\omega}(t)\mathbf{x} =: \boldsymbol{\omega}(t) \times \mathbf{x}, \quad (2.1.24)$$

as in Equation (2.1.7) for free rotation.

### **Kinetic energy and angular momentum of a rigid body**

The kinetic energy (2.1.23) of a rigid body rotating about its centre of mass may be expressed in the spatial frame in analogy to Equation (2.1.14) for free rotation,

$$K = \frac{1}{2} \int_{O(t)\mathcal{B}} \rho(O^{-1}(t)\mathbf{x}) |\boldsymbol{\omega}(t) \times \mathbf{x}|^2 d^3\mathbf{x}. \quad (2.1.25)$$

However, its additional time dependence makes this integral unwieldy. Instead, one takes advantage of the preservation of scalar products by rotations to write

$$|\dot{\mathbf{x}}|^2 = |O^{-1}\dot{\mathbf{x}}|^2,$$

## 32 2 : NEWTON, LAGRANGE, HAMILTON

and one computes as in Equation (2.1.24) the *body velocity*

$$O^{-1}\dot{\mathbf{x}}(\mathbf{X}, t) = O^{-1}\dot{O}(t)\mathbf{X} =: \widehat{\Omega}(t)\mathbf{X} =: \boldsymbol{\Omega}(t) \times \mathbf{X}. \quad (2.1.26)$$

Here, skew-symmetry  $\widehat{\Omega}^T = -\widehat{\Omega}$  of the matrix

$$\widehat{\Omega} = O^{-1}\dot{O} = \widehat{\omega}_{body}$$

follows because the matrix  $O$  is orthogonal, that is,  $OO^T = Id$ . Skew-symmetry of  $\widehat{\Omega}$  allows one to introduce the *body angular velocity vector*  $\boldsymbol{\Omega}(t)$  whose components  $\Omega_i$ , with  $i = 1, 2, 3$ , are given in body coordinates by

$$(O^{-1}\dot{O})_{jk} = \widehat{\Omega}_{jk} = -\Omega_i \epsilon_{ijk}. \quad (2.1.27)$$

In terms of body angular velocity vector  $\boldsymbol{\Omega}(t)$  the kinetic energy of a rigid body becomes

$$K = \frac{1}{2} \int_{\mathcal{B}} \rho(\mathbf{X}) |\boldsymbol{\Omega}(t) \times \mathbf{X}|^2 d^3\mathbf{X} =: \frac{1}{2} \langle\langle \boldsymbol{\Omega}(t), \boldsymbol{\Omega}(t) \rangle\rangle, \quad (2.1.28)$$

where  $\langle\langle \cdot, \cdot \rangle\rangle$  is a mass-weighted symmetric pairing, defined for any two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$  as the following integration over the body,

$$\langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle := \int_{\mathcal{B}} \rho(\mathbf{X}) (\mathbf{a} \times \mathbf{X}) \cdot (\mathbf{b} \times \mathbf{X}) d^3\mathbf{X}. \quad (2.1.29)$$

**Definition 2.1.9 (Moment of inertia tensor)** *The mass-weighted pairing, or inner product in (2.1.29)*

$$\langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle = \mathbb{I} \mathbf{a} \cdot \mathbf{b},$$

defines the symmetric *moment of inertia tensor*  $\mathbb{I}$  for the rigid body.

**Exercise.** By definition,  $\mathbb{I}$  is constant in the body frame. What is its time dependence in the spatial frame? ★

**Remark 2.1.13 (Principal axis frame)** The moment of inertia tensor becomes *diagonal*,

$$\mathbb{I} = \text{diag}(I_1, I_2, I_3),$$

upon aligning the body reference coordinates with its *principal axis frame*. In the principal axis coordinates of  $\mathbb{I}$ , the kinetic energy (2.1.28) takes the elegant form

$$K = \frac{1}{2} \langle\langle \boldsymbol{\Omega}, \boldsymbol{\Omega} \rangle\rangle = \frac{1}{2} \mathbb{I} \boldsymbol{\Omega} \cdot \boldsymbol{\Omega} = \frac{1}{2} (I_1 \Omega_1^2 + I_2 \Omega_2^2 + I_3 \Omega_3^2).$$

□

**Definition 2.1.10 (Body angular momentum)** The *body angular momentum* is defined as the derivative of the kinetic energy (2.1.28) with respect to body angular velocity. This produces the linear relation

$$\begin{aligned} \boldsymbol{\Pi} &= \frac{\partial K}{\partial \boldsymbol{\Omega}} = - \int_{\mathcal{B}} \rho(\mathbf{X}) \mathbf{X} \times (\mathbf{X} \times \boldsymbol{\Omega}(t)) d^3 \mathbf{X} \\ &= \left( \int_{\mathcal{B}} \rho(\mathbf{X}) (|\mathbf{X}|^2 Id - \mathbf{X} \otimes \mathbf{X}) d^3 \mathbf{X} \right) \boldsymbol{\Omega}(t) \\ &= \mathbb{I} \boldsymbol{\Omega}. \end{aligned} \quad (2.1.30)$$

This  $\mathbb{I}$  is the continuum version of the moment of inertia tensor defined for particle systems in Equation (2.1.14). It is called the *moment of inertia tensor of the rigid body*.

**Remark 2.1.14** In general, the body angular momentum vector  $\boldsymbol{\Pi}$  is not parallel to the body angular velocity vector  $\boldsymbol{\Omega}$ . Their misalignment is measured by  $\boldsymbol{\Pi} \times \boldsymbol{\Omega} \neq 0$ . □

### Angular momentum conservation

The rigid body rotates freely along  $O(t)$  in the absence of any externally applied forces or torques, so Newton's second law implies that the motion of the rigid body conserves total angular momentum when expressed in the fixed space coordinates in  $\mathbb{R}^3$ . This conservation law is expressed in the *spatial* frame as

$$\frac{d\boldsymbol{\pi}}{dt} = 0,$$

## 34 2 : NEWTON, LAGRANGE, HAMILTON

where  $\boldsymbol{\pi}(t)$  is the *angular momentum vector in space*. The angular momentum vector in space  $\boldsymbol{\pi}(t)$  is related to the angular momentum vector in the body  $\boldsymbol{\Pi}(t)$  by the mutual rotation of their axes at any time. That is,  $\boldsymbol{\pi}(t) = O(t)\boldsymbol{\Pi}(t)$ . Likewise, the *angular velocity vector in space* satisfies  $\boldsymbol{\omega}(t) = O(t)\boldsymbol{\Omega}(t)$ . The angular velocity vector in space is related to its corresponding angular momentum vector by

$$\begin{aligned}\boldsymbol{\pi}(t) = O(t)\boldsymbol{\Pi}(t) &= O(t)\mathbb{I}\boldsymbol{\Omega}(t) \\ &= (O(t)\mathbb{I}O^{-1}(t))\boldsymbol{\omega}(t) =: \mathbb{I}_{space}(t)\boldsymbol{\omega}(t).\end{aligned}$$

Thus, the *moment of inertia tensor in space*  $\mathbb{I}_{space}(t)$  transforms as a symmetric tensor,

$$\mathbb{I}_{space}(t) = O(t)\mathbb{I}O^{-1}(t), \quad (2.1.31)$$

so it is *time-dependent* and the relation of the spatial angular velocity vector  $\boldsymbol{\omega}(t)$  to the motion of the rigid body may be found by using (2.1.26), as

$$\dot{\mathbf{x}}(\mathbf{X}, t) = \dot{O}(t)\mathbf{X} =: \dot{O}O^{-1}(t)\mathbf{x} =: \hat{\boldsymbol{\omega}}(t)\mathbf{x} =: \boldsymbol{\omega}(t) \times \mathbf{x}. \quad (2.1.32)$$

As expected, the motion in space of a point at  $\mathbf{x}$  within the rigid body is a rotation by the time-dependent angular velocity  $\boldsymbol{\omega}(t)$ . We may formally confirm the relation of the spatial angular velocity vector to the body angular velocity vector  $\boldsymbol{\omega}(t) = O(t)\boldsymbol{\Omega}(t)$  by using (2.1.26) in the following calculation:

$$\begin{aligned}\boldsymbol{\Omega}(t) \times \mathbf{X} &= \hat{\boldsymbol{\Omega}}(t)\mathbf{X} = O^{-1}\dot{O}\mathbf{X} \\ &= O^{-1}\hat{\boldsymbol{\omega}}(t)\mathbf{x} = O^{-1}(\boldsymbol{\omega}(t) \times \mathbf{x}) = (O^{-1}\boldsymbol{\omega}(t) \times O^{-1}\mathbf{x}).\end{aligned}$$

Consequently, the conservation of spatial angular momentum in the absence of external torques implies

$$\begin{aligned}\frac{d\boldsymbol{\pi}}{dt} &= \frac{d}{dt}(O(t)\boldsymbol{\Pi}) \\ &= O(t)\left(\frac{d\boldsymbol{\Pi}}{dt} + O^{-1}\dot{O}\boldsymbol{\Pi}\right) \\ &= O(t)\left(\frac{d\boldsymbol{\Pi}}{dt} + \hat{\boldsymbol{\Omega}}\boldsymbol{\Pi}\right) \\ &= O(t)\left(\frac{d\boldsymbol{\Pi}}{dt} + \boldsymbol{\Omega} \times \boldsymbol{\Pi}\right) = 0.\end{aligned}$$

Hence, the body angular momentum satisfies *Euler's equations* for a rigid body,

$$\frac{d\mathbf{\Pi}}{dt} + \mathbf{\Omega} \times \mathbf{\Pi} = 0. \quad (2.1.33)$$

**Remark 2.1.15 (Body angular momentum equation)** Viewed in the moving frame, the rigid body occupies a fixed domain  $\mathcal{B}$ , so its moment of inertia tensor in that frame  $\mathbb{I}$  is constant. Its body angular momentum vector  $\mathbf{\Pi} = \mathbb{I}\mathbf{\Omega}$  evolves according to (2.1.33) by rotating around the body angular velocity vector  $\mathbf{\Omega} = \mathbb{I}^{-1}\mathbf{\Pi}$ . That is, conservation of spatial angular momentum  $\boldsymbol{\pi}(t) = O(t)\mathbf{\Pi}(t)$  relative to a fixed frame implies the body angular momentum  $\mathbf{\Pi}$  appears constant in a frame rotating with the body angular velocity  $\mathbf{\Omega} = \mathbb{I}^{-1}\mathbf{\Pi}$ .  $\square$

**Proposition 2.1.2 (Conservation laws)** *The dynamics of Equation (2.1.33) conserves both the square of the body angular momentum  $|\mathbf{\Pi}|^2$  and the kinetic energy  $K = \mathbf{\Omega} \cdot \mathbf{\Pi}/2$ .*

**Proof.** These two conservation laws may be verified by direct calculations:

$$\begin{aligned} \frac{d|\mathbf{\Pi}|^2}{dt} &= 2\mathbf{\Pi} \cdot \frac{d\mathbf{\Pi}}{dt} = 2\mathbf{\Pi} \cdot \mathbf{\Pi} \times \mathbf{\Omega} = 0, \\ \frac{d(\mathbf{\Omega} \cdot \mathbf{\Pi})}{dt} &= 2\mathbf{\Omega} \cdot \frac{d\mathbf{\Pi}}{dt} = 2\mathbf{\Omega} \cdot \mathbf{\Pi} \times \mathbf{\Omega} = 0, \end{aligned}$$

where one uses the symmetry of the moment of inertia tensor in the second line.  $\blacksquare$

**Remark 2.1.16 (Reconstruction formula)** Having found the evolution of  $\mathbf{\Pi}(t)$  and thus  $\mathbf{\Omega}(t)$  by solving (2.1.33), one may compute the net angle of rotation  $O(t)$  in body coordinates from the skew-symmetric angular velocity matrix  $\widehat{\mathbf{\Omega}}$  in (2.1.27) and its defining relation,

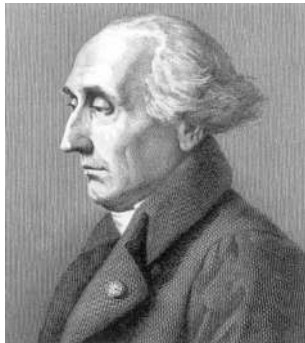
$$\dot{O}(t) = O\widehat{\mathbf{\Omega}}(t).$$

## 36 2 : NEWTON, LAGRANGE, HAMILTON

Solving this linear differential equation with time-dependent coefficients yields the paths of rotations  $O(t) \in SO(3)$ . Having these, one may finally construct the trajectories in space taken by points  $\mathbf{X}$  in the body  $\mathcal{B}$  given by  $\mathbf{x}(\mathbf{X}, t) = O(t)\mathbf{X} \in \mathbb{R}^3$ .  $\square$

## 2.2 Lagrange

### 2.2.1 The principle of stationary action



Joseph-Louis Lagrange

In Lagrangian mechanics, a mechanical system in a *configuration space* with generalised coordinates and velocities,

$$q^a, \dot{q}^a, \quad a = 1, 2, \dots, 3N,$$

is characterised by its Lagrangian  $L(q(t), \dot{q}(t))$  – a smooth, real-valued function. The motion of a Lagrangian system is determined by the principle of stationary action, formulated using the operation of variational derivative.

**Definition 2.2.1 (Variational derivative)** *The variational derivative of a functional  $S[q]$  is defined as its linearisation in an arbitrary direction  $\delta q$  in the configuration space. That is,  $S[q]$  is defined as*

$$\delta S[q]; := \lim_{s \rightarrow 0} \frac{S[q + s\delta q] - S[q]}{s} = \left. \frac{d}{ds} \right|_{s=0} S[q + s\delta q] =: \left\langle \frac{\delta S}{\delta q}, \delta q \right\rangle,$$

where the pairing  $\langle \cdot, \cdot \rangle$  is obtained in the process of linearisation.

**Theorem 2.2.1 (Principle of stationary action)** *The Euler–Lagrange equations,*

$$[L]_{q^a} := \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} = 0, \quad (2.2.1)$$



follow from stationarity of the **action integral**,  $S$ , defined as the integral over a time interval  $t \in (t_1, t_2)$ ,

$$S := \int_{t_1}^{t_2} L(q, \dot{q}) dt. \quad (2.2.2)$$

Then the **principle of stationary action**,

$$\delta S = 0,$$

implies  $[L]_{q^a} = 0$ , for variations  $\delta q^a$  that vanish at the endpoints in time.

**Proof.** Applying the variational derivative in Definition 2.2.1 to the action integral in (2.2.2) yields

$$\begin{aligned} \delta S[q] &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^a} \delta q^a + \frac{\partial L}{\partial \dot{q}^a} \delta \dot{q}^a \right) dt \\ &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^a} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^a} \right) \delta q^a dt + \left[ \frac{\partial L}{\partial \dot{q}^a} \delta q^a \right]_{t_1}^{t_2} \\ &=: - \int_{t_1}^{t_2} [L]_{q^a} \delta q^a dt. \end{aligned} \quad (2.2.3)$$

Here one integrates by parts and in the last step applies the condition that the variations  $\delta q^a$  vanish at the endpoints in time. Because the variations  $\delta q^a$  are otherwise arbitrary, one concludes that the Euler–Lagrange Equations (2.2.1) are satisfied. ■

**Remark 2.2.1** The principle of stationary action is sometimes also called **Hamilton’s principle**. □

**Example 2.2.1 (Simple mechanical systems)** *The Lagrangian for the motion of a simple mechanical system is given in the separated form,*

$$L(q, \dot{q}) = \frac{m}{2} |\dot{q}|^2 - V(q) \quad \text{with Euclidean norm} \quad |\dot{q}|^2 := \dot{q}^b \delta_{bc} \dot{q}^c.$$

The Lagrangian in this case has partial derivatives

$$\frac{\partial L}{\partial \dot{q}^a} = m \delta_{ac} \dot{q}^c \quad \text{and} \quad \frac{\partial L}{\partial q^a} = - \frac{\partial V(q)}{\partial q^a}.$$

**38** 2 : NEWTON, LAGRANGE, HAMILTON

Consequently, its Euler–Lagrange equations  $[L]_{q^a} = 0$  are

$$\begin{aligned} [L]_{q^a} &:= \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} \\ &= m \delta_{ac} \ddot{q}^c + \frac{\partial V(q)}{\partial q^a} = 0. \end{aligned}$$

This is Newton’s law of acceleration for a potential force.

**Example 2.2.2 (Geodesic motion in a Riemannian space)** *The Lagrangian for the motion of a free particle in a Riemannian space is its kinetic energy with respect to the Riemannian metric,*

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^b g_{bc}(q) \dot{q}^c.$$

The Lagrangian in this case has partial derivatives

$$\frac{\partial L}{\partial \dot{q}^a} = g_{ac}(q) \dot{q}^c \quad \text{and} \quad \frac{\partial L}{\partial q^a} = \frac{1}{2} \frac{\partial g_{bc}(q)}{\partial q^a} \dot{q}^b \dot{q}^c.$$

Consequently, its Euler–Lagrange equations  $[L]_{q^a} = 0$  are

$$\begin{aligned} [L]_{q^a} &:= \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} \\ &= g_{ae}(q) \ddot{q}^e + \frac{\partial g_{ae}(q)}{\partial q^b} \dot{q}^b \dot{q}^e - \frac{1}{2} \frac{\partial g_{bc}(q)}{\partial q^a} \dot{q}^b \dot{q}^e = 0. \end{aligned}$$

Symmetrising the middle term and contracting with co-metric  $g^{ca}$  satisfying  $g^{ca} g_{ae} = \delta_e^c$  yields

$$\ddot{q}^c + \Gamma_{be}^c(q) \dot{q}^b \dot{q}^e = 0, \quad (2.2.4)$$

where  $\Gamma_{be}^c$  are the **Christoffel symbols**, given in terms of the metric by

$$\Gamma_{be}^c(q) = \frac{1}{2} g^{ca} \left[ \frac{\partial g_{ae}(q)}{\partial q^b} + \frac{\partial g_{ab}(q)}{\partial q^e} - \frac{\partial g_{be}(q)}{\partial q^a} \right]. \quad (2.2.5)$$

These Euler–Lagrange equations are the **geodesic equations** of a free particle moving in a Riemannian space.

## 2.3 Noether's theorem

### 2.3.1 Lie symmetries and conservation laws



Emmy Noether

Recall from Definition 1.2.3 that a Lie group depends smoothly on its parameters. (See Appendix B for more details.)

**Definition 2.3.1 (Lie symmetry)** A smooth transformation of variables  $\{t, q\}$  depending on a single parameter  $s$  defined by

$$\{t, q\} \mapsto \{\bar{t}(t, q, s), \bar{q}(t, q, s)\},$$

that leaves the action  $S = \int L dt$  invariant is called a **Lie symmetry** of the action.

**Theorem 2.3.1 (Noether's theorem)** Each Lie symmetry of the action for a Lagrangian system defined on a manifold  $M$  with Lagrangian  $L$  corresponds to a constant of the motion [No1918].

**Example 2.3.1** Suppose the variation of the action in (2.2.3) vanishes ( $\delta S = 0$ ) because of a Lie symmetry which does not preserve the endpoints. Then on solutions of the Euler–Lagrange equations, the endpoint term must vanish for another reason. For example, if the Lie symmetry leaves time invariant, so that

$$\{t, q\} \mapsto \{t, \bar{q}(t, q, s)\},$$

then the endpoint term must vanish,

$$\left[ \frac{\partial L}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2} = 0.$$

Hence, the quantity

$$A(q, \dot{q}, \delta q) = \frac{\partial L}{\partial \dot{q}^a} \delta q^a$$

## 40 2 : NEWTON, LAGRANGE, HAMILTON

is a **constant of motion** for solutions of the Euler–Lagrange equations. In particular, if  $\delta q^a = c^a$  for constants  $c^a$ ,  $a = 1, \dots, n$ , that is, for spatial translations in  $n$  dimensions, then the quantities  $\partial L / \partial \dot{q}^a$  (the corresponding momentum components) are constants of motion.

**Remark 2.3.1** This result first appeared in Noether [No1918]. See, e.g., [Ol2000, SaCa1981] for good discussions of the history, framework and applications of Noether’s theorem. We shall see in a moment that Lie symmetries that reparameterise time may also yield constants of motion.  $\square$

### 2.3.2 Infinitesimal transformations of a Lie group

#### Definition 2.3.2 (Infinitesimal Lie transformations)



Sophus Lie

Consider the Lie group of transformations

$$\{t, q\} \mapsto \{\bar{t}(t, q, s), \bar{q}(t, q, s)\},$$

and suppose the identity transformation is arranged to occur for  $s = 0$ . The derivatives with respect to the group parameters  $s$  at the identity,

$$\begin{aligned} \tau(t, q) &= \left. \frac{d}{ds} \right|_{s=0} \bar{t}(t, q, s), \\ \xi^a(t, q) &= \left. \frac{d}{ds} \right|_{s=0} \bar{q}^a(t, q, s), \end{aligned}$$

are called the **infinitesimal transformations** of the action of a Lie group on the time and space variables.

Thus, at linear order in a Taylor expansion in the group parameter  $s$  one has

$$\bar{t} = t + s\tau(t, q), \quad \bar{q}^a = q^a + s\xi^a(t, q), \quad (2.3.1)$$

where  $\tau$  and  $\xi^a$  are functions of coordinates and time, but do not depend on velocities. Then, to first order in  $s$  the velocities of the

transformed trajectories are computed as

$$\frac{d\bar{q}^a}{d\bar{t}} = \frac{\dot{q}^a + s\dot{\xi}^a}{1 + s\dot{\tau}} = \dot{q}^a + s(\dot{\xi}^a - \dot{q}^a\dot{\tau}), \quad (2.3.2)$$

where order  $O(s^2)$  terms are neglected and one defines the total time derivatives

$$\dot{\tau} \equiv \frac{\partial \tau}{\partial t}(t, q) + \dot{q}^b \frac{\partial \tau}{\partial q^b}(t, q) \quad \text{and} \quad \dot{\xi}^a \equiv \frac{\partial \xi^a}{\partial t}(t, q) + \dot{q}^b \frac{\partial \xi^a}{\partial q^b}(t, q).$$

**Remark 2.3.2** The result (2.3.2) for the infinitesimal transformation of the derivative when both the independent and dependent variables are transformed is the operation introduced by Lie of *prolongation* of the infinitesimal Lie group actions (2.3.1). Prolongation is a generalisation of the tangent lift explained in Definition A.3.3.  $\square$

We are now in a position to prove Noether's Theorem 2.3.1.

**Proof.** The variation of the action corresponding to the Lie symmetry with infinitesimal transformations (2.3.1) is

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial t} \delta t + \frac{\partial L}{\partial q^a} \delta q^a + \frac{\partial L}{\partial \dot{q}^a} \delta \dot{q}^a + L \frac{d\delta t}{dt} \right) dt \\ &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial t} \tau + \frac{\partial L}{\partial q^a} \xi^a + \frac{\partial L}{\partial \dot{q}^a} (\dot{\xi}^a - \dot{q}^a \dot{\tau}) + L \dot{\tau} \right) dt \\ &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^a} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^a} \right) (\xi^a - \dot{q}^a \tau) + \frac{d}{dt} \left( L \tau + \frac{\partial L}{\partial \dot{q}^a} (\xi^a - \dot{q}^a \tau) \right) dt \\ &= \int_{t_1}^{t_2} [L]_{q^a} (\xi^a - \dot{q}^a \tau) + \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}^a} \xi^a - \left( \frac{\partial L}{\partial \dot{q}^a} \dot{q}^a - L \right) \tau \right] dt. \end{aligned}$$

Thus, stationarity  $\delta S = 0$  and the Euler-Lagrange equations  $[L]_{q^a} = 0$  imply

$$0 = \left[ \frac{\partial L}{\partial \dot{q}^a} \xi^a - \left( \frac{\partial L}{\partial \dot{q}^a} \dot{q}^a - L \right) \tau \right]_{t_1}^{t_2},$$

## 42 2 : NEWTON, LAGRANGE, HAMILTON

so that the quantity

$$C(t, q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^a} \xi^a - \left( \frac{\partial L}{\partial \dot{q}^a} \dot{q}^a - L \right) \tau \quad (2.3.3)$$

$$\equiv \langle p, \delta q \rangle - H \delta t \quad (2.3.4)$$

has the same value at every time along the solution path. That is,  $C(t, q, \dot{q})$  is a constant of the motion. ■

**Remark 2.3.3** The abbreviated notation in Equation (2.3.4) for  $\delta q$  and  $\delta t$  is standard. If  $\delta t$  is absent and  $\delta q$  is a constant (corresponding to translations in space in a certain direction) then  $\delta S = 0$  implies that the *linear momentum*  $p = \partial L / \partial \dot{q}$  in that direction is conserved for solutions of the Euler–Lagrange equations  $[L]_q = 0$ . □

**Exercise.** Show that conservation of energy results from Noether’s theorem if, in Hamilton’s principle, the variations are chosen as

$$\delta q(t) = \left. \frac{d}{ds} \right|_{s=0} q(\bar{t}(t, q, s)),$$

corresponding to symmetry of the Lagrangian under space-time-dependent transformations of time along a given curve, so that  $t \rightarrow \bar{t}(t, q, s)$  and  $q(t) \rightarrow q(\bar{t}(t, q, s))$  with  $\bar{t}(t, q, 0) = t$ . ★

**Answer.** Under reparameterisations of time along the curve

$$q(t) \rightarrow q(\bar{t}(t, q, s)),$$

the action  $S = \int_{t_1}^{t_2} L(q, \dot{q}) dt$  changes infinitesimally according to

$$\delta S = \left[ \left( L(q, \dot{q}) - \frac{\partial L}{\partial \dot{q}} \dot{q} \right) \delta t \right]_{t_1}^{t_2},$$

with variations in position and time defined by

$$\begin{aligned}\delta q(t) &= \left. \frac{d}{ds} \right|_{s=0} q(\bar{t}(t, q, s)) \\ &= \dot{q}(t)\delta t \quad \text{and} \quad \delta t = \left. \frac{d\bar{t}(t, q, s)}{ds} \right|_{s=0}.\end{aligned}$$

For translations in time,  $\delta t$  is a constant and stationarity of the action  $\delta S = 0$  implies that the energy

$$E(t, q, \dot{q}) \equiv \frac{\partial L}{\partial \dot{q}^a} \dot{q}^a - L \quad (2.3.5)$$

is a constant of motion along solutions of the Euler–Lagrange equations.  $\blacktriangle$

**Exercise. (Infinitesimal Galilean transformations)** From their finite transformations in Definition 1.2.1, compute the infinitesimal transformations of the Galilean group under composition of first rotations, then boosts, then translations in space and time, in the case when they act on a velocity-space-time point  $(\dot{\mathbf{q}}, \mathbf{q}, t)$ .  $\star$

**Answer.** The composition of translations  $g_1(\mathbf{q}_0(s), t_0(s))$ , Galilean boosts  $g_3(\mathbf{v}_0(s))$  and rotations  $g_2(O(s))$  acting on a velocity-space-time point  $(\dot{\mathbf{q}}, \mathbf{q}, t)$  is given by

$$g_1 g_3 g_2(\dot{\mathbf{q}}, \mathbf{q}, t) = \left( O(s)\dot{\mathbf{q}} + \mathbf{v}_0(s), O(s)\mathbf{q} + t\mathbf{v}_0(s) + \mathbf{q}_0(s), t + t_0(s) \right).$$

One computes the infinitesimal transformations as

$$\begin{aligned}\tau &= \left. \frac{dt}{ds} \right|_{s=0} = t_0, \\ \boldsymbol{\xi} &= \left. \frac{d\mathbf{q}}{ds} \right|_{s=0} = \mathbf{q}_0 + \mathbf{v}_0 t + \boldsymbol{\Xi} \times \mathbf{q}, \\ \dot{\boldsymbol{\xi}} - \dot{\mathbf{q}}\dot{\tau} &= \left. \frac{d\dot{\mathbf{q}}}{ds} \right|_{s=0} = \mathbf{v}_0 + \boldsymbol{\Xi} \times \dot{\mathbf{q}}.\end{aligned}$$

## 44 2 : NEWTON, LAGRANGE, HAMILTON

The infinitesimal velocity transformation may also be computed from Equation (2.3.2).

Consequently, the infinitesimal transformation by the Galilean group of a function  $F(t, \mathbf{q}, \dot{\mathbf{q}})$  is given by operation of the following vector field, obtained as the first term in a Taylor series,

$$\begin{aligned} \left. \frac{d}{ds} \right|_{s=0} F(t(s), \mathbf{q}(s), \dot{\mathbf{q}}(s)) & \quad (2.3.6) \\ &= t_0 \frac{\partial F}{\partial t} + (\mathbf{q}_0 + \mathbf{v}_0 t + \boldsymbol{\Xi} \times \mathbf{q}) \cdot \frac{\partial F}{\partial \mathbf{q}} \\ & \quad + (\mathbf{v}_0 + \boldsymbol{\Xi} \times \dot{\mathbf{q}}) \cdot \frac{\partial F}{\partial \dot{\mathbf{q}}}. \end{aligned}$$

▲

**Exercise. (Galilean Lie symmetries)** Since the Galilean transformations form a Lie group, one may expect them to be a source of Lie symmetries of the Lagrangian in Hamilton's principle. Compute the corresponding Noether conservation laws. ★

**Answer.** As we have already seen, symmetries under space and time translations imply conservation of linear momentum and energy, respectively. Likewise, symmetry under rotations implies conservation of angular momentum.

Suppose the Lagrangian in Hamilton's principle (2.2.2) is invariant under  $S^1$  rotations about a spatial direction  $\boldsymbol{\Xi}$ . The infinitesimal transformation of such a rotation is  $\delta \mathbf{q} = \boldsymbol{\Xi} \times \mathbf{q}$ . In this case, the conserved Noether quantity (2.3.4) is

$$J^{\boldsymbol{\Xi}}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \delta \mathbf{q} = \mathbf{q} \times \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \boldsymbol{\Xi}, \quad (2.3.7)$$



which is the angular momentum about the  $\Xi$ -axis.

Finally, symmetry under Galilean boosts implies vanishing of total momentum for a system of  $N$  particles, labelled by an index  $j = 1, 2, \dots, N$ . The last statement may be proved explicitly from Noether's theorem and the infinitesimal Galilean boost transformations, as follows:

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial t} \delta t + \sum_j \left( \frac{\partial L}{\partial \mathbf{q}^j} \cdot \delta \mathbf{q}^j + \frac{\partial L}{\partial \dot{\mathbf{q}}^j} \cdot \delta \dot{\mathbf{q}}^j \right) + L \frac{d\delta t}{dt} \right) dt \\ &= \int_{t_1}^{t_2} \sum_j \left( \frac{\partial L}{\partial \mathbf{q}^j} \cdot \mathbf{v}_0 t + \frac{\partial L}{\partial \dot{\mathbf{q}}^j} \cdot \mathbf{v}_0 \right) dt \\ &= \int_{t_1}^{t_2} \sum_j \left( \frac{\partial L}{\partial \mathbf{q}^j} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}^j} \right) \cdot \mathbf{v}_0 t + \frac{d}{dt} \sum_j \left( \frac{\partial L}{\partial \dot{\mathbf{q}}^j} \cdot \mathbf{v}_0 t \right) dt \\ &= \int_{t_1}^{t_2} - \sum_j [L]_{\mathbf{q}^j} \cdot \mathbf{v}_0 t dt + \left[ \left( \sum_j \frac{\partial L}{\partial \dot{\mathbf{q}}^j} \right) \cdot \mathbf{v}_0 t \right]_{t_1}^{t_2}. \end{aligned}$$

So the Euler-Lagrange equations  $[L]_{\mathbf{q}^j} = 0$  and stationarity  $\delta S = 0$  for any time  $t \in [t_1, t_2]$  together imply

$$\mathbf{P} t \cdot \mathbf{v}_0 = \text{constant}, \quad \text{with} \quad \mathbf{P} := \sum_j \frac{\partial L}{\partial \dot{\mathbf{q}}^j}.$$

Let's explore the meaning of this result for a system of  $N$  particles with constant total mass  $M = \sum_j m_j$ . The centre of mass of the system is defined as

$$\mathbf{Q}_{CM} := M^{-1} \sum_j m_j \mathbf{q}^j.$$

For a simple mechanical system with Lagrangian

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \left( \sum_j m_j |\dot{\mathbf{q}}^j|^2 \right) - V(\{\mathbf{q}^j\}),$$

## 46 2 : NEWTON, LAGRANGE, HAMILTON

one finds that

$$\begin{aligned}\mathbf{P} &= \sum_j \frac{\partial L}{\partial \dot{\mathbf{q}}^j} = \sum_j m_j \dot{\mathbf{q}}^j \\ &= \frac{d}{dt} \sum_j m_j \mathbf{q}^j(t) =: \frac{d}{dt} (M \mathbf{Q}_{CM}).\end{aligned}$$

Hence, for such a system Noether's theorem yields

$$\mathbf{P}t - M \mathbf{Q}_{CM} = 0,$$

provided the action principle is *also* invariant under spatial translations, so that  $\mathbf{P}$  is *also* a constant of the motion. Consequently, the motion of a simple mechanical system of particles may always be taken as being relative to a fixed centre of mass.  $\blacktriangle$

## 2.4 Lagrangian form of rigid-body motion

In the absence of external torques, Euler's equations in (2.1.33) for rigid-body motion in principal axis coordinates are

$$\begin{aligned}I_1 \dot{\Omega}_1 &= (I_2 - I_3) \Omega_2 \Omega_3, \\ I_2 \dot{\Omega}_2 &= (I_3 - I_1) \Omega_3 \Omega_1, \\ I_3 \dot{\Omega}_3 &= (I_1 - I_2) \Omega_1 \Omega_2,\end{aligned}\tag{2.4.1}$$

or, equivalently,

$$\mathbb{I} \dot{\boldsymbol{\Omega}} = \mathbb{I} \boldsymbol{\Omega} \times \boldsymbol{\Omega},\tag{2.4.2}$$

where  $\boldsymbol{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$  is the body angular velocity vector and  $I_1, I_2, I_3$  are the moments of inertia in the principal axis frame of the rigid body. We ask whether these equations may be expressed using Hamilton's principle on  $\mathbb{R}^3$ . For this, we will need to define the variational derivative of a functional  $S[(\boldsymbol{\Omega})]$ .

**Definition 2.4.1 (Variational derivative)** *The variational derivative of a functional  $S[\boldsymbol{\Omega}]$  is defined as its linearisation in an arbitrary direction  $\delta\boldsymbol{\Omega}$  in the vector space of body angular velocities. That is,*

$$\delta S[\boldsymbol{\Omega}] := \lim_{s \rightarrow 0} \frac{S[\boldsymbol{\Omega} + s\delta\boldsymbol{\Omega}] - S[\boldsymbol{\Omega}]}{s} = \left. \frac{d}{ds} \right|_{s=0} S[\boldsymbol{\Omega} + s\delta\boldsymbol{\Omega}] =: \left\langle \frac{\delta S}{\delta\boldsymbol{\Omega}}, \delta\boldsymbol{\Omega} \right\rangle,$$

where the new pairing, also denoted as  $\langle \cdot, \cdot \rangle$ , is between the space of body angular velocities and its dual, the space of body angular momenta.

**Theorem 2.4.1 (Euler's rigid-body equations)** *Euler's rigid-body equations are equivalent to Hamilton's principle*

$$\delta S(\boldsymbol{\Omega}) = \delta \int_a^b l(\boldsymbol{\Omega}) dt = 0, \quad (2.4.3)$$

in which the Lagrangian  $l(\boldsymbol{\Omega})$  appearing in the **action integral**  $S(\boldsymbol{\Omega}) = \int_a^b l(\boldsymbol{\Omega}) dt$  is given by the kinetic energy in principal axis coordinates,

$$l(\boldsymbol{\Omega}) = \frac{1}{2} \langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega} \rangle = \frac{1}{2} \mathbb{I}\boldsymbol{\Omega} \cdot \boldsymbol{\Omega} = \frac{1}{2} (I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2), \quad (2.4.4)$$

and variations of  $\boldsymbol{\Omega}$  are restricted to be of the form

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega} \times \boldsymbol{\Xi}, \quad (2.4.5)$$

where  $\boldsymbol{\Xi}(t)$  is a curve in  $\mathbb{R}^3$  that vanishes at the endpoints in time.

**Proof.** Since  $l(\boldsymbol{\Omega}) = \frac{1}{2} \langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega} \rangle$ , and  $\mathbb{I}$  is symmetric, one obtains

$$\begin{aligned} \delta \int_a^b l(\boldsymbol{\Omega}) dt &= \int_a^b \left\langle \mathbb{I}\boldsymbol{\Omega}, \delta\boldsymbol{\Omega} \right\rangle dt \\ &= \int_a^b \left\langle \mathbb{I}\boldsymbol{\Omega}, \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega} \times \boldsymbol{\Xi} \right\rangle dt \\ &= \int_a^b \left[ \left\langle -\frac{d}{dt} \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Xi} \right\rangle + \left\langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega} \times \boldsymbol{\Xi} \right\rangle \right] dt \\ &= \int_a^b \left\langle -\frac{d}{dt} \mathbb{I}\boldsymbol{\Omega} + \mathbb{I}\boldsymbol{\Omega} \times \boldsymbol{\Omega}, \boldsymbol{\Xi} \right\rangle dt + \left\langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Xi} \right\rangle \Big|_{t_a}^{t_b}, \end{aligned}$$

## 48 2 : NEWTON, LAGRANGE, HAMILTON

upon integrating by parts. The last term vanishes, upon using the endpoint conditions,

$$\Xi(a) = 0 = \Xi(b).$$

Since  $\Xi$  is otherwise arbitrary, (2.4.3) is equivalent to

$$-\frac{d}{dt}(\mathbb{I}\Omega) + \mathbb{I}\Omega \times \Omega = 0,$$

which recovers Euler's Equations (2.4.1) in vector form. ■

**Proposition 2.4.1 (Derivation of the restricted variation)** *The restricted variation in (2.4.5) arises via the following steps:*

- (i) Vary the definition of body angular velocity,  $\widehat{\Omega} = O^{-1}\dot{O}$ .
- (ii) Take the time derivative of the variation,  $\widehat{\Xi} = O^{-1}O'$ .
- (iii) Use the equality of cross derivatives,  $O'^{\cdot} = d^2O/dtds = O'^{\cdot}$ .
- (iv) Apply the hat map.

**Proof.** One computes directly that

$$\begin{aligned}\widehat{\Omega}' &= (O^{-1}\dot{O})' = -O^{-1}O'O^{-1}\dot{O} + O^{-1}O'^{\cdot} = -\widehat{\Xi}\widehat{\Omega} + O^{-1}O'^{\cdot}, \\ \widehat{\Xi}^{\cdot} &= (O^{-1}O')^{\cdot} = -O^{-1}\dot{O}O^{-1}O' + O^{-1}O'^{\cdot} = -\widehat{\Omega}\widehat{\Xi} + O^{-1}O'^{\cdot}.\end{aligned}$$

On taking the difference, the cross derivatives cancel and one finds a variational formula equivalent to (2.4.5),

$$\widehat{\Omega}' - \widehat{\Xi}^{\cdot} = [\widehat{\Omega}, \widehat{\Xi}] \quad \text{with} \quad [\widehat{\Omega}, \widehat{\Xi}] := \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega}. \quad (2.4.6)$$

Under the bracket relation (2.1.12) for the hat map, this equation recovers the vector relation (2.4.5) in the form

$$\Omega' - \dot{\Xi} = \Omega \times \Xi. \quad (2.4.7)$$

## 2.4 LAGRANGIAN FORM OF RIGID-BODY MOTION 49

Thus, Euler's equations for the rigid body in  $T\mathbb{R}^3$ ,

$$\mathbb{I}\dot{\Omega} = \mathbb{I}\Omega \times \Omega, \quad (2.4.8)$$

do follow from the variational principle (2.4.3) with variations of the form (2.4.5) derived from the definition of body angular velocity  $\hat{\Omega}$ . ■

**Remark 2.4.1** The body angular velocity is expressed in terms of the spatial angular velocity by  $\Omega(t) = O^{-1}(t)\omega(t)$ . Consequently, the kinetic energy Lagrangian in (2.4.4) transforms as

$$l(\Omega) = \frac{1}{2} \Omega \cdot \mathbb{I}\Omega = \frac{1}{2} \omega \cdot \mathbb{I}_{space}(t)\omega =: l_{space}(\omega),$$

where

$$\mathbb{I}_{space}(t) = O(t)\mathbb{I}O^{-1}(t),$$

as in (2.1.31). □

**Exercise.** Show that Hamilton's principle for the action

$$S(\omega) = \int_a^b l_{space}(\omega) dt$$

yields conservation of spatial angular momentum

$$\pi = \mathbb{I}_{space}(t)\omega(t).$$

Hint: First derive  $\delta\mathbb{I}_{space} = [\xi, \mathbb{I}_{space}]$  with right-invariant  $\xi = \delta OO^{-1}$ . ★

**Exercise. (Noether's theorem for the rigid body)** What conservation law does Theorem 2.3.1 (Noether's theorem) imply for the rigid-body Equations (2.4.2)?

## 50 2 : NEWTON, LAGRANGE, HAMILTON

Hint: Transform the endpoint terms arising on integrating the variation  $\delta S$  by parts in the proof of Theorem 2.4.1 into the spatial representation by setting  $\Xi = O^{-1}(t)\Gamma$  and  $\Omega = O^{-1}(t)\omega$ . ★

**Remark 2.4.2 (Reconstruction of  $O(t) \in SO(3)$ )** The Euler solution is expressed in terms of the time-dependent angular velocity vector in the body,  $\Omega$ . The body angular velocity vector  $\Omega(t)$  yields the tangent vector  $\dot{O}(t) \in T_{O(t)}SO(3)$  along the integral curve in the rotation group  $O(t) \in SO(3)$  by the relation

$$\dot{O}(t) = O(t)\widehat{\Omega}(t), \quad (2.4.9)$$

where the left-invariant skew-symmetric  $3 \times 3$  matrix  $\widehat{\Omega}$  is defined by the hat map (2.1.27)

$$(O^{-1}\dot{O})_{jk} = \widehat{\Omega}_{jk} = -\Omega_i \epsilon_{ijk}. \quad (2.4.10)$$

Equation (2.4.9) is the *reconstruction formula* for  $O(t) \in SO(3)$ .

Once the time dependence of  $\Omega(t)$  and hence  $\widehat{\Omega}(t)$  is determined from the Euler equations, solving formula (2.4.9) as a linear differential equation with time-dependent coefficients yields the integral curve  $O(t) \in SO(3)$  for the orientation of the rigid body. □

### 2.4.1 Hamilton–Pontryagin constrained variations

Formula (2.4.6) for the variation  $\widehat{\Omega}$  of the skew-symmetric matrix

$$\widehat{\Omega} = O^{-1}\dot{O}$$

## 2.4 LAGRANGIAN FORM OF RIGID-BODY MOTION 51

may be imposed as a constraint in Hamilton's principle and thereby provide a variational derivation of Euler's Equations (2.1.33) for rigid-body motion in principal axis coordinates. This constraint is incorporated into the matrix Euler equations, as follows.

**Proposition 2.4.2 (Matrix Euler equations)** *Euler's rigid-body equation may be written in matrix form as*

$$\frac{d\Pi}{dt} = - [\widehat{\Omega}, \Pi] \quad \text{with} \quad \Pi = \mathbb{I}\widehat{\Omega} = \frac{\delta l}{\delta \widehat{\Omega}}, \quad (2.4.11)$$

for the Lagrangian  $l(\widehat{\Omega})$  given by

$$l = \frac{1}{2} \langle \mathbb{I}\widehat{\Omega}, \widehat{\Omega} \rangle. \quad (2.4.12)$$

Here, the bracket

$$[\widehat{\Omega}, \Pi] := \widehat{\Omega}\Pi - \Pi\widehat{\Omega} \quad (2.4.13)$$

denotes the commutator and  $\langle \cdot, \cdot \rangle$  denotes the **trace pairing**, e.g.,

$$\langle \Pi, \widehat{\Omega} \rangle =: \frac{1}{2} \text{trace} (\Pi^T \widehat{\Omega}). \quad (2.4.14)$$

**Remark 2.4.3** Note that the symmetric part of  $\Pi$  does not contribute in the pairing and if set equal to zero initially, it will remain zero.  $\square$

**Proposition 2.4.3 (Constrained variational principle)** *The matrix Euler Equations (2.4.11) are equivalent to stationarity  $\delta S = 0$  of the following constrained action:*

$$\begin{aligned} S(\widehat{\Omega}, O, \dot{O}, \Pi) &= \int_a^b l(\widehat{\Omega}, O, \dot{O}, \Pi) dt & (2.4.15) \\ &= \int_a^b \left[ l(\widehat{\Omega}) + \langle \Pi, (O^{-1}\dot{O} - \widehat{\Omega}) \rangle \right] dt. \end{aligned}$$

**Remark 2.4.4** The integrand of the constrained action in (2.4.15) is similar to the formula for the Legendre transform, but its functional dependence is different. This variational approach is related

## 52 2 : NEWTON, LAGRANGE, HAMILTON

to the classic *Hamilton–Pontryagin principle* for control theory in [YoMa2006]. It is also used in [BoMa2009] to develop algorithms for geometric numerical integrations of rotating motion.  $\square$

**Proof.** The variations of  $S$  in formula (2.4.15) are given by

$$\begin{aligned} \delta S = \int_a^b \left\{ \left\langle \frac{\delta l}{\delta \widehat{\Omega}} - \Pi, \delta \widehat{\Omega} \right\rangle \right. \\ \left. + \left\langle \delta \Pi, (O^{-1}\dot{O} - \Omega) \right\rangle + \left\langle \Pi, \delta(O^{-1}\dot{O}) \right\rangle \right\} dt, \end{aligned}$$

where

$$\delta(O^{-1}\dot{O}) = \widehat{\Xi} \cdot + [\widehat{\Omega}, \widehat{\Xi}], \quad (2.4.16)$$

and  $\widehat{\Xi} = (O^{-1}\delta O)$  from Equation (2.4.6).

Substituting for  $\delta(O^{-1}\dot{O})$  into the last term of  $\delta S$  produces

$$\begin{aligned} \int_a^b \left\langle \Pi, \delta(O^{-1}\dot{O}) \right\rangle dt &= \int_a^b \left\langle \Pi, \widehat{\Xi} \cdot + [\widehat{\Omega}, \widehat{\Xi}] \right\rangle dt \\ &= \int_a^b \left\langle -\Pi \cdot - [\widehat{\Omega}, \Pi], \widehat{\Xi} \right\rangle dt \\ &\quad + \left\langle \Pi, \widehat{\Xi} \right\rangle \Big|_a^b, \end{aligned} \quad (2.4.17)$$

where one uses the cyclic properties of the trace operation for matrices,

$$\text{trace} \left( \Pi^T \widehat{\Xi} \widehat{\Omega} \right) = \text{trace} \left( \widehat{\Omega} \Pi^T \widehat{\Xi} \right). \quad (2.4.18)$$

Thus, stationarity of the Hamilton–Pontryagin variational principle for vanishing endpoint conditions  $\widehat{\Xi}(a) = 0 = \widehat{\Xi}(b)$  implies the following set of equations:

$$\frac{\delta l}{\delta \widehat{\Omega}} = \Pi, \quad O^{-1}\dot{O} = \widehat{\Omega}, \quad \Pi \cdot = -[\widehat{\Omega}, \Pi]. \quad (2.4.19)$$

■



**Remark 2.4.5 (Interpreting the formulas in (2.4.19))** The first formula in (2.4.19) defines the angular momentum matrix  $\Pi$  as the *fibre derivative* of the Lagrangian with respect to the angular velocity matrix  $\widehat{\Omega}$ . The second formula is the reconstruction formula (2.4.9) for the solution curve  $O(t) \in SO(3)$ , given the solution  $\widehat{\Omega}(t) = O^{-1}\dot{O}$ . And the third formula is Euler's equation for rigid-body motion in matrix form.  $\square$

We transform the endpoint terms in (2.4.20), arising on integrating the variation  $\delta S$  by parts in the proof of Theorem 2.4.3, into the spatial representation by setting  $\widehat{\Xi}(t) =: O(t)\widehat{\xi}O^{-1}(t)$  and  $\widehat{\Pi}(t) =: O(t)\widehat{\pi}(t)O^{-1}(t)$ , as follows:

$$\langle \Pi, \widehat{\Xi} \rangle = \text{trace} \left( \Pi^T \widehat{\Xi} \right) = \text{trace} \left( \pi^T \widehat{\xi} \right) = \langle \pi, \widehat{\xi} \rangle. \quad (2.4.20)$$

Thus, the vanishing of both endpoints for a *constant* infinitesimal spatial rotation  $\widehat{\xi} = (\delta O O^{-1}) = \text{const}$  implies

$$\pi(a) = \pi(b). \quad (2.4.21)$$

This is Noether's theorem for the rigid body.

**Theorem 2.4.2 (Noether's theorem for the rigid body)** *Invariance of the constrained Hamilton–Pontryagin action under spatial rotations implies conservation of spatial angular momentum,*

$$\pi = O^{-1}(t)\Pi(t)O(t) =: \text{Ad}_{O^{-1}(t)}^* \Pi(t). \quad (2.4.22)$$

**Proof.**

$$\begin{aligned} \frac{d}{dt} \langle \pi, \widehat{\xi} \rangle &= \frac{d}{dt} \langle O^{-1}\Pi O, \widehat{\xi} \rangle = \frac{d}{dt} \text{trace} \left( \Pi^T O^{-1}\widehat{\xi}O \right) \\ &= \left\langle \frac{d}{dt} \Pi + [\widehat{\Omega}, \Pi], O^{-1}\widehat{\xi}O \right\rangle = 0 \\ &=: \left\langle \frac{d}{dt} \Pi - \text{ad}_{\widehat{\Omega}}^* \Pi, \text{Ad}_{O^{-1}} \widehat{\xi} \right\rangle, \\ \frac{d}{dt} \langle \text{Ad}_{O^{-1}}^* \Pi, \widehat{\xi} \rangle &= \left\langle \text{Ad}_{O^{-1}}^* \left( \frac{d}{dt} \Pi - \text{ad}_{\widehat{\Omega}}^* \Pi \right), \widehat{\xi} \right\rangle. \end{aligned} \quad (2.4.23)$$

The proof of Noether's theorem for the rigid body is already on the second line. However, the last line gives a general result.  $\blacksquare$

## 54 2 : NEWTON, LAGRANGE, HAMILTON

**Remark 2.4.6** The proof of Noether's theorem for the rigid body when the constrained Hamilton–Pontryagin action is invariant under spatial rotations also proves a general result in Equation (2.4.23), with  $\widehat{\Omega} = O^{-1}\dot{O}$  for a Lie group  $O$ , that

$$\frac{d}{dt} \left( \text{Ad}_{O^{-1}}^* \Pi \right) = \text{Ad}_{O^{-1}}^* \left( \frac{d}{dt} \Pi - \text{ad}_{\widehat{\Omega}}^* \Pi \right). \quad (2.4.24)$$

This equation will be useful in the remainder of the text. In particular, it provides the solution of a differential equation defined on the dual of a Lie algebra. Namely, for a Lie group  $O$  with Lie algebra  $\mathfrak{o}$ , the equation for  $\Pi \in \mathfrak{o}^*$  and  $\widehat{\Omega} = O^{-1}\dot{O} \in \mathfrak{o}$

$$\frac{d}{dt} \Pi - \text{ad}_{\widehat{\Omega}}^* \Pi = 0 \quad \text{has solution} \quad \Pi(t) = \text{Ad}_{O(t)}^* \pi, \quad (2.4.25)$$

in which the constant  $\pi \in \mathfrak{o}^*$  is obtained from the initial conditions.  $\square$

## 2.4.2 Manakov's formulation of the $SO(n)$ rigid body

**Proposition 2.4.4 (Manakov [Man1976])** *Euler's equations for a rigid body on  $SO(n)$  take the matrix commutator form,*

$$\frac{dM}{dt} = [M, \Omega] \quad \text{with} \quad M = \mathbb{A}\Omega + \Omega\mathbb{A}, \quad (2.4.26)$$

where the  $n \times n$  matrices  $M, \Omega$  are skew-symmetric (forgoing superfluous hats) and  $\mathbb{A}$  is symmetric.

**Proof.** Manakov's commutator form of the  $SO(n)$  rigid-body Equations (2.4.26) follows as the Euler–Lagrange equations for Hamilton's principle  $\delta S = 0$  with  $S = \int l dt$  for the Lagrangian

$$l = -\frac{1}{2} \text{tr}(\Omega\mathbb{A}\Omega),$$

where  $\Omega = O^{-1}\dot{O} \in \mathfrak{so}(n)$  and the  $n \times n$  matrix  $\mathbb{A}$  is symmetric. Taking matrix variations in Hamilton's principle yields

$$\delta S = -\frac{1}{2} \int_a^b \text{tr}(\delta\Omega (\mathbb{A}\Omega + \Omega\mathbb{A})) dt = -\frac{1}{2} \int_a^b \text{tr}(\delta\Omega M) dt,$$

after cyclically permuting the order of matrix multiplication under the trace and substituting  $M := \mathbb{A}\Omega + \Omega\mathbb{A}$ . Using the variational formula (2.4.16) for  $\delta\Omega$  now leads to

$$\delta S = -\frac{1}{2} \int_a^b \text{tr}((\Xi' + \Omega\Xi - \Xi\Omega)M) dt.$$

Integrating by parts and permuting under the trace then yields the equation

$$\delta S = \frac{1}{2} \int_a^b \text{tr}(\Xi(\dot{M} + \Omega M - M\Omega)) dt.$$

Finally, invoking stationarity for arbitrary  $\Xi$  implies the commutator form (2.4.26). ■

### 2.4.3 Matrix Euler–Poincaré equations

Manakov’s commutator form of the rigid-body equations recalls much earlier work by Poincaré [Po1901], who also noticed that the matrix commutator form of Euler’s rigid-body equations suggests an additional mathematical structure going back to Lie’s theory of groups of transformations depending continuously on parameters. In particular, Poincaré [Po1901] remarked that the commutator form of Euler’s rigid-body equations would make sense for any Lie algebra, not just for  $so(3)$ . The proof of Manakov’s commutator form (2.4.26) by Hamilton’s principle is essentially the same as Poincaré’s proof in [Po1901], which is translated into English in Appendix D.

**Theorem 2.4.3 (Matrix Euler–Poincaré equations)** *The Euler–Lagrange equations for Hamilton’s principle  $\delta S = 0$  with  $S = \int l(\Omega) dt$  may be expressed in matrix commutator form,*

$$\frac{dM}{dt} = [M, \Omega] \quad \text{with} \quad M = \frac{\delta l}{\delta \Omega}, \quad (2.4.27)$$

*for any Lagrangian  $l(\Omega)$ , where  $\Omega = g^{-1}\dot{g} \in \mathfrak{g}$  and  $\mathfrak{g}$  is the matrix Lie algebra of any matrix Lie group  $G$ .*

**Proof.** The proof here is the same as the proof of Manakov's commutator formula via Hamilton's principle, modulo replacing  $O^{-1}\dot{O} \in so(n)$  with  $g^{-1}\dot{g} \in \mathfrak{g}$ . ■

**Remark 2.4.7** Poincaré's observation leading to the matrix Euler–Poincaré Equation (2.4.27) was reported in two pages with no references [Po1901]. The proof above shows that the matrix Euler–Poincaré equations possess a natural variational principle. Note that if  $\Omega = g^{-1}\dot{g} \in \mathfrak{g}$ , then  $M = \delta l / \delta \Omega \in \mathfrak{g}^*$ , where the dual is defined in terms of the matrix trace pairing. □

**Exercise.** Retrace the proof of the variational principle for the Euler–Poincaré equation, replacing the left-invariant quantity  $g^{-1}\dot{g}$  with the right-invariant quantity  $\dot{g}g^{-1}$ . ★

#### 2.4.4 An isospectral eigenvalue problem for the $SO(n)$ rigid body

The solution of the  $SO(n)$  rigid-body dynamics

$$\frac{dM}{dt} = [M, \Omega] \quad \text{with} \quad M = \mathbb{A}\Omega + \Omega\mathbb{A},$$

for the evolution of the  $n \times n$  skew-symmetric matrices  $M, \Omega$ , with constant symmetric  $\mathbb{A}$ , is given by a similarity transformation (later to be identified as coadjoint motion),

$$M(t) = O(t)^{-1}M(0)O(t) =: \text{Ad}_{O(t)}^*M(0),$$

with  $O(t) \in SO(n)$  and  $\Omega := O^{-1}\dot{O}(t)$ . Consequently, the evolution of  $M(t)$  is *isospectral*. This means that

- The initial eigenvalues of the matrix  $M(0)$  are preserved by the motion; that is,  $d\lambda/dt = 0$  in

$$M(t)\psi(t) = \lambda\psi(t),$$

provided its eigenvectors  $\psi \in \mathbb{R}^n$  evolve according to

$$\psi(t) = O(t)^{-1}\psi(0).$$

The proof of this statement follows from the corresponding property of similarity transformations.

- Its matrix invariants are preserved:

$$\frac{d}{dt}\text{tr}(M - \lambda\text{Id})^K = 0,$$

for every non-negative integer power  $K$ .

This is clear because the invariants of the matrix  $M$  may be expressed in terms of its eigenvalues; but these are invariant under a similarity transformation.

**Proposition 2.4.5** *Isospectrality allows the quadratic rigid-body dynamics (2.4.26) on  $SO(n)$  to be rephrased as a system of two coupled linear equations: the eigenvalue problem for  $M$  and an evolution equation for its eigenvectors  $\psi$ , as follows:*

$$M\psi = \lambda\psi \quad \text{and} \quad \dot{\psi} = -\Omega\psi, \quad \text{with} \quad \Omega = O^{-1}\dot{O}(t).$$

**Proof.** Applying isospectrality in the time derivative of the first equation yields

$$(\dot{M} + [\Omega, M])\psi + (M - \lambda\text{Id})(\dot{\psi} + \Omega\psi) = 0.$$

Now substitute the second equation to recover (2.4.26). ■

### 2.4.5 Manakov's integration of the $SO(n)$ rigid body

Manakov [Man1976] observed that Equations (2.4.26) may be “deformed” into

$$\frac{d}{dt}(M + \lambda A) = [(M + \lambda A), (\Omega + \lambda B)], \quad (2.4.28)$$

where  $A, B$  are also  $n \times n$  matrices and  $\lambda$  is a scalar constant parameter. For these deformed rigid-body equations on  $SO(n)$  to hold for any value of  $\lambda$ , the coefficient of each power must vanish.

- The coefficient of  $\lambda^2$  is

$$0 = [A, B].$$

Therefore,  $A$  and  $B$  must commute. For this, let them be constant and diagonal:

$$A_{ij} = \text{diag}(a_i)\delta_{ij}, \quad B_{ij} = \text{diag}(b_i)\delta_{ij} \quad (\text{no sum}).$$

- The coefficient of  $\lambda$  is

$$0 = \frac{dA}{dt} = [A, \Omega] + [M, B].$$

Therefore, by antisymmetry of  $M$  and  $\Omega$ ,

$$(a_i - a_j)\Omega_{ij} = (b_i - b_j)M_{ij},$$

which implies that

$$\Omega_{ij} = \frac{b_i - b_j}{a_i - a_j} M_{ij} \quad (\text{no sum}).$$

Hence, angular velocity  $\Omega$  is a linear function of angular momentum,  $M$ .

- Finally, the coefficient of  $\lambda^0$  recovers the Euler equation

$$\frac{dM}{dt} = [M, \Omega],$$

## 2.4 LAGRANGIAN FORM OF RIGID-BODY MOTION 59

but now with the restriction that the moments of inertia are of the form

$$\Omega_{ij} = \frac{b_i - b_j}{a_i - a_j} M_{ij} \quad (\text{no sum}).$$

This relation turns out to possess only five free parameters for  $n = 4$ .

Under these conditions, Manakov's deformation of the  $SO(n)$  rigid-body equation into the commutator form (2.4.28) implies for every non-negative integer power  $K$  that

$$\frac{d}{dt}(M + \lambda A)^K = [(M + \lambda A)^K, (\Omega + \lambda B)].$$

Since the commutator is antisymmetric, its trace vanishes and  $K$  conservation laws emerge, as

$$\frac{d}{dt} \text{tr}(M + \lambda A)^K = 0,$$

after commuting the trace operation with the time derivative. Consequently,

$$\text{tr}(M + \lambda A)^K = \text{constant},$$

for each power of  $\lambda$ . That is, all the coefficients of each power of  $\lambda$  are constant in time for the  $SO(n)$  rigid body. Manakov [Man1976] proved that these constants of motion are sufficient to completely determine the solution for  $n = 4$ .

**Remark 2.4.8** This result generalises considerably. For example, Manakov's method determines the solution for all the algebraically solvable rigid bodies on  $SO(n)$ . The moments of inertia of these bodies possess only  $2n - 3$  parameters. (Recall that in Manakov's case for  $SO(4)$  the moment of inertia possesses only five parameters.)  $\square$

## 60 2 : NEWTON, LAGRANGE, HAMILTON

**Exercise.** Try computing the constants of motion  $\text{tr}(M + \lambda A)^K$  for the values  $K = 2, 3, 4$ .

Hint: Keep in mind that  $M$  is a skew-symmetric matrix,  $M^T = -M$ , so the trace of the product of any diagonal matrix times an odd power of  $M$  vanishes. ★

**Answer.** The traces of the powers  $\text{trace}(M + \lambda A)^n$  are given by

$$\boxed{n = 2} : \text{tr } M^2 + 2\lambda \text{tr}(AM) + \lambda^2 \text{tr } A^2,$$

$$\boxed{n = 3} : \text{tr } M^3 + 3\lambda \text{tr}(AM^2) + 3\lambda^2 \text{tr } A^2 M + \lambda^3 \text{tr } A^3,$$

$$\begin{aligned} \boxed{n = 4} : \text{tr } M^4 + 4\lambda \text{tr}(AM^3) \\ + \lambda^2(2\text{tr } A^2 M^2 + 4\text{tr } AMAM) \\ + \lambda^3 \text{tr } A^3 M + \lambda^4 \text{tr } A^4. \end{aligned}$$

The number of conserved quantities for  $n = 2, 3, 4$  are, respectively, one ( $C_2 = \text{tr } M^2$ ), one ( $C_3 = \text{tr } AM^2$ ) and two ( $C_4 = \text{tr } M^4$  and  $I_4 = 2\text{tr } A^2 M^2 + 4\text{tr } AMAM$ ). ▲

**Exercise.** How do the Euler equations look on  $so(4)^*$  as a matrix equation? Is there an analogue of the hat map for  $so(4)$ ?

Hint: The Lie algebra  $so(4)$  is locally isomorphic to  $so(3) \times so(3)$ . ★

## 2.5 Hamilton

The *Legendre transform* of the Lagrangian (2.4.4) in the variational principle (2.4.3) for Euler's rigid-body dynamics (2.4.8) on  $\mathbb{R}^3$  will reveal its well-known Hamiltonian formulation.



**Definition 2.5.1 (Legendre transformation)** *The Legendre transformation  $\mathbb{F}l : \mathbb{R}^3 \rightarrow \mathbb{R}^{3*} \simeq \mathbb{R}^3$  is defined by the fibre derivative,*

$$\mathbb{F}l(\Omega) = \frac{\delta l}{\delta \Omega} = \Pi.$$

The Legendre transformation defines the *body angular momentum* by the variations of the rigid body's reduced Lagrangian with respect to the body angular velocity. For the Lagrangian in (2.4.3), the  $\mathbb{R}^3$  components of the body angular momentum are

$$\Pi_i = I_i \Omega_i = \frac{\partial l}{\partial \Omega_i}, \quad i = 1, 2, 3. \quad (2.5.1)$$

**Remark 2.5.1** This is also how body angular momentum was defined in the Newtonian setting. See Definition 2.1.10.  $\square$

**Exercise.** Express the Lagrangian (2.4.4) in terms of the matrices  $O(t)$  and  $\dot{O}(t)$ . Show that this Lagrangian is left-invariant under  $(O, \dot{O}) \mapsto (RO, R\dot{O})$  for any orthogonal matrix  $R^T = R^{-1}$ . Compute the Euler–Lagrange equations for this Lagrangian in geodesic form (2.2.4).  $\star$

**Exercise.** Compute the Legendre transformation and pass to the canonical Hamiltonian formulation using the Lagrangian  $l(\Omega) = L(O, \dot{O})$  and the following definitions of the canonical momentum and Hamiltonian,

$$P = \frac{\partial L(O, \dot{O})}{\partial \dot{O}} \quad \text{and} \quad H(P, O) = \langle P, \dot{O} \rangle - L(O, \dot{O}),$$

in combination with the chain rule for  $\Omega = O^{-1}\dot{O}$ .  $\star$

## 62 2 : NEWTON, LAGRANGE, HAMILTON

## 2.5.1 Hamiltonian form of rigid-body motion

**Definition 2.5.2 (Dynamical systems in Hamiltonian form)** *A dynamical system on a manifold  $M$*

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}), \quad \mathbf{x} \in M,$$

is said to be in **Hamiltonian form**, if it can be expressed as

$$\dot{\mathbf{x}}(t) = \{\mathbf{x}, H\}, \quad \text{for } H : M \mapsto \mathbb{R},$$

in terms of a Poisson bracket operation  $\{\cdot, \cdot\}$  among smooth real functions  $\mathcal{F}(M) : M \mapsto \mathbb{R}$  on the manifold  $M$ ,

$$\{\cdot, \cdot\} : \mathcal{F}(M) \times \mathcal{F}(M) \mapsto \mathcal{F}(M),$$

so that  $\dot{F} = \{F, H\}$  for any  $F \in \mathcal{F}(M)$ .

**Definition 2.5.3 (Poisson bracket)** *A Poisson bracket operation  $\{\cdot, \cdot\}$  is defined as possessing the following properties:*

- It is  **bilinear**.
- It is  **skew-symmetric**,  $\{F, H\} = -\{H, F\}$ .
- It satisfies the  **Leibniz rule** (product rule),

$$\{FG, H\} = \{F, H\}G + F\{G, H\},$$

for the product of any two functions  $F$  and  $G$  on  $M$ .

- It satisfies the  **Jacobi identity**,

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0, \quad (2.5.2)$$

for any three functions  $F, G$  and  $H$  on  $M$ .

**Remark 2.5.2** This definition of a Poisson bracket does not require it to be the standard canonical bracket in position  $q$  and conjugate momentum  $p$ , although it does include that case as well.  $\square$

## 2.5.2 Lie–Poisson Hamiltonian rigid-body dynamics

Let

$$h(\mathbf{\Pi}) := \langle \mathbf{\Pi}, \mathbf{\Omega} \rangle - l(\mathbf{\Omega}), \quad (2.5.3)$$

where the pairing  $\langle \cdot, \cdot \rangle : \mathbb{R}^{3*} \times \mathbb{R}^3 \rightarrow \mathbb{R}$  denotes the vector dot product on  $\mathbb{R}^3$ ,

$$\langle \mathbf{\Pi}, \mathbf{\Omega} \rangle := \mathbf{\Pi} \cdot \mathbf{\Omega}, \quad (2.5.4)$$

in which indices are suppressed within the brackets. Hence, one finds the expected expression for the rigid-body Hamiltonian

$$h = \frac{1}{2} \langle \mathbf{\Pi}, \mathbb{I}^{-1} \mathbf{\Pi} \rangle = \frac{1}{2} \mathbf{\Pi} \cdot \mathbb{I}^{-1} \mathbf{\Pi} := \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3}. \quad (2.5.5)$$

The Legendre transform  $\mathbb{F}l$  for this case is a diffeomorphism, so one may solve for

$$\frac{\partial h}{\partial \mathbf{\Pi}} = \mathbf{\Omega} + \left\langle \mathbf{\Pi}, \frac{\partial \mathbf{\Omega}}{\partial \mathbf{\Pi}} \right\rangle - \left\langle \frac{\partial l}{\partial \mathbf{\Omega}}, \frac{\partial \mathbf{\Omega}}{\partial \mathbf{\Pi}} \right\rangle = \mathbf{\Omega} \quad (2.5.6)$$

upon using the definition of angular momentum  $\mathbf{\Pi} = \partial l / \partial \mathbf{\Omega}$  in (2.5.1). In  $\mathbb{R}^3$  coordinates, the relation (2.5.6) expresses the body angular velocity as the derivative of the reduced Hamiltonian with respect to the body angular momentum, namely,

$$\frac{\partial h}{\partial \mathbf{\Pi}} = \mathbf{\Omega}. \quad (2.5.7)$$

Hence, the reduced Euler–Lagrange equations for  $l$  may be expressed equivalently in angular momentum vector components in  $\mathbb{R}^3$  and Hamiltonian  $h$  as

$$\frac{d}{dt}(\mathbb{I}\mathbf{\Omega}) = \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega} \iff \dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \frac{\partial h}{\partial \mathbf{\Pi}} := \{\mathbf{\Pi}, h\}.$$

This expression suggests we introduce the following *rigid-body Poisson bracket* on functions of the  $\mathbf{\Pi}$ 's:

$$\{f, h\}(\mathbf{\Pi}) := -\mathbf{\Pi} \cdot \left( \frac{\partial f}{\partial \mathbf{\Pi}} \times \frac{\partial h}{\partial \mathbf{\Pi}} \right). \quad (2.5.8)$$

## 64 2 : NEWTON, LAGRANGE, HAMILTON

For the Hamiltonian (2.5.5), one checks that the Euler equations in terms of the rigid-body angular momenta,

$$\begin{aligned}\dot{\Pi}_1 &= \left( \frac{1}{I_3} - \frac{1}{I_2} \right) \Pi_2 \Pi_3, \\ \dot{\Pi}_2 &= \left( \frac{1}{I_1} - \frac{1}{I_3} \right) \Pi_3 \Pi_1, \\ \dot{\Pi}_3 &= \left( \frac{1}{I_2} - \frac{1}{I_1} \right) \Pi_1 \Pi_2,\end{aligned}\tag{2.5.9}$$

that is, the equations

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbb{I}^{-1} \mathbf{\Pi},\tag{2.5.10}$$

are equivalent to

$$\dot{f} = \{f, h\}, \quad \text{with } f = \mathbf{\Pi}.$$

### 2.5.3 Lie–Poisson bracket

The Poisson bracket proposed in (2.5.8) is an example of a *Lie–Poisson bracket*.

It satisfies the defining relations of a Poisson bracket for a number of reasons, not least because it is the hat map to  $\mathbb{R}^3$  of the following bracket defined by the general form in Equation (2.4.23) in terms of the  $so(3)^* \times so(3)$  pairing  $\langle \cdot, \cdot \rangle$  in Equation (2.4.20). Namely,

$$\begin{aligned}\frac{dF}{dt} &= \left\langle \frac{d}{dt} \mathbf{\Pi}, \frac{\partial F}{\partial \mathbf{\Pi}} \right\rangle = \left\langle \text{ad}_{\hat{\Omega}}^* \mathbf{\Pi}, \frac{\partial F}{\partial \mathbf{\Pi}} \right\rangle \\ &= \left\langle \mathbf{\Pi}, \text{ad}_{\hat{\Omega}} \frac{\partial F}{\partial \mathbf{\Pi}} \right\rangle = \left\langle \mathbf{\Pi}, \left[ \hat{\Omega}, \frac{\partial F}{\partial \mathbf{\Pi}} \right] \right\rangle \\ &= - \left\langle \mathbf{\Pi}, \left[ \frac{\partial F}{\partial \mathbf{\Pi}}, \frac{\partial H}{\partial \mathbf{\Pi}} \right] \right\rangle,\end{aligned}\tag{2.5.11}$$

where we have used the equation corresponding to (2.5.7) under the inverse of the hat map

$$\hat{\Omega} = \frac{\partial H}{\partial \mathbf{\Pi}}$$

and applied antisymmetry of the matrix commutator. Writing Equation (2.5.11) as

$$\frac{dF}{dt} = - \left\langle \Pi, \left[ \frac{\partial F}{\partial \Pi}, \frac{\partial H}{\partial \Pi} \right] \right\rangle =: \{F, H\} \quad (2.5.12)$$

defines the Lie–Poisson bracket  $\{\cdot, \cdot\}$  on smooth functions  $(F, H) : so(3)^* \rightarrow \mathbb{R}$ . This bracket satisfies the defining relations of a Poisson bracket because it is a linear functional of the commutator product of skew-symmetric matrices, which is bilinear, skew-symmetric, satisfies the Leibniz rule (because of the partial derivatives) and also satisfies the Jacobi identity.

These Lie–Poisson brackets may be written in tabular form as

$$\{\Pi_i, \Pi_j\} = \begin{array}{c|ccc} \{\cdot, \cdot\} & \Pi_1 & \Pi_2 & \Pi_3 \\ \hline \Pi_1 & 0 & -\Pi_3 & \Pi_2 \\ \Pi_2 & \Pi_3 & 0 & -\Pi_1 \\ \Pi_3 & -\Pi_2 & \Pi_1 & 0 \end{array} \quad (2.5.13)$$

or, in index notation,

$$\{\Pi_i, \Pi_j\} = -\epsilon_{ijk}\Pi_k = \hat{\Pi}_{ij}. \quad (2.5.14)$$

**Remark 2.5.3** The Lie–Poisson bracket in the form (2.5.12) would apply to any Lie algebra. This Lie–Poisson Hamiltonian form of the rigid-body dynamics substantiates Poincaré’s observation in [Po1901] that the corresponding equations could have been written on the dual of any Lie algebra by using the  $\text{ad}^*$  operation for that Lie algebra.  $\square$

The corresponding Poisson bracket in (2.5.8) in  $\mathbb{R}^3$ -vector form also satisfies the defining relations of a Poisson bracket because it is an example of a *Nambu bracket*, to be discussed next.

#### 2.5.4 Nambu’s $\mathbb{R}^3$ Poisson bracket

The rigid-body Poisson bracket (2.5.8) is a special case of the Poisson bracket for functions of  $\mathbf{x} \in \mathbb{R}^3$  introduced in [Na1973],

$$\{f, h\} = -\nabla c \cdot \nabla f \times \nabla h. \quad (2.5.15)$$

## 66 2 : NEWTON, LAGRANGE, HAMILTON

This bracket generates the motion

$$\dot{\mathbf{x}} = \{\mathbf{x}, h\} = \nabla c \times \nabla h. \quad (2.5.16)$$

For this bracket the motion takes place along the intersections of level surfaces of the functions  $c$  and  $h$  in  $\mathbb{R}^3$ . In particular, for the rigid body, the motion takes place along intersections of angular momentum spheres  $c = |\mathbf{x}|^2/2$  and energy ellipsoids  $h = \mathbf{x} \cdot \mathbb{I}\mathbf{x}$ . (See the cover illustration of [MaRa1994].)

**Exercise.** Consider the Nambu  $\mathbb{R}^3$  bracket

$$\{f, h\} = -\nabla c \cdot \nabla f \times \nabla h. \quad (2.5.17)$$

Let  $c = \mathbf{x}^T \cdot \mathbb{C}\mathbf{x}/2$  be a quadratic form on  $\mathbb{R}^3$ , and let  $\mathbb{C}$  be the associated symmetric  $3 \times 3$  matrix. Show by direct computation that this Nambu bracket satisfies the Jacobi identity. ★

**Exercise.** Find the general conditions on the function  $c(\mathbf{x})$  so that the  $\mathbb{R}^3$  bracket

$$\{f, h\} = -\nabla c \cdot \nabla f \times \nabla h$$

satisfies the defining properties of a Poisson bracket. Is this  $\mathbb{R}^3$  bracket also a derivation satisfying the Leibniz relation for a product of functions on  $\mathbb{R}^3$ ? If so, why? ★

**Answer.** The bilinear skew-symmetric Nambu  $\mathbb{R}^3$  bracket yields the divergenceless vector field

$$X_{c,h} = \{\cdot, h\} = (\nabla c \times \nabla h) \cdot \nabla \quad \text{with} \quad \text{div}(\nabla c \times \nabla h) = 0.$$

Divergenceless vector fields are derivative operators that satisfy the Leibniz product rule. They also satisfy the Jacobi identity for any choice of  $C^2$  functions  $c$  and  $h$ . Hence, the Nambu  $\mathbb{R}^3$  bracket is a bilinear skew-symmetric operation satisfying the defining properties of a Poisson bracket. ▲

**Theorem 2.5.1 (Jacobi identity)** *The Nambu  $\mathbb{R}^3$  bracket (2.5.17) satisfies the Jacobi identity.*

**Proof.** The isomorphism  $X_H = \{ \cdot, H \}$  between the Lie algebra of divergenceless vector fields and functions under the  $\mathbb{R}^3$  bracket is the key to proving this theorem. The Lie derivative among vector fields is identified with the Nambu bracket by

$$\mathcal{L}_{X_G} X_H = [X_G, X_H] = -X_{\{G,H\}}.$$

Repeating the Lie derivative produces

$$\mathcal{L}_{X_F}(\mathcal{L}_{X_G} X_H) = [X_F, [X_G, X_H]] = X_{\{F,\{G,H\}\}}.$$

The result follows because both the left- and right-hand sides in this equation satisfy the Jacobi identity. ■

**Exercise.** How is the  $\mathbb{R}^3$  bracket related to the canonical Poisson bracket?

Hint: Restrict to level surfaces of the function  $c(\mathbf{x})$ . ★

**Exercise. (Casimirs of the  $\mathbb{R}^3$  bracket)** The Casimirs (or distinguished functions, as Lie called them) of a Poisson bracket satisfy

$$\{c, h\}(\mathbf{x}) = 0, \quad \text{for all } h(\mathbf{x}).$$

Suppose the function  $c(\mathbf{x})$  is chosen so that the  $\mathbb{R}^3$  bracket (2.5.15) defines a proper Poisson bracket. What are the Casimirs for the  $\mathbb{R}^3$  bracket (2.5.15)? Why? ★

**Exercise. (Geometric interpretation of Nambu motion)**

- Show that the Nambu motion equation (2.5.16)

$$\dot{\mathbf{x}} = \{\mathbf{x}, h\} = \nabla c \times \nabla h$$

for the  $\mathbb{R}^3$  bracket (2.5.15) is invariant under a certain linear combination of the functions  $c$  and  $h$ . Interpret this invariance geometrically.

- Show that the rigid-body equations for

$$\mathbb{I} = \text{diag}(1, 1/2, 1/3)$$

may be interpreted as intersections in  $\mathbb{R}^3$  of the spheres  $x_1^2 + x_2^2 + x_3^2 = \text{constant}$  and the hyperbolic cylinders  $x_1^2 - x_3^2 = \text{constant}$ . See [HoMa1991] for more discussions of this geometric interpretation of solutions under the  $\mathbb{R}^3$  bracket.

- A special case of the equations for three-wave interactions is [AlluMaRo1998]

$$\dot{x}_1 = s_1 \gamma_1 x_2 x_3, \quad \dot{x}_2 = s_2 \gamma_2 x_3 x_1, \quad \dot{x}_3 = s_3 \gamma_3 x_1 x_2,$$

for a set of constants  $\gamma_1 + \gamma_2 + \gamma_3 = 0$  and signs  $s_1, s_2, s_3 = \pm 1$ . Write these equations as a Nambu motion equation on  $\mathbb{R}^3$  of the form (2.5.16). Interpret their solutions geometrically as intersections of level surfaces of quadratic functions for various values and signs of the  $\gamma$ 's. ★



### 2.5.5 Clebsch variational principle for the rigid body

**Proposition 2.5.1 (Clebsch variational principle)** *The Euler rigid-body Equations (2.4.2) on  $T\mathbb{R}^3$  are equivalent to the **constrained variational principle**,*

$$\delta S(\Omega, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P}) = \delta \int_a^b l(\Omega, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P}) dt = 0, \quad (2.5.18)$$

for a *constrained action integral*

$$\begin{aligned} S(\Omega, \mathbf{Q}, \dot{\mathbf{Q}}) &= \int_a^b l(\Omega, \mathbf{Q}, \dot{\mathbf{Q}}) dt \\ &= \int_a^b \frac{1}{2} \Omega \cdot \mathbb{I} \Omega + \mathbf{P} \cdot (\dot{\mathbf{Q}} + \Omega \times \mathbf{Q}) dt. \end{aligned} \quad (2.5.19)$$

**Remark 2.5.4 (Reconstruction as constraint)**

- The first term in the Lagrangian (2.5.19),

$$l(\Omega) = \frac{1}{2} (I_1 \Omega_1^2 + I_2 \Omega_2^2 + I_3 \Omega_3^2) = \frac{1}{2} \Omega^T \mathbb{I} \Omega, \quad (2.5.20)$$

is again the (rotational) kinetic energy of the rigid body in (2.1.23).

- The second term in the Lagrangian (2.5.19) introduces the Lagrange multiplier  $\mathbf{P}$  which imposes the constraint

$$\dot{\mathbf{Q}} + \Omega \times \mathbf{Q} = 0.$$

This *reconstruction formula* has the solution

$$\mathbf{Q}(t) = O^{-1}(t) \mathbf{Q}(0),$$

which satisfies

$$\begin{aligned} \dot{\mathbf{Q}}(t) &= -(O^{-1} \dot{O}) O^{-1}(t) \mathbf{Q}(0) \\ &= -\hat{\Omega}(t) \mathbf{Q}(t) = -\Omega(t) \times \mathbf{Q}(t). \end{aligned} \quad (2.5.21)$$

□

## 70 2 : NEWTON, LAGRANGE, HAMILTON

**Proof.** The variations of  $S$  are given by

$$\begin{aligned}\delta S &= \int_a^b \left( \frac{\delta l}{\delta \Omega} \cdot \delta \Omega + \frac{\delta l}{\delta \mathbf{P}} \cdot \delta \mathbf{P} + \frac{\delta l}{\delta \mathbf{Q}} \cdot \delta \mathbf{Q} \right) dt \\ &= \int_a^b \left[ \left( \mathbb{I} \Omega - \mathbf{P} \times \mathbf{Q} \right) \cdot \delta \Omega \right. \\ &\quad \left. + \delta \mathbf{P} \cdot \left( \dot{\mathbf{Q}} + \Omega \times \mathbf{Q} \right) - \delta \mathbf{Q} \cdot \left( \dot{\mathbf{P}} + \Omega \times \mathbf{P} \right) \right] dt.\end{aligned}$$

Thus, stationarity of this *implicit variational principle* implies the following set of equations:

$$\mathbb{I} \Omega = \mathbf{P} \times \mathbf{Q}, \quad \dot{\mathbf{Q}} = -\Omega \times \mathbf{Q}, \quad \dot{\mathbf{P}} = -\Omega \times \mathbf{P}. \quad (2.5.22)$$

These *symmetric equations* for the rigid body first appeared in the theory of optimal control of rigid bodies [BlCrMaRa1998]. Euler's form of the rigid-body equations emerges from these, upon elimination of  $\mathbf{Q}$  and  $\mathbf{P}$ , as

$$\begin{aligned}\mathbb{I} \dot{\Omega} &= \dot{\mathbf{P}} \times \mathbf{Q} + \mathbf{P} \times \dot{\mathbf{Q}} \\ &= \mathbf{Q} \times (\Omega \times \mathbf{P}) + \mathbf{P} \times (\mathbf{Q} \times \Omega) \\ &= -\Omega \times (\mathbf{P} \times \mathbf{Q}) = -\Omega \times \mathbb{I} \Omega,\end{aligned}$$

which are Euler's equations for the rigid body in  $T\mathbb{R}^3$ . ■

**Remark 2.5.5** The Clebsch variational principle for the rigid body is a natural approach in developing geometric algorithms for numerical integrations of rotating motion. Geometric integrators for rotations are derived using the Clebsch approach in [CoHo2007]. □

**Remark 2.5.6** The Clebsch approach is also a natural path across to the Hamiltonian formulation of the rigid-body equations. This becomes clear in the course of the following exercise. □

**Exercise.** Given that the canonical Poisson brackets in Hamilton's approach are

$$\{Q_i, P_j\} = \delta_{ij} \quad \text{and} \quad \{Q_i, Q_j\} = 0 = \{P_i, P_j\},$$

what are the Poisson brackets for  $\mathbf{\Pi} = \mathbf{P} \times \mathbf{Q} \in \mathbb{R}^3$  in (2.5.22)? Show that these Poisson brackets recover the rigid-body Poisson bracket (2.5.8). ★

**Answer.** The components of the angular momentum  $\mathbf{\Pi} = \mathbb{I}\mathbf{\Omega}$  in (2.5.22) are

$$\Pi_a = \epsilon_{abc} P_b Q_c,$$

and their canonical Poisson brackets are (noting the similarity with the hat map)

$$\{\Pi_a, \Pi_i\} = \{\epsilon_{abc} P_b Q_c, \epsilon_{ijk} P_j Q_k\} = -\epsilon_{ail} \Pi_l.$$

Consequently, the derivative property of the canonical Poisson bracket yields

$$\{f, h\}(\mathbf{\Pi}) = \frac{\partial f}{\partial \Pi_a} \{\Pi_a, \Pi_i\} \frac{\partial h}{\partial \Pi_b} = -\epsilon_{abc} \Pi_c \frac{\partial f}{\partial \Pi_a} \frac{\partial h}{\partial \Pi_b}, \quad (2.5.23)$$

which is indeed the Lie–Poisson bracket in (2.5.8) on functions of the  $\mathbf{\Pi}$ 's. The correspondence with the hat map noted above shows that this Poisson bracket satisfies the Jacobi identity as a result of the Jacobi identity for the vector cross product on  $\mathbb{R}^3$ . ▲

**Remark 2.5.7** This exercise proves that the map  $T^*\mathbb{R}^3 \rightarrow \mathbb{R}^3$  given by  $\mathbf{\Pi} = \mathbf{P} \times \mathbf{Q} \in \mathbb{R}^3$  in (2.5.22) is Poisson. That is, the map takes Poisson brackets on one manifold into Poisson brackets on another manifold. Later we will recognise such an occurrence as one of the properties of a *momentum map*. □

## 72 2 : NEWTON, LAGRANGE, HAMILTON

**Exercise.** The Euler–Lagrange equations in matrix commutator form of Manakov’s formulation of the rigid body on  $SO(n)$  are

$$\frac{dM}{dt} = [M, \Omega],$$

where the  $n \times n$  matrices  $M, \Omega$  are skew-symmetric. Show that these equations may be derived from Hamilton’s principle  $\delta S = 0$  with constrained action integral

$$S(\Omega, Q, P) = \int_a^b l(\Omega) + \text{tr}(P^T (\dot{Q} - Q\Omega)) dt,$$

for which  $M = \delta l / \delta \Omega = P^T Q - Q^T P$  and  $Q, P \in SO(n)$  satisfy the following symmetric equations reminiscent of those in (2.5.22),

$$\dot{Q} = Q\Omega \quad \text{and} \quad \dot{P} = P\Omega, \quad (2.5.24)$$

as a result of the constraints. ★

**Exercise.** Write Manakov’s deformation of the rigid-body Equations (2.4.28) in the symmetric form (2.5.24). ★

### 2.5.6 Rotating motion with potential energy

Manakov’s method for showing the integrability of the  $n$ -dimensional rigid body illustrates the conditions necessary to prove isospectral integrability for any Lie–Poisson system. For example, consider the problem of a rigid body in a quadratic potential, first studied in [Bo1985].

The Lagrangian of an arbitrary rigid body rotating about a fixed point at the origin of spatial coordinates  $x \in \mathbb{R}^n$  in a field with a quadratic potential

$$\phi(x) = \frac{1}{2} \text{tr}(x^T \mathbb{S}_0 x)$$

is defined in the body coordinates by the difference between its kinetic and potential energies in the form

$$l = \underbrace{\frac{1}{2} \text{tr}(\Omega^T \mathbb{A} \Omega)}_{\text{kinetic}} - \underbrace{\frac{1}{2} \text{tr}(\mathbb{S} \mathbb{A})}_{\text{potential}} .$$

Here,  $\Omega(t) = O^{-1}(t) \dot{O}(t) \in so(n)$ , the  $n \times n$  constant matrices  $\mathbb{A}$  and  $\mathbb{S}_0$  are symmetric, and  $\mathbb{S}(t) = O^{-1}(t) \mathbb{S}_0 O(t)$ .

The reduced Euler–Lagrange equations for this Lagrangian are computed by taking matrix variations in its Hamilton’s principle  $\delta S = 0$  with  $S = \int l dt$ , to find

$$\delta S = \frac{1}{2} \int_a^b \text{tr}(\delta \Omega M) dt + \frac{1}{2} \int_a^b \text{tr}(\Xi [\mathbb{S}, \mathbb{A}]) dt ,$$

with matrix commutator  $[\mathbb{S}, \mathbb{A}] := \mathbb{S} \mathbb{A} - \mathbb{A} \mathbb{S}$ , variation  $\Xi := O^{-1} \delta O \in so(n)$  and variational derivative  $M := \partial l / \partial \Omega = \mathbb{A} \Omega + \Omega \mathbb{A}$ .

Integrating by parts, invoking homogeneous endpoint conditions, then rearranging as in the proof of Proposition 2.4.2 and using the variational relation (2.4.16), rewritten here as

$$\delta \Omega = \frac{d\Xi}{dt} + [\Omega, \Xi] ,$$

finally yields the following formula for the variation,

$$\delta S = -\frac{1}{2} \int_a^b \text{tr} \left( \left( \frac{dM}{dt} - [M, \Omega] - [\mathbb{S}, \mathbb{A}] \right) \Xi \right) dt .$$

Hence, Hamilton’s principle for  $\delta S = 0$  with arbitrary  $\Xi$  implies an equation for the evolution of  $M$  given by

$$\frac{dM}{dt} = [M, \Omega] + [\mathbb{S}(t), \mathbb{A}] . \quad (2.5.25)$$

## 74 2 : NEWTON, LAGRANGE, HAMILTON

A differential equation for  $\mathbb{S}(t)$  follows from the time derivative of its definition  $\mathbb{S}(t) := O^{-1}(t)\mathbb{S}_0O(t)$ , as

$$\frac{d\mathbb{S}}{dt} = [\mathbb{S}, \Omega]. \quad (2.5.26)$$

The last two equations constitute a closed dynamical system for  $M(t)$  and  $\mathbb{S}(t)$ , with initial conditions specified by the values of  $\Omega(0)$  and  $\mathbb{S}(0) = \mathbb{S}_0$  for  $O(0) = \text{Id}$  at time  $t = 0$ .

Following Manakov's idea [Man1976], these equations may be combined into a commutator of polynomials,

$$\frac{d}{dt}(\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2) = [\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2, \Omega + \lambda \mathbb{A}]. \quad (2.5.27)$$

The commutator form (2.5.27) implies for every non-negative integer power  $K$  that

$$\frac{d}{dt}(\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2)^K = [(\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2)^K, (\Omega + \lambda \mathbb{A})].$$

Since the commutator is antisymmetric, its trace vanishes and  $K$  conservation laws emerge, as

$$\frac{d}{dt} \text{tr}(\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2)^K = 0,$$

after commuting the trace operation with the time derivative. Consequently,

$$\text{tr}(\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2)^K = \text{constant}, \quad (2.5.28)$$

for each power of  $\lambda$ . That is, all the coefficients of each power of  $\lambda$  are constant in time for the motion of a rigid body in a quadratic field.

**Exercise.** Show that the Hamiltonian formulation of this system is Lie–Poisson, with Hamiltonian function

$$H(M, \mathbb{S}) = \frac{1}{2} \text{tr}(\Omega^T M) + \frac{1}{2} \text{tr}(\mathbb{S}, \mathbb{A}).$$

Determine the Lie algebra involved. ★

**Exercise.** Explicitly compute the conservation laws in (2.5.28) for  $n = 4$ . ★

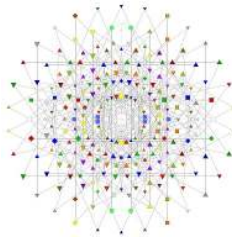
**Exercise.** What is the dimension of the generic solution of the system of equations (2.5.25) and (2.5.26)? That is, what is the sum of the dimensions of  $so(n)$  and the symmetric  $n \times n$  matrices, minus the number of conservation laws? ★

**Exercise.** Write the equations of motion and their Lie–Poisson Hamiltonian formulation in  $\mathbb{R}^3$ -vector form for the case when  $\Omega(t) = O^{-1}(t)\dot{O}(t) \in so(3)$  by using the hat map. List the conservation laws in this case. ★

**Exercise.** How would the variational calculation of the system (2.5.25) and (2.5.26) have changed if the Lie group had been unitary instead of orthogonal and the matrices  $\mathbb{S}_0$ ,  $\mathbb{A}$  and  $\mathbb{S}(t)$  were Hermitian, rather than symmetric? ★







# 3

## QUATERNIONS

---

### Contents

3.1	Operating with quaternions	78
3.1.1	Multiplying quaternions using Pauli matrices	79
3.1.2	Quaternionic conjugate	82
3.1.3	Decomposition of three-vectors	85
3.1.4	Alignment dynamics for Newton's second law	86
3.1.5	Quaternionic dynamics of Kepler's problem	90
3.2	Quaternionic conjugation	93
3.2.1	Cayley–Klein parameters	93
3.2.2	Pure quaternions, Pauli matrices and $SU(2)$	99
3.2.3	Tilde map: $\mathbb{R}^3 \simeq su(2) \simeq so(3)$	102
3.2.4	Dual of the tilde map: $\mathbb{R}^{3*} \simeq su(2)^* \simeq so(3)^*$	103
3.2.5	Pauli matrices and Poincaré's sphere $\mathbb{C}^2 \rightarrow S^2$	103
3.2.6	Poincaré's sphere and Hopf's fibration	105
3.2.7	Coquaternions	108

---

### 3.1 Operating with quaternions



William Rowan Hamilton

Hamilton had great hopes for the utility of quaternions, although he was not entirely sure how their utility would emerge. This was evidenced by his appeal to the following quotation of John Wallis [Wa1685] in the preface of Hamilton's book on quaternions:

We find therefore that in Equations, whether Lateral or Quadratick, which in the strict Sense, and first Prospect, appear Impossible; some mitigation may be allowed to make them Possible; and in such a mitigated interpretation they may yet be useful.

– Wallis, 1685

However, not all of his peers had such great hopes for quaternions and history treated them rather unkindly. Decades later, Lord Kelvin condemned them harshly [OcoRo1998]:

Quaternions came from Hamilton after his best work had been done, and though beautifully ingenious, they have been an unmixed evil to those who have touched them in any way. – Lord Kelvin (William Thomson), 1890

Hamilton's hope that quaternions "may yet be useful" was eventually redeemed by their broad modern applications. The relation between quaternions and vectors is now understood, as we shall explain, and quaternions are used for their special advantages in the robotics and avionics industries to track objects moving continuously along a curve of tumbling rotations. They are also heavily used in graphics [Ha2006, Ku1999].

Hamilton was correct: quaternions are special. For example, they form the only associative division ring containing both real and complex numbers. For us, they also form a natural introduction to

geometric mechanics. In particular, quaternions will introduce us to *mechanics on Lie groups*; namely, mechanics on the Lie group  $SU(2)$  of  $2 \times 2$  special unitary matrices.

### 3.1.1 Multiplying quaternions using Pauli matrices

Every quaternion  $q \in \mathbb{H}$  is a real linear combination of the *basis quaternions*, denoted as  $(\mathbb{J}_0, \mathbb{J}_1, \mathbb{J}_2, \mathbb{J}_3)$ . The *multiplication rules* for their basis are given by the triple product

$$\mathbb{J}_1\mathbb{J}_2\mathbb{J}_3 = -\mathbb{J}_0, \quad (3.1.1)$$

and the squares

$$\mathbb{J}_1^2 = \mathbb{J}_2^2 = \mathbb{J}_3^2 = -\mathbb{J}_0, \quad (3.1.2)$$

where  $\mathbb{J}_0$  is the identity element. Thus,  $\mathbb{J}_1\mathbb{J}_2 = \mathbb{J}_3$  holds, with cyclic permutations of  $(1, 2, 3)$ . According to a famous story, Hamilton inscribed a version of their triple product formula on Brougham (pronounced “Broom”) bridge in Dublin [OcoRo1998].

Quaternions combine a real scalar  $q \in \mathbb{R}$  and a real three-vector  $\mathbf{q} \in \mathbb{R}^3$  with components  $q_a$   $a = 1, 2, 3$ , into a *tetrad*

$$q = [q_0, \mathbf{q}] = q_0\mathbb{J}_0 + q_1\mathbb{J}_1 + q_2\mathbb{J}_2 + q_3\mathbb{J}_3 \in \mathbb{H}. \quad (3.1.3)$$

The multiplication table of the quaternion basis elements may be expressed as

	$\mathbb{J}_0$	$\mathbb{J}_1$	$\mathbb{J}_2$	$\mathbb{J}_3$	
$\mathbb{J}_0$	$\mathbb{J}_0$	$\mathbb{J}_1$	$\mathbb{J}_2$	$\mathbb{J}_3$	. (3.1.4)
$\mathbb{J}_1$	$\mathbb{J}_1$	$-\mathbb{J}_0$	$\mathbb{J}_3$	$-\mathbb{J}_2$	
$\mathbb{J}_2$	$\mathbb{J}_2$	$-\mathbb{J}_3$	$-\mathbb{J}_0$	$\mathbb{J}_1$	
$\mathbb{J}_3$	$\mathbb{J}_3$	$\mathbb{J}_2$	$-\mathbb{J}_1$	$-\mathbb{J}_0$	

**Definition 3.1.1 (Multiplication of quaternions)** *The multiplication rule for two quaternions,*

$$q = [q_0, \mathbf{q}] \quad \text{and} \quad r = [r_0, \mathbf{r}] \in \mathbb{H},$$

## 80 3 : QUATERNIONS

may be defined in vector notation as

$$\mathbf{qr} = [q_0, \mathbf{q}][r_0, \mathbf{r}] = [q_0r_0 - \mathbf{q} \cdot \mathbf{r}, q_0\mathbf{r} + r_0\mathbf{q} + \mathbf{q} \times \mathbf{r}]. \quad (3.1.5)$$

**Remark 3.1.1** The antisymmetric and symmetric parts of the quaternionic product correspond to *vector operations*<sup>1</sup>:

$$\frac{1}{2}(\mathbf{qr} - \mathbf{rq}) = [0, \mathbf{q} \times \mathbf{r}], \quad (3.1.6)$$

$$\frac{1}{2}(\mathbf{qr} + \mathbf{rq}) = [q_0r_0 - \mathbf{q} \cdot \mathbf{r}, q_0\mathbf{r} + r_0\mathbf{q}]. \quad (3.1.7)$$

The product of quaternions is not commutative. (It has a nonzero antisymmetric part.)  $\square$

**Theorem 3.1.1 (Isomorphism with Pauli matrix product)** *The multiplication rule (3.1.5) may be represented in a  $2 \times 2$  matrix basis as*

$$\mathbf{q} = [q_0, \mathbf{q}] = q_0\sigma_0 - i\mathbf{q} \cdot \boldsymbol{\sigma}, \quad \text{with } \mathbf{q} \cdot \boldsymbol{\sigma} := \sum_{a=1}^3 q_a\sigma_a, \quad (3.1.8)$$

where  $\sigma_0$  is the  $2 \times 2$  identity matrix and  $\sigma_a$ , with  $a = 1, 2, 3$ , are the Hermitian **Pauli spin matrices**,

$$\begin{aligned} \sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \sigma_1 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \sigma_2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, & \sigma_3 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned} \quad (3.1.9)$$

**Proof.** The isomorphism is implied by the product relation for the Pauli matrices

$$\sigma_a\sigma_b = \delta_{ab}\sigma_0 + i\epsilon_{abc}\sigma_c \quad \text{for } a, b, c = 1, 2, 3, \quad (3.1.10)$$

<sup>1</sup>Hamilton introduced the word *vector* in 1846 as a synonym for a *pure quaternion*, whose scalar part vanishes.

## 3.1 OPERATING WITH QUATERNIONS 81

where  $\epsilon_{abc}$  is the totally antisymmetric tensor density with  $\epsilon_{123} = 1$ . The Pauli matrices also satisfy  $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_0$  and one has  $\sigma_1\sigma_2\sigma_3 = i\sigma_0$  as well as cyclic permutations of  $\{1, 2, 3\}$ . Identifying  $\mathbb{J}_0 = \sigma_0$  and  $\mathbb{J}_a = -i\sigma_a$ , with  $a = 1, 2, 3$ , provides the basic quaternionic properties. ■

**Exercise.** Verify by antisymmetry of  $\epsilon_{abc}$  the *commutator relation* for the Pauli matrices

$$[\sigma_a, \sigma_b] := \sigma_a\sigma_b - \sigma_b\sigma_a = 2i\epsilon_{abc}\sigma_c \quad \text{for } a, b, c = 1, 2, 3, \quad (3.1.11)$$

and their *anticommutator relation*

$$\{\sigma_a, \sigma_b\}_+ := \sigma_a\sigma_b + \sigma_b\sigma_a = 2\delta_{ab}\sigma_0 \quad \text{for } a, b = 1, 2, 3. \quad (3.1.12)$$

The corresponding relations among quaternions are given in (3.1.6) and (3.1.7), respectively. ★

**Exercise.** Verify the quaternionic multiplication rule expressed in the tetrad-bracket notation in (3.1.5) by using the isomorphism (3.1.8) and the product relation for the Pauli matrices in Equation (3.1.10). ★

**Answer.**

$$\begin{aligned} \mathbf{qr} &= (q_0\sigma_0 - iq_a\sigma_a)(r_0\sigma_0 - ir_b\sigma_b) \\ &= (q_0r_0 - \mathbf{q} \cdot \mathbf{r})\sigma_0 - i(q_0\mathbf{r} + r_0\mathbf{q} + \mathbf{q} \times \mathbf{r}) \cdot \boldsymbol{\sigma}. \end{aligned}$$



## 82 3 : QUATERNIONS

**Exercise.** Use Equations (3.1.11), (3.1.12) and isomorphism (3.1.8) to verify relations (3.1.6) and (3.1.7). ★

**Exercise.** Use formula (3.1.10) to verify the decomposition of a vector in Pauli matrices

$$\mathbf{q} \sigma_0 = (\mathbf{q} \cdot \boldsymbol{\sigma}) \boldsymbol{\sigma} - i \mathbf{q} \times \boldsymbol{\sigma}, \quad (3.1.13)$$

which is valid for three-vectors  $\mathbf{q} \in \mathbb{R}^3$ . Verify also that

$$-|\mathbf{q} \times \boldsymbol{\sigma}|^2 = 2|\mathbf{q}|^2 \sigma_0 = 2(\mathbf{q} \cdot \boldsymbol{\sigma})^2. \quad \star$$

**Exercise.** Use Equations (3.1.11) to verify the commutation relation

$$[\mathbf{p} \cdot \boldsymbol{\sigma}, \mathbf{q} \cdot \boldsymbol{\sigma}] = 2i \mathbf{p} \times \mathbf{q} \cdot \boldsymbol{\sigma}$$

for three-vectors  $\mathbf{p}, \mathbf{q} \in \mathbb{R}^3$ . ★

### 3.1.2 Quaternionic conjugate

**Remark 3.1.2 (Quaternionic product is associative)** The quaternionic product is *associative*:

$$\mathbf{p}(\mathbf{q}\boldsymbol{\tau}) = (\mathbf{p}\mathbf{q})\boldsymbol{\tau}. \quad (3.1.14)$$

However, the quaternionic product is not commutative,

$$[\mathbf{p}, \mathbf{q}] := \mathbf{p}\mathbf{q} - \mathbf{q}\mathbf{p} = [0, 2\mathbf{p} \times \mathbf{q}], \quad (3.1.15)$$

as we saw earlier in (3.1.6). □

**Definition 3.1.2 (Quaternionic conjugate)** One defines the *conjugate* of  $\mathbf{q} := [q_0, \mathbf{q}]$  in analogy to complex variables as

$$\mathbf{q}^* = [q_0, -\mathbf{q}]. \quad (3.1.16)$$

Following this analogy, the scalar and vector parts of a quaternion are defined as

$$\operatorname{Re} \mathbf{q} := \frac{1}{2}(\mathbf{q} + \mathbf{q}^*) = [q_0, 0], \quad (3.1.17)$$

$$\operatorname{Im} \mathbf{q} := \frac{1}{2}(\mathbf{q} - \mathbf{q}^*) = [0, \mathbf{q}]. \quad (3.1.18)$$

**Lemma 3.1.1 (Properties of quaternionic conjugation)** Two important properties of quaternionic conjugation are easily demonstrated. Namely,

$$(\mathbf{pq})^* = \mathbf{q}^* \mathbf{p}^* \quad (\text{note reversed order}), \quad (3.1.19)$$

$$\begin{aligned} \operatorname{Re}(\mathbf{pq}^*) &:= \frac{1}{2}(\mathbf{pq}^* + \mathbf{qp}^*) \\ &= [p_0 q_0 + \mathbf{p} \cdot \mathbf{q}, 0] \quad (\text{yields real part}). \end{aligned} \quad (3.1.20)$$

Note that conjugation reverses the order in the product of two quaternions.

**Definition 3.1.3 (Dot product of quaternions)** The quaternionic dot product, or inner product, is defined as

$$\begin{aligned} \mathbf{p} \cdot \mathbf{q} &= [p_0, \mathbf{p}] \cdot [q_0, \mathbf{q}] \\ &:= [p_0 q_0 + \mathbf{p} \cdot \mathbf{q}, 0] = \operatorname{Re}(\mathbf{pq}^*). \end{aligned} \quad (3.1.21)$$

**Definition 3.1.4 (Pairing of quaternions)** The quaternionic dot product (3.1.21) defines a real symmetric pairing  $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \mapsto \mathbb{R}$ , denoted as

$$\langle \mathbf{p}, \mathbf{q} \rangle = \operatorname{Re}(\mathbf{pq}^*) := \operatorname{Re}(\mathbf{qp}^*) = \langle \mathbf{q}, \mathbf{p} \rangle. \quad (3.1.22)$$

In particular,  $\langle \mathbf{q}, \mathbf{q} \rangle = \operatorname{Re}(\mathbf{qq}^*) =: |\mathbf{q}|^2$  is a positive real number.

## 84 3 : QUATERNIONS

**Definition 3.1.5 (Magnitude of a quaternion)** *The magnitude of a quaternion  $\mathbf{q}$  may be defined by*

$$|\mathbf{q}| := (\mathbf{q} \cdot \mathbf{q})^{1/2} = (q_0^2 + \mathbf{q} \cdot \mathbf{q})^{1/2}. \quad (3.1.23)$$

**Remark 3.1.3** A level set of  $|\mathbf{q}|$  defines a three-sphere  $S^3$ .  $\square$

**Definition 3.1.6 (Quaternionic inverse)** *We have the product*

$$|\mathbf{q}|^2 := \mathbf{q}\mathbf{q}^* = (\mathbf{q} \cdot \mathbf{q})\epsilon, \quad (3.1.24)$$

where  $\epsilon = [1, 0]$  is the *identity quaternion*. Hence, one may define

$$\mathbf{q}^{-1} := \mathbf{q}^*/|\mathbf{q}|^2 \quad (3.1.25)$$

to be the *inverse* of quaternion  $\mathbf{q}$ .

**Exercise.** Does a quaternion  $\mathbf{q}$  have a square root? Prove it.



**Exercise.** Show that the magnitude of the product of two quaternions is the product of their magnitudes.



**Answer.** From the definitions of the quaternionic multiplication rule (3.1.5), inner product (3.1.21) and magnitude (3.1.23), one verifies that

$$\begin{aligned} |\mathbf{p}\mathbf{q}|^2 &= (p_0q_0 - \mathbf{p} \cdot \mathbf{q})^2 + |p_0\mathbf{q} + \mathbf{p}q_0 + \mathbf{p} \times \mathbf{q}|^2 \\ &= (p_0^2 + |\mathbf{p}|^2)(q_0^2 + |\mathbf{q}|^2) = |\mathbf{p}|^2|\mathbf{q}|^2. \end{aligned}$$



**Definition 3.1.7** *A quaternion  $\mathbf{q}$  with magnitude  $|\mathbf{q}| = 1$  is called a **unit quaternion**.*



**Definition 3.1.8** A quaternion with no scalar (or real) part  $\mathbf{q} = [0, \mathbf{q}]$  is called a **pure quaternion**, or equivalently a **vector** (a term introduced by Hamilton in 1846 [Ne1997]).

**Exercise.** Show that the antisymmetric and symmetric parts of the product of two pure quaternions  $\mathbf{v} = [0, \mathbf{v}]$  and  $\mathbf{w} = [0, \mathbf{w}]$  yield, respectively, the cross product and (minus) the scalar product of the two corresponding vectors  $\mathbf{v}, \mathbf{w}$ . ★

**Answer.** The quaternionic product of pure quaternions  $\mathbf{v} = [0, \mathbf{v}]$  and  $\mathbf{w} = [0, \mathbf{w}]$  is defined as

$$\mathbf{vw} = [-\mathbf{v} \cdot \mathbf{w}, \mathbf{v} \times \mathbf{w}].$$

Its antisymmetric (vector) part yields the cross product of the corresponding vectors:

$$\text{Im}(\mathbf{vw}) = \frac{1}{2}(\mathbf{vw} - \mathbf{wv}) = [0, \mathbf{v} \times \mathbf{w}]$$

(vanishes for  $\mathbf{v} \parallel \mathbf{w}$ ).

Its symmetric (or real) part yields minus the scalar product of the vectors:

$$\text{Re}(\mathbf{vw}) = \frac{1}{2}(\mathbf{vw} + \mathbf{wv}) = [-\mathbf{v} \cdot \mathbf{w}, 0]$$

(vanishes for  $\mathbf{v} \perp \mathbf{w}$ ). ▲

**Remark 3.1.4** ( $\mathbb{H}_0 \simeq \mathbb{R}^3$ ) Being three-dimensional linear spaces possessing the same vector and scalar products, pure quaternions in  $\mathbb{H}_0$  (with no real part) are equivalent to vectors in  $\mathbb{R}^3$ . □

### 3.1.3 Decomposition of three-vectors

Pure quaternions have been identified with vectors in  $\mathbb{R}^3$ . Under this identification, the two types of products of pure quaternions

## 86 3 : QUATERNIONS

$[0, \mathbf{v}]$  and  $[0, \mathbf{w}]$  are given by

$$[0, \mathbf{v}] \cdot [0, \mathbf{w}] = [\mathbf{v} \cdot \mathbf{w}, 0] \quad \text{and} \quad [0, \mathbf{v}][0, \mathbf{w}] = [-\mathbf{v} \cdot \mathbf{w}, \mathbf{v} \times \mathbf{w}].$$

Thus, the dot ( $\cdot$ ) and cross ( $\times$ ) products of three-vectors may be identified with these two products of pure quaternions. The product of an arbitrary quaternion  $[\alpha, \boldsymbol{\chi}]$  with a pure unit quaternion  $[0, \hat{\boldsymbol{\omega}}]$  produces another pure quaternion, provided  $\boldsymbol{\chi} \cdot \hat{\boldsymbol{\omega}} = 0$ . In this case, one computes

$$[\alpha, \boldsymbol{\chi}][0, \hat{\boldsymbol{\omega}}] = [-\boldsymbol{\chi} \cdot \hat{\boldsymbol{\omega}}, \alpha \hat{\boldsymbol{\omega}} + \boldsymbol{\chi} \times \hat{\boldsymbol{\omega}}] =: [0, \mathbf{v}], \quad \text{for } \boldsymbol{\chi} \cdot \hat{\boldsymbol{\omega}} = 0. \quad (3.1.26)$$

**Remark 3.1.5** Quaternions are summoned whenever a three-vector  $\mathbf{v}$  is decomposed into its components parallel ( $\parallel$ ) and perpendicular ( $\perp$ ) to a unit three-vector direction  $\hat{\boldsymbol{\omega}}$ , according to

$$\mathbf{v} = \alpha \hat{\boldsymbol{\omega}} + \boldsymbol{\chi} \times \hat{\boldsymbol{\omega}} = [\alpha, \boldsymbol{\chi}][0, \hat{\boldsymbol{\omega}}] = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}. \quad (3.1.27)$$

Here  $\alpha = \hat{\boldsymbol{\omega}} \cdot \mathbf{v}$  and  $\boldsymbol{\chi} = \hat{\boldsymbol{\omega}} \times \mathbf{v}$  so that  $\boldsymbol{\chi} \cdot \hat{\boldsymbol{\omega}} = 0$  and one uses  $\hat{\boldsymbol{\omega}} \cdot \hat{\boldsymbol{\omega}} = 1$  to find  $\mathbf{v} \cdot \mathbf{v} = \alpha^2 + \boldsymbol{\chi}^2$  with  $\boldsymbol{\chi} := |\boldsymbol{\chi}|$ . The vector decomposition (3.1.27) is precisely the quaternionic product (3.1.26), in which the vectors  $\mathbf{v}$  and  $\hat{\boldsymbol{\omega}}$  are treated as pure quaternions.  $\square$

This remark may be summarised by the following.

**Proposition 3.1.1 (Vector decomposition)** *Quaternionic left multiplication of  $[0, \hat{\boldsymbol{\omega}}]$  by  $[\alpha, \boldsymbol{\chi}] = [\hat{\boldsymbol{\omega}} \cdot \mathbf{v}, \hat{\boldsymbol{\omega}} \times \mathbf{v}]$  decomposes the pure quaternion  $[0, \mathbf{v}]$  into components that are  $\parallel$  and  $\perp$  to the pure unit quaternion  $[0, \hat{\boldsymbol{\omega}}]$ .*

### 3.1.4 Alignment dynamics for Newton's second law

*Newton's second law of motion* is a set of ordinary differential equations for vectors of position and velocity  $(\mathbf{r}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$ . Namely,

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} \quad \text{and} \quad \frac{d\mathbf{v}}{dt} = \mathbf{f},$$

where  $\mathbf{f}$  is the *force per unit mass*. Quaternions  $[\alpha_v, \boldsymbol{\chi}_v]$  and  $[\alpha_f, \boldsymbol{\chi}_f]$  are defined by the dot ( $\cdot$ ) and cross ( $\times$ ) products of the three-

## 3.1 OPERATING WITH QUATERNIONS 87

vectors of velocity and force with the radial unit vector  $\hat{\mathbf{r}}$  as

$$[\alpha_v, \chi_v] = \left[ \hat{\mathbf{r}} \cdot \frac{d\mathbf{r}}{dt}, \hat{\mathbf{r}} \times \frac{d\mathbf{r}}{dt} \right] = [\hat{\mathbf{r}} \cdot \mathbf{v}, \hat{\mathbf{r}} \times \mathbf{v}], \text{ with } \chi_v \cdot \hat{\mathbf{r}} = 0,$$

$$[\alpha_f, \chi_f] = \left[ \hat{\mathbf{r}} \cdot \frac{d\mathbf{v}}{dt}, \hat{\mathbf{r}} \times \frac{d\mathbf{v}}{dt} \right] = [\hat{\mathbf{r}} \cdot \mathbf{f}, \hat{\mathbf{r}} \times \mathbf{f}], \text{ with } \chi_f \cdot \hat{\mathbf{r}} = 0.$$

In these equations,  $\alpha_v = \hat{\mathbf{r}} \cdot \mathbf{v} = dr/dt$  represents radial velocity,  $\chi_v = \hat{\mathbf{r}} \times \mathbf{v}$  represents angular velocity,  $\alpha_f = \hat{\mathbf{r}} \cdot \mathbf{f}$  represents radial force per unit mass and  $\chi_f = \hat{\mathbf{r}} \times \mathbf{f}$  represents *twist*, or torque per unit (mass  $\times$  length), which vanishes for central forces.

Upon decomposing into components that are parallel ( $\parallel$ ) and perpendicular ( $\perp$ ) to the position unit vector  $\hat{\mathbf{r}}$ , Newton's second law  $(d\mathbf{r}/dt, d\mathbf{v}/dt) = (\mathbf{v}, \mathbf{f})$  may be expressed as a pair of quaternionic equations,

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} = \alpha_v \hat{\mathbf{r}} + \chi_v \times \hat{\mathbf{r}} = [\alpha_v, \chi_v][0, \hat{\mathbf{r}}], \text{ using } \chi_v \cdot \hat{\mathbf{r}} = 0,$$

$$\frac{d\mathbf{v}}{dt} = \mathbf{f} = \alpha_f \hat{\mathbf{r}} + \chi_f \times \hat{\mathbf{r}} = [\alpha_f, \chi_f][0, \hat{\mathbf{r}}], \text{ using } \chi_f \cdot \hat{\mathbf{r}} = 0.$$

In this representation the alignment parameters for force  $[\alpha_f, \chi_f]$  drive those for velocity  $[\alpha_v, \chi_v]$ . That is, upon using

$$d[0, \hat{\mathbf{r}}]/dt = [0, \chi_v/r][0, \hat{\mathbf{r}}],$$

one finds

$$\left( \frac{d}{dt}[\alpha_v, \chi_v] + [\alpha_v, \chi_v][0, \chi_v/r] - [\alpha_f, \chi_f] \right)[0, \hat{\mathbf{r}}] = [0, \mathbf{0}]. \quad (3.1.28)$$

Since  $\chi_v \cdot \hat{\mathbf{r}} = 0 = \chi_f \cdot \hat{\mathbf{r}}$ , the term in parentheses vanishes.

The force vector  $\mathbf{f}$  in the orthonormal frame  $\mathbf{e}_a = (\hat{\mathbf{r}}, \hat{\mathbf{r}} \times \hat{\chi}_v, \hat{\chi}_v)$  may be expanded as

$$\mathbf{f} = \alpha_f \hat{\mathbf{r}} + \beta_f \hat{\mathbf{r}} \times \hat{\chi}_v + \gamma_f \hat{\chi}_v, \quad (3.1.29)$$

in which the coefficients  $\beta_f$  and  $\gamma_f$  are defined as

$$\beta_f = \mathbf{f} \cdot \hat{\mathbf{r}} \times \hat{\chi}_v \quad \text{and} \quad \gamma_f = \mathbf{f} \cdot \hat{\chi}_v. \quad (3.1.30)$$

## 88 3 : QUATERNIONS

In this orthonormal frame the twist vector becomes

$$\boldsymbol{\chi}_f := \hat{\mathbf{r}} \times \mathbf{f} = -\beta_f \hat{\boldsymbol{\chi}}_v + \gamma_f \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v,$$

in which  $\beta_f$  and  $\gamma_f$  represent components of the twist that are parallel and perpendicular to the direction of angular frequency  $\hat{\boldsymbol{\chi}}_v$ . As a pure quaternion, the twist vector  $[0, \boldsymbol{\chi}_f]$  may be decomposed as

$$[0, \boldsymbol{\chi}_f] = [-\beta_f, \gamma_f \hat{\mathbf{r}}][0, \hat{\boldsymbol{\chi}}_v].$$

Expanding the quaternionic representation of Newton's second law in Equation (3.1.28) provides the *alignment dynamics*,

$$\frac{d}{dt}[\alpha_v, \boldsymbol{\chi}_v] + [-\chi_v^2/r, \alpha_v \boldsymbol{\chi}_v/r] = [\alpha_f, -\beta_f \hat{\boldsymbol{\chi}}_v + \gamma_f \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v]. \quad (3.1.31)$$

The scalar part of this equation and the magnitude of its vector part yield

$$\frac{d\alpha_v}{dt} = \frac{d^2 r}{dt^2} = \frac{\chi_v^2}{r} + \alpha_f \quad \text{and} \quad \frac{d\chi_v}{dt} = -\frac{\alpha_v \chi_v}{r} - \beta_f, \quad (3.1.32)$$

with  $\alpha_v = \hat{\mathbf{r}} \cdot \mathbf{v}$ ,  $\chi_v = |\hat{\mathbf{r}} \times \mathbf{v}|$  and force components  $\alpha_f, \beta_f$  defined in (3.1.29). Its unit-vector parts yield an evolution equation for the alignment dynamics of the following *orthonormal frame*,

$$\frac{d}{dt} \begin{pmatrix} \hat{\mathbf{r}} \\ \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v \\ \hat{\boldsymbol{\chi}}_v \end{pmatrix} = \begin{pmatrix} 0 & -\chi_v/r & 0 \\ \chi_v/r & 0 & -\gamma_f/\chi_v \\ 0 & \gamma_f/\chi_v & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{r}} \\ \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v \\ \hat{\boldsymbol{\chi}}_v \end{pmatrix}, \quad (3.1.33)$$

with  $\gamma_f = \mathbf{f} \cdot \hat{\boldsymbol{\chi}}_v = \hat{\mathbf{v}} \cdot \mathbf{f} \times \hat{\mathbf{r}}$ . Thus, the alignment dynamics of Newton's second law is expressed as rotation of the orthonormal frame  $\mathbf{e}_a = (\hat{\mathbf{r}}, \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v, \hat{\boldsymbol{\chi}}_v)$ ,  $a = 1, 2, 3$  given by

## 3.1 OPERATING WITH QUATERNIONS 89

$$\frac{d\mathbf{e}_a}{dt} = \widehat{\mathcal{D}} \mathbf{e}_a = \mathbf{D} \times \mathbf{e}_a, \quad (3.1.34)$$

with angular frequency vector

$$\mathbf{D} = \frac{\chi_v}{r} \hat{\boldsymbol{\chi}}_v + \frac{\gamma_f}{\chi_v} \hat{\mathbf{r}} \quad (3.1.35)$$

lying in the  $(\hat{\mathbf{r}}, \hat{\boldsymbol{\chi}}_v)$  plane. The components of the antisymmetric matrix  $\widehat{\mathcal{D}}$  in Equation (3.1.34) are related to the components of the vector  $\mathbf{D}$  by the hat map  $\widehat{\mathcal{D}}_{ij} = -\epsilon_{ijk} D_k$ .

The three unit vectors in the orthonormal frame

$$\mathbf{e}_a = (\hat{\mathbf{r}}, \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v, \hat{\boldsymbol{\chi}}_v)^T$$

point along the position vector, opposite the nonradial component of the velocity vector, and along the angular velocity vector, respectively. For central forces, there are no torques and  $\beta_f = 0 = \gamma_f$  in Equations (3.1.32) and (3.1.33). In that case, Equation (3.1.33) implies that the direction of the angular velocity  $\hat{\boldsymbol{\chi}}_v = \hat{\mathbf{r}} \times \hat{\mathbf{v}}$  remains fixed. The other two unit vectors in the orthonormal frame  $\mathbf{e}_a$  then rotate around  $\hat{\boldsymbol{\chi}}_v$  at angular frequency  $\chi_v/r$ . For  $\chi_v \neq 0$ , one may assume  $r(t) \neq 0$  for  $r(0) \neq 0$ . (One may have  $r = 0$  for  $\chi_v \equiv 0$ , but in that case the motions are only one-dimensional and rotations are not defined.)

**Remark 3.1.6** The alignment Equation (3.1.33) governs the evolution of an orthonormal frame for any application of Newton's second law. However, this idea goes much further than particle dynamics. For example, in ideal incompressible fluid dynamics, the alignment of the vorticity vector with the velocity shear tensor in ideal incompressible fluids may be analysed by following similar steps to those used here for Newtonian mechanics [GiHoKeRo2006].

□

### 3.1.5 Quaternionic dynamics of Kepler's problem



Johannes Kepler

Newton's dynamical equation for the reduced Kepler problem is

$$\ddot{\mathbf{r}} + \frac{\mu \mathbf{r}}{r^3} = 0, \quad (3.1.36)$$

in which  $\mu$  is the gravitational constant times the reduced mass of the system of two particles in the centre of mass frame. Scale invariance of this equation under the changes  $R \rightarrow s^2 R$  and  $T \rightarrow s^3 T$  in the units of space  $R$  and time  $T$  for any constant  $s$  means that it admits families of solutions whose space and time scales are related by  $T^2/R^3 = \text{const}$  (Kepler's third law).

The reduced Kepler problem conserves the quantities

$$\begin{aligned} E &= \frac{1}{2} |\dot{\mathbf{r}}|^2 - \frac{\mu}{r} \quad (\text{energy}), \\ \mathbf{L} &= \mathbf{r} \times \dot{\mathbf{r}} \quad (\text{specific angular momentum}). \end{aligned}$$

Constancy of magnitude  $L$  means the orbit sweeps out equal areas in equal times (Kepler's second law). Constancy of direction  $\hat{\mathbf{L}}$  means the orbital position  $\mathbf{r}$  and velocity  $\dot{\mathbf{r}}$  span a plane with unit normal vector  $\hat{\mathbf{L}}$ . In that orbital plane one may specify plane polar coordinates  $(r, \theta)$  with unit vectors  $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}})$  in the plane and  $\hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}} = \hat{\mathbf{L}}$  normal to it.

The radial quaternionic variables in Equation (3.1.31) for Kepler's problem are  $\alpha_v = \dot{r}$  and  $\alpha_f = -\mu/r^2$ . The angular velocity is  $\chi_v = \mathbf{L}/r$  and the angular frequency vector in (3.1.35) is

$$\mathbf{D} = \chi_v/r = \mathbf{L}/r^2,$$

whose magnitude is  $\chi_v/r = L/r^2$ . The quantity  $\chi_f$  and its components  $\gamma_f$  and  $\beta_f$  all vanish in (3.1.29) because gravity is a central force.

## 3.1 OPERATING WITH QUATERNIONS 91

Substituting these values into the scalar alignment equations in (3.1.32) yields

$$\frac{d\alpha_v}{dt} = \frac{d^2 r}{dt^2} = \frac{L}{r^3} - \frac{\mu}{r^2} \quad \text{and} \quad \frac{d}{dt} \left( \frac{L}{r} \right) = -\frac{\dot{r}L}{r^2}. \quad (3.1.37)$$

The former equation is the balance of radial centrifugal and gravitational forces. The latter implies  $dL/dt = 0$ .

The unit vectors for polar coordinates in the orbital plane are

$$\hat{\mathbf{r}}, \quad \hat{\boldsymbol{\chi}}_v = \hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}} \quad \text{and} \quad \hat{\mathbf{r}} \times \hat{\boldsymbol{\chi}}_v = -\hat{\boldsymbol{\theta}}. \quad (3.1.38)$$

These orthogonal unit vectors form an orthonormal frame, whose alignment dynamics is governed by Equation (3.1.33) as

$$\frac{d}{dt} \begin{pmatrix} \hat{\mathbf{r}} \\ -\hat{\boldsymbol{\theta}} \\ \hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}} \end{pmatrix} = \begin{pmatrix} 0 & -L/r^2 & 0 \\ L/r^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{r}} \\ -\hat{\boldsymbol{\theta}} \\ \hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}} \end{pmatrix} = \begin{pmatrix} L/r^2 \hat{\boldsymbol{\theta}} \\ L/r^2 \hat{\mathbf{r}} \\ 0 \end{pmatrix}.$$

That is, the normal to the orbital plane is the constant unit vector  $\hat{\mathbf{r}} \times \hat{\boldsymbol{\theta}} = \hat{\mathbf{L}}$ , while

$$\frac{d\hat{\mathbf{r}}}{dt} = \dot{\theta} \hat{\boldsymbol{\theta}}, \quad \frac{d\hat{\boldsymbol{\theta}}}{dt} = -\dot{\theta} \hat{\mathbf{r}} \quad \text{and} \quad \dot{\theta} = \frac{L}{r^2}.$$

Newton's equation of motion (3.1.36) for the Kepler problem may now be written equivalently as

$$0 = \ddot{\mathbf{r}} + \frac{\mu \mathbf{r}}{r^3} = \ddot{\mathbf{r}} + \frac{\mu}{L} \dot{\theta} \hat{\mathbf{r}} = \frac{d}{dt} \left( \dot{\mathbf{r}} - \frac{\mu}{L} \hat{\boldsymbol{\theta}} \right).$$

This equation implies conservation of the following vector in the plane of motion:

$$\mathbf{K} = \dot{\mathbf{r}} - \frac{\mu}{L} \hat{\boldsymbol{\theta}} \quad (\text{Hamilton's vector}).$$

The vector in the plane given by the cross product of the two conserved vectors  $\mathbf{K}$  and  $\mathbf{L}$ ,

$$\mathbf{J} = \mathbf{K} \times \mathbf{L} = \dot{\mathbf{r}} \times \mathbf{L} - \mu \hat{\mathbf{r}} \quad (\text{Laplace-Runge-Lenz vector}),$$

## 92 3 : QUATERNIONS

is also conserved. From their definitions, these conserved quantities are related by

$$K^2 = 2E + \frac{\mu^2}{L^2} = \frac{J^2}{L^2}.$$

Choose the conserved Laplace–Runge–Lenz vector  $\mathbf{J}$  in the plane of the orbit as the reference line for the measurement of the polar angle  $\theta$ . The scalar product of  $\mathbf{r}$  and  $\mathbf{J}$  then yields an elegant result for the Kepler orbit in plane polar coordinates:

$$\mathbf{r} \cdot \mathbf{J} = rJ \cos \theta = \mathbf{r} \cdot (\dot{\mathbf{r}} \times \mathbf{L}) - \mu \mathbf{r} \cdot \hat{\mathbf{r}},$$

which implies

$$r(\theta) = \frac{L^2}{\mu + J \cos \theta} = \frac{l_{\perp}}{1 + e \cos \theta}.$$

As expected, the orbit is a *conic section* whose origin is at one of the two foci. This is *Kepler's first law*.

The Laplace–Runge–Lenz vector  $\mathbf{J}$  is directed from the focus of the orbit to its perihelion (point of closest approach). The eccentricity of the conic section is  $e = J/\mu = KL/\mu$  and its semi-latus rectum (normal distance from the line through the foci to the orbit) is  $l_{\perp} = L^2/\mu$ . The eccentricity vanishes ( $e = 0$ ) for a circle and correspondingly  $K = 0$  implies that  $\dot{\mathbf{r}} = \mu \hat{\boldsymbol{\theta}}/L$ . The eccentricity takes values  $0 < e < 1$  for an ellipse,  $e = 1$  for a parabola and  $e > 1$  for a hyperbola.

**Exercise. (Monopole Kepler problem [LeFl2003])** Consider the Kepler problem with a magnetic monopole, whose dynamical equation is, for constants  $\lambda$  and  $\mu$ ,

$$\ddot{\mathbf{r}} + \frac{\lambda}{r^3} \mathbf{L} + \left( \frac{\mu}{r^3} - \frac{\lambda^2}{r^4} \right) \mathbf{r} = 0. \quad (3.1.39)$$

Take vector cross products of this equation with  $\mathbf{r}$  and  $\mathbf{L} = \mathbf{r} \times \dot{\mathbf{r}}$  to find its conserved Hamilton's vector and



Laplace–Runge–Lenz vector. Are energy and angular momentum conserved? Are negative energy orbits ellipses? Do Kepler’s three laws still hold when  $\lambda \neq 0$ ? ★

**Exercise.** Derive Equation (3.1.39) from Hamilton’s principle.

Hint: Assume there exists a generalised function  $\mathbf{A}(\mathbf{r})$  whose curl satisfies  $\text{curl } \mathbf{A} = \mathbf{r}/r^3$ . ★

**Exercise.** Write the Hamiltonian formulation and Poisson brackets for Equation (3.1.39). Compute the Poisson brackets for  $\{\dot{r}_i, \dot{r}_j\}$ . ★

**Exercise.** Write the alignment Equations (3.1.33) for the dynamics of Equation (3.1.39). ★

---

## 3.2 Quaternionic conjugation

### 3.2.1 Cayley–Klein parameters

**Definition 3.2.1 (Quaternionic conjugation)** *Quaternionic conjugation* is defined as the map under the quaternionic product (recalling that it is associative),

$$\mathfrak{r} \rightarrow \mathfrak{r}' = \hat{\mathfrak{q}} \mathfrak{r} \hat{\mathfrak{q}}^*, \quad (3.2.1)$$

## 94 3 : QUATERNIONS

where  $\hat{q} = [q_0, \mathbf{q}]$  is a unit quaternion,  $\hat{q} \cdot \hat{q} = 1$ , so  $\hat{q}\hat{q}^* = \mathbf{e} = [1, 0]$ . The inverse map is

$$\mathbf{r} = \hat{q}^* \mathbf{r} \hat{q}.$$

**Exercise.** Show that the product of a quaternion  $\mathbf{r} = [r_0, \mathbf{r}]$  with a unit quaternion  $\hat{q} = [q_0, \mathbf{q}]$ , whose inverse is  $\hat{q}^* = [q_0, -\mathbf{q}]$ , satisfies

$$\mathbf{r}\hat{q}^* = [\mathbf{r} \cdot \hat{q}, -r_0\mathbf{q} + q_0\mathbf{r} + \mathbf{q} \times \mathbf{r}],$$

$$\hat{q} \mathbf{r} \hat{q}^* = [r_0|\hat{q}|^2, \mathbf{r} + 2q_0\mathbf{q} \times \mathbf{r} + 2\mathbf{q} \times (\mathbf{q} \times \mathbf{r})],$$

where  $\mathbf{r} \cdot \hat{q} = r_0q_0 + \mathbf{r} \cdot \mathbf{q}$  and  $|\hat{q}|^2 = \hat{q} \cdot \hat{q} = q_0^2 + \mathbf{q} \cdot \mathbf{q} = 1$  according to the definitions of the dot product in (3.1.21) and magnitude in (3.1.23). ★

**Remark 3.2.1** The same products using the pure unit quaternion  $\hat{\mathbf{z}} = [0, \hat{\mathbf{z}}]$  with  $\hat{\mathbf{z}} = (0, 0, 1)^T$  and the unit quaternion  $\hat{q} = [q_0, \mathbf{q}]$  satisfy

$$\hat{\mathbf{z}}\hat{q}^* = [q_3, q_0\hat{\mathbf{z}} + \mathbf{q} \times \hat{\mathbf{z}}],$$

$$\hat{q} \hat{\mathbf{z}} \hat{q}^* = [0, \hat{\mathbf{z}} + 2q_0\mathbf{q} \times \hat{\mathbf{z}} + 2\mathbf{q} \times (\mathbf{q} \times \hat{\mathbf{z}})],$$

which produces a complete set of unit vectors. □

**Remark 3.2.2** Conjugation  $\hat{q} \mathbf{r} \hat{q}^*$  is a wise choice, as opposed to, say, choosing the apparently less meaningful triple product

$$\hat{q} \mathbf{r} \hat{q} = [0, \mathbf{r}] + (r_0q_0 - \mathbf{r} \cdot \mathbf{q})[q_0, \mathbf{q}]$$

for quaternions  $\mathbf{r} = [r_0, \mathbf{r}]$  and  $\hat{q} = [q_0, \mathbf{q}]$  with  $|\mathbf{q}|^2 = q_0^2 + \mathbf{q} \cdot \mathbf{q} = 1$ . □

**Exercise.** For  $\mathfrak{q}^* = [q_0, -\mathfrak{q}]$ , such that  $\mathfrak{q}^*\mathfrak{q} = \mathbb{J}_0|\mathfrak{q}|^2$ , verify that

$$2\mathfrak{q}^* = -\mathbb{J}_0\mathfrak{q}\mathbb{J}_0^* + \mathbb{J}_1\mathfrak{q}\mathbb{J}_1^* + \mathbb{J}_2\mathfrak{q}\mathbb{J}_2^* + \mathbb{J}_3\mathfrak{q}\mathbb{J}_3^*.$$

What does this identity mean geometrically? Does the complex conjugate  $z^*$  for  $z \in \mathbb{C}$  satisfy such an identity? Prove it. ★

**Lemma 3.2.1** *As a consequence of Remark 3.2.1 and the Exercise just before it, one finds that conjugation  $\hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^*$  of a quaternion  $\mathfrak{r}$  by a unit quaternion  $\hat{\mathfrak{q}}$  preserves the sphere  $S_{|\mathfrak{r}|}^3$  given by any level set of  $|\mathfrak{r}|$ . That is, the value of  $|\mathfrak{r}|^2$  is invariant under conjugation by a unit quaternion:*

$$|\hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^*|^2 = |\mathfrak{r}|^2 = r_0^2 + \mathbf{r} \cdot \mathbf{r}. \quad (3.2.2)$$

**Definition 3.2.2 (Conjugacy classes)** *The set*

$$C(\mathfrak{r}) := \left\{ \mathfrak{r}' \in \mathbb{H} \mid \mathfrak{r}' = \hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^* \right\} \quad (3.2.3)$$

*is called the **conjugacy class** of the quaternion  $\mathfrak{r}$ .*

**Corollary 3.2.1** *The conjugacy classes of the three-sphere  $S_{|\mathfrak{r}|}^3$  under conjugation by a unit quaternion  $\hat{\mathfrak{q}}$  are the two-spheres given by*

$$\left\{ \mathbf{r} \in \mathbb{R}^3 \mid \mathbf{r} \cdot \mathbf{r} = |\mathfrak{r}|^2 - r_0^2 \right\}. \quad (3.2.4)$$

**Proof.** The proof is a straightforward exercise. ■

**Remark 3.2.3** The expressions in Remark 3.2.1 correspond to *spatial rotations* when  $r_0 = 0$  so that  $\mathfrak{r} = [0, \mathbf{r}]$ . □

## 96 3 : QUATERNIONS

**Lemma 3.2.2 (Euler–Rodrigues formula)** *If  $\mathbf{r} = [0, \mathbf{r}]$  is a pure quaternion and  $\hat{\mathbf{q}} = [q_0, \mathbf{q}]$  is a unit quaternion, then under quaternionic conjugation,*

$$\begin{aligned}\mathbf{r}' &= \hat{\mathbf{q}} \mathbf{r} \hat{\mathbf{q}}^* = [0, \mathbf{r}'] \\ &= [0, \mathbf{r} + 2q_0(\mathbf{q} \times \mathbf{r}) + 2\mathbf{q} \times (\mathbf{q} \times \mathbf{r})].\end{aligned}\quad (3.2.5)$$

For  $\hat{\mathbf{q}} := \pm[\cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}}]$ , we have

$$[0, \mathbf{r}'] = \left[ \cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}} \right] [0, \mathbf{r}] \left[ \cos \frac{\theta}{2}, -\sin \frac{\theta}{2} \hat{\mathbf{n}} \right],$$

so that

$$\begin{aligned}\mathbf{r}' &= \mathbf{r} + 2 \cos \frac{\theta}{2} \sin \frac{\theta}{2} (\hat{\mathbf{n}} \times \mathbf{r}) + 2 \sin^2 \frac{\theta}{2} (\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{r})) \\ &= \mathbf{r} + \sin \theta (\hat{\mathbf{n}} \times \mathbf{r}) + (1 - \cos \theta) (\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{r})) \quad (3.2.6) \\ &=: O_{\hat{\mathbf{n}}}^{\theta} \mathbf{r}.\end{aligned}$$

This is the famous **Euler–Rodrigues formula** for the rotation  $O_{\hat{\mathbf{n}}}^{\theta} \mathbf{r}$  of a vector  $\mathbf{r}$  by an angle  $\theta$  about the unit vector  $\hat{\mathbf{n}}$ .

**Exercise.** Verify the Euler–Rodrigues formula (3.2.6) by a direct computation using quaternionic multiplication. ★

**Exercise.** Write formula (3.2.5) for conjugation of a pure quaternion by a unit quaternion  $q_0^2 + \mathbf{q} \cdot \mathbf{q} = 1$  as a  $3 \times 3$  matrix operation acting on a vector. ★

**Answer.** As a  $3 \times 3$  matrix operation acting on a vector,  $\mathbf{r}' = O_{3 \times 3} \mathbf{r}$ , formula (3.2.5) becomes

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} + 2q_0(\mathbf{q} \times \mathbf{r}) + 2\mathbf{q} \times (\mathbf{q} \times \mathbf{r}) \\ &= \left[ (2q_0^2 - 1)Id + 2q_0\hat{\mathbf{q}} + 2\mathbf{q}\mathbf{q}^T \right] \mathbf{r} =: O_{3 \times 3} \mathbf{r}, \end{aligned}$$

where  $\hat{\mathbf{q}} = \mathbf{q} \times$ , or in components  $\hat{q}_{lm} = -q^k \epsilon_{klm}$  by the hat map in (2.1.10) and (2.1.11). When  $\mathbf{q} = [q_0, \mathbf{q}]$  is a unit quaternion, the Euler–Rodrigues formula implies  $O_{3 \times 3} \in SO(3)$ .  $\blacktriangle$

**Definition 3.2.3 (Euler parameters)** *In the Euler–Rodrigues formula (3.2.6) for the rotation of vector  $\mathbf{r}$  by angle  $\theta$  about  $\hat{\mathbf{n}}$ , the quantities  $\theta, \hat{\mathbf{n}}$  are called the **Euler parameters**.*

**Definition 3.2.4 (Cayley–Klein parameters)** *The unit quaternion  $\hat{\mathbf{q}} = [q_0, \mathbf{q}]$  corresponding to the rotation of a pure quaternion  $\mathbf{r} = [0, \mathbf{r}]$  by angle  $\theta$  about  $\hat{\mathbf{n}}$  using quaternionic conjugation is*

$$\hat{\mathbf{q}} := \pm \left[ \cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}} \right]. \quad (3.2.7)$$

*The quantities  $q_0 = \pm \cos \frac{\theta}{2}$  and  $\mathbf{q} = \pm \sin \frac{\theta}{2} \hat{\mathbf{n}}$  in (3.2.7) are called the **Cayley–Klein parameters**.*

**Remark 3.2.4 (Cayley–Klein coordinates of a quaternion)** An arbitrary quaternion may be written in terms of its magnitude and its Cayley–Klein parameters as

$$\mathbf{q} = |\mathbf{q}| \hat{\mathbf{q}} = |\mathbf{q}| \left[ \cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}} \right]. \quad (3.2.8)$$

□

The calculation of the Euler–Rodrigues formula (3.2.6) shows the equivalence of quaternionic conjugation and rotations of vectors. Moreover, compositions of quaternionic products imply the following.

## 98 3 : QUATERNIONS

**Corollary 3.2.2** *Composition of rotations*

$$O_{\hat{\mathbf{n}}'}^{\theta'} O_{\hat{\mathbf{n}}}^{\theta} \mathbf{r} = \hat{\mathbf{q}}' (\hat{\mathbf{q}} \tau \hat{\mathbf{q}}^*) \hat{\mathbf{q}}'^*$$

is equivalent to multiplication of  $(\pm)$  unit quaternions.

**Exercise.** Show directly by quaternionic multiplication that

$$O_{\hat{\mathbf{y}}}^{\pi/2} O_{\hat{\mathbf{x}}}^{\pi/2} = O_{\hat{\mathbf{n}}}^{2\pi/3} \quad \text{with} \quad \hat{\mathbf{n}} = (\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}})/\sqrt{3}. \quad \star$$

**Answer.** One multiplies the corresponding unit quaternions, yielding

$$\begin{aligned} \left[ \frac{1}{\sqrt{2}}, \hat{\mathbf{y}} \frac{1}{\sqrt{2}} \right] \left[ \frac{1}{\sqrt{2}}, \hat{\mathbf{x}} \frac{1}{\sqrt{2}} \right] &= \left[ \frac{1}{2}, \frac{1}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}}) \right] \\ &= \left[ \frac{1}{2}, \frac{1}{\sqrt{3}}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}}) \frac{\sqrt{3}}{2} \right] \\ &= \left[ \cos \frac{\pi}{3}, \frac{1}{\sqrt{3}}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}}) \sin \frac{\pi}{3} \right] \\ &= \left[ \cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}} \right], \end{aligned}$$

which is the Cayley–Klein form for a rotation of

$$\theta = 2\pi/3 \quad \text{about} \quad \hat{\mathbf{n}} = (\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}})/\sqrt{3}. \quad \blacktriangle$$

**Exercise.** Compute  $O_{\hat{\mathbf{y}}}^{\pi} O_{\hat{\mathbf{x}}}^{\pi} - O_{\hat{\mathbf{x}}}^{\pi} O_{\hat{\mathbf{y}}}^{\pi}$  by quaternionic multiplication. Does it vanish? Prove it.  $\star$

**Remark 3.2.5 (Cayley–Klein parameters for three-vectors)** Consider the unit Cayley–Klein quaternion,  $\hat{p} := \pm[\cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\chi}]$ . Then the decompositions for quaternions (3.1.26) and for vectors (3.1.27) may be set equal to find

$$\begin{aligned} [0, \hat{\mathbf{v}}] &:= |\mathbf{v}|^{-1}[0, \mathbf{v}] = \hat{p}[0, \hat{\omega}]\hat{p}^* \\ &= [0, \cos \theta \hat{\omega} + \sin \theta \hat{\chi} \times \hat{\omega}] \\ &= |\mathbf{v}|^{-1}[\alpha, \boldsymbol{\chi}][0, \hat{\omega}] \\ &= (\alpha^2 + \chi^2)^{-1/2} [0, \alpha \hat{\omega} + \boldsymbol{\chi} \times \hat{\omega}]. \end{aligned}$$

Thus, the unit vector  $\hat{\mathbf{v}} = |\mathbf{v}|^{-1}\mathbf{v}$  is a rotation of  $\hat{\omega}$  by angle  $\theta$  around  $\hat{\chi}$  with

$$\cos \theta = \frac{\alpha}{(\alpha^2 + \chi^2)^{1/2}} \quad \text{and} \quad \sin \theta = \frac{\chi}{(\alpha^2 + \chi^2)^{1/2}}.$$

Hence, the alignment parameters  $\alpha$  and  $\chi$  in (3.1.26) and (3.1.27) define the three-vector  $\mathbf{v}$  in  $[0, \mathbf{v}] = [\alpha, \boldsymbol{\chi}][0, \hat{\omega}]$  as a stretching of  $\hat{\omega}$  by  $(\alpha^2 + \chi^2)^{1/2}$  and a rotation of  $\hat{\omega}$  by  $\theta = \tan^{-1} \chi/\alpha$  about  $\hat{\chi}$ . The Cayley–Klein angle  $\theta$  is the *relative angle* between the directions  $\hat{\mathbf{v}}$  and  $\hat{\omega}$ .  $\square$

### 3.2.2 Pure quaternions, Pauli matrices and $SU(2)$

**Exercise.** Write the product of two pure unit quaternions as a multiplication of Pauli matrices.  $\star$

**Answer.** By the quaternionic multiplication rule (3.1.5), one finds

$$[0, \hat{\mathbf{v}}][0, \hat{\mathbf{w}}] = [-\hat{\mathbf{v}} \cdot \hat{\mathbf{w}}, \hat{\mathbf{v}} \times \hat{\mathbf{w}}] =: [\cos \theta, \hat{\mathbf{n}} \sin \theta]. \quad (3.2.9)$$

## 100 3 : QUATERNIONS

Here  $\hat{\mathbf{v}} \cdot \hat{\mathbf{w}} = -\cos \theta$ , so that  $\theta$  is the relative angle between the unit three-vectors  $\hat{\mathbf{v}}$  and  $\hat{\mathbf{w}}$ , and  $\hat{\mathbf{v}} \times \hat{\mathbf{w}} = \hat{\mathbf{n}} \sin \theta$  is their cross product, satisfying

$$|\hat{\mathbf{v}} \times \hat{\mathbf{w}}|^2 = |\hat{\mathbf{v}}|^2 |\hat{\mathbf{w}}|^2 - (\hat{\mathbf{v}} \cdot \hat{\mathbf{w}})^2 = 1 - \cos^2 \theta = \sin^2 \theta.$$

This is equivalent to following the multiplication of Pauli matrices,

$$\begin{aligned} (-i\hat{\mathbf{v}} \cdot \boldsymbol{\sigma})(-i\hat{\mathbf{w}} \cdot \boldsymbol{\sigma}) &= -\hat{\mathbf{v}} \cdot \hat{\mathbf{w}} \sigma_0 - i\hat{\mathbf{v}} \times \hat{\mathbf{w}} \cdot \boldsymbol{\sigma} \\ &= -(\cos \theta \sigma_0 + i \sin \theta \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}), \end{aligned} \quad (3.2.10)$$

with, e.g.,  $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} = \sum_{a=1}^3 \hat{n}_a \sigma_a$ . ▲

**Proposition 3.2.1 (De Moivre's theorem for quaternions)** *De Moivre's theorem for complex numbers of unit modulus is*

$$(\cos \theta + i \sin \theta)^m = (\cos m\theta + i \sin m\theta).$$

*The analogue of De Moivre's theorem for unit quaternions is*

$$[\cos \theta, \sin \theta \hat{\mathbf{n}}]^m = [\cos m\theta, \sin m\theta \hat{\mathbf{n}}].$$

**Proof.** The proof follows immediately from the Cayley–Klein representation of a unit quaternion. ■

**Theorem 3.2.1** *The unit quaternions form a representation of the matrix Lie group  $SU(2)$ .*

**Proof.** The matrix representation of a unit quaternion is given in (3.1.8). Let  $\hat{\mathbf{q}} = [q_0, \mathbf{q}]$  be a unit quaternion ( $|\hat{\mathbf{q}}|^2 = q_0^2 + \mathbf{q} \cdot \mathbf{q} = 1$ ) and define the matrix  $Q$  by

$$\begin{aligned} Q &= q_0 \sigma_0 - i\mathbf{q} \cdot \boldsymbol{\sigma} \\ &= \begin{bmatrix} q_0 - iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & q_0 + iq_3 \end{bmatrix}. \end{aligned} \quad (3.2.11)$$



The matrix  $Q$  is a unitary  $2 \times 2$  matrix ( $QQ^\dagger = Id$ ) with unit determinant ( $\det Q = 1$ ). That is,  $Q \in SU(2)$ . In fact, we may rewrite the map (3.2.1) for quaternionic conjugation of a vector  $\mathbf{r} = [0, \mathbf{r}]$  by a unit quaternion equivalently in terms of unitary conjugation of the Hermitian Pauli spin matrices as

$$\mathbf{r}' = \hat{q} \mathbf{r} \hat{q}^* \iff \mathbf{r}' \cdot \boldsymbol{\sigma} = Q \mathbf{r} \cdot \boldsymbol{\sigma} Q^\dagger, \quad (3.2.12)$$

with

$$\mathbf{r} \cdot \boldsymbol{\sigma} = \begin{bmatrix} r_3 & r_1 - ir_2 \\ r_1 + ir_2 & -r_3 \end{bmatrix}. \quad (3.2.13)$$

This is the standard representation of  $SO(3)$  rotations as a double covering ( $\pm J$ ) by  $SU(2)$  matrices, which is now seen to be equivalent to quaternionic multiplication. ■

**Remark 3.2.6** A variant of the map (3.2.11), known as the Kustaanheimo–Stiefel map [KuSt1965], establishes a relation between the solutions of a constrained isotropic harmonic oscillator in four dimensions and those of the Kepler problem in three dimensions. The KS map is beyond our present scope. However, for an interesting discussion of it, see [Co2003]. □

**Remark 3.2.7** Composition of  $SU(2)$  matrices by matrix multiplication forms a Lie subgroup of the Lie group of  $2 \times 2$  complex matrices  $GL(2, \mathbb{C})$ , see, e.g., [MaRa1994]. □

**Exercise.** Check that the matrix  $Q$  in (3.2.11) is a special unitary matrix so that  $Q \in SU(2)$ . That is, show that  $Q$  is unitary and has unit determinant. ★

## 102 3 : QUATERNIONS

**Exercise.** Verify the conjugacy formula (3.2.12) arising from the isomorphism between unit quaternions and  $SU(2)$ . ★

**Remark 3.2.8** The  $(\pm)$  in the Cayley–Klein parameters reflects the 2:1 covering of the map  $SU(2) \rightarrow SO(3)$ . □

### 3.2.3 Tilde map: $\mathbb{R}^3 \simeq su(2) \simeq so(3)$

The following *tilde map* may be defined by considering the isomorphism (3.1.8) for a pure quaternion  $[0, \mathbf{q}]$ . Namely,

$$\begin{aligned} \mathbf{q} \in \mathbb{R}^3 \mapsto -i\mathbf{q} \cdot \boldsymbol{\sigma} &= -i \sum_{j=1}^3 q_j \sigma_j & (3.2.14) \\ &= \begin{bmatrix} -iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & iq_3 \end{bmatrix} =: \tilde{\mathbf{q}} \in su(2). \end{aligned}$$

The tilde map (3.2.14) is a Lie algebra isomorphism between  $\mathbb{R}^3$  with the cross product of vectors and the Lie algebra  $su(2)$  of  $2 \times 2$  skew-Hermitian traceless matrices. Just as in the hat map one writes

$$JJ^\dagger(t) = Id \implies JJ^\dagger + (JJ^\dagger)^\dagger = 0,$$

so the tangent space at the identity for the  $SU(2)$  matrices comprises  $2 \times 2$  skew-Hermitian traceless matrices, whose basis is  $-i\boldsymbol{\sigma}$ , the imaginary number  $(-i)$  times the three Pauli matrices. This completes the *circle of the isomorphisms* between Pauli matrices and quaternions, and between *pure* quaternions and vectors in  $\mathbb{R}^3$ . In particular, their Lie products are all isomorphic. That is,

$$\text{Im}(\mathbf{p}\mathbf{q}) = \frac{1}{2}(\mathbf{p}\mathbf{q} - \mathbf{q}\mathbf{p}) = [\tilde{\mathbf{p}}, \tilde{\mathbf{q}}] = (\mathbf{p} \times \mathbf{q})^\sim. \quad (3.2.15)$$

In addition, recalling that  $\text{Re}(\mathfrak{p}\mathfrak{q}^*) = [\mathfrak{p} \cdot \mathfrak{q}, 0]$  helps prove the following identities:

$$\det(\mathfrak{q} \cdot \sigma) = |\mathfrak{q}|^2, \quad (\tilde{\mathfrak{p}}\tilde{\mathfrak{q}}) = -\mathfrak{p} \cdot \mathfrak{q}.$$

### 3.2.4 Dual of the tilde map: $\mathbb{R}^{3*} \simeq su(2)^* \simeq so(3)^*$

One may identify  $su(2)^*$  with  $\mathbb{R}^3$  via the map  $\mu \in su(2)^* \rightarrow \check{\mu} \in \mathbb{R}^3$  defined by

$$\check{\mu} \cdot \mathfrak{q} := \langle \mu, \tilde{\mathfrak{q}} \rangle_{su(2)^* \times su(2)}$$

for any  $\mathfrak{q} \in \mathbb{R}^3$ .

Then, for example,

$$\check{\mu} \cdot (\mathfrak{p} \times \mathfrak{q}) := \langle \mu, [\tilde{\mathfrak{p}}, \tilde{\mathfrak{q}}] \rangle_{su(2)^* \times su(2)},$$

which foreshadows the adjoint and coadjoint actions of  $SU(2)$  to appear in our discussions of rigid-body dynamics in Chapter 5 and momentum maps in Chapter 11.

### 3.2.5 Pauli matrices and Poincaré's sphere $\mathbb{C}^2 \rightarrow S^2$

The Lie algebra isomorphisms given by the Pauli matrix representation of the quaternions (3.1.8) and the tilde map (3.2.14) are related to a map  $\mathbb{C}^2 \mapsto S^2$  first introduced by Poincaré [Po1892] and later studied by Hopf [Ho1931]. Consider for  $a_k \in \mathbb{C}^2$ , with  $k = 1, 2$  the four real combinations written in terms of the Pauli matrices

$$n_\alpha = \sum_{k,l=1}^2 a_k^* \{\sigma_\alpha\}_{kl} a_l \quad \text{with } \alpha = 0, 1, 2, 3. \quad (3.2.16)$$

The  $n_\alpha \in \mathbb{R}^4$  have components

$$\begin{aligned} n_0 &= |a_1|^2 + |a_2|^2, \\ n_3 &= |a_1|^2 - |a_2|^2, \\ n_1 + i n_2 &= 2a_1^* a_2. \end{aligned} \quad (3.2.17)$$

## 104 3 : QUATERNIONS

**Remark 3.2.9** One may motivate the definition of  $n_\alpha \in \mathbb{R}^4$  in (3.2.16) by introducing the following Hermitian matrix,

$$\rho = \mathbf{a} \otimes \mathbf{a}^* = \frac{1}{2} \left( n_0 \sigma_0 + \mathbf{n} \cdot \boldsymbol{\sigma} \right), \quad (3.2.18)$$

in which the vector  $\mathbf{n}$  is defined as

$$\mathbf{n} = \text{tr } \rho \boldsymbol{\sigma} = a_l a_k^* \boldsymbol{\sigma}_{kl}. \quad (3.2.19)$$

The last equation recovers (3.2.16). We will return to the interpretation of this map when we discuss momentum maps in Chapter 11. For now, we simply observe that the components of the singular Hermitian matrix ( $\det \rho = 0$ )

$$\rho = \mathbf{a} \otimes \mathbf{a}^* = \frac{1}{2} \begin{bmatrix} n_0 + n_3 & n_1 - in_2 \\ n_1 + in_2 & n_0 - n_3 \end{bmatrix}$$

are all invariant under the diagonal action

$$S^1 : \mathbf{a} \rightarrow e^{i\phi} \mathbf{a}, \mathbf{a}^* \rightarrow e^{-i\phi} \mathbf{a}^*.$$

□

A fixed value  $n_0 = \text{const}$  defines a three-sphere  $S^3 \in \mathbb{R}^4$ . Moreover, because  $\det \rho = 0$  the remaining three components satisfy an additional relation which defines the *Poincaré sphere*  $S^2 \in S^3$  as

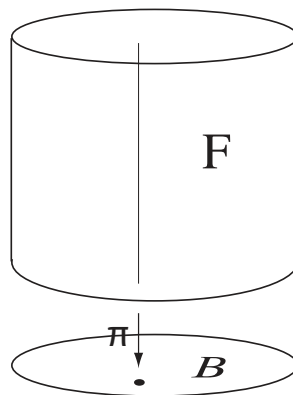
$$n_0^2 = n_1^2 + n_2^2 + n_3^2 = |\mathbf{n}|^2. \quad (3.2.20)$$

Each point on this sphere defines a direction introduced by Poincaré to represent polarised light. The north (resp. south) pole represents right (resp. left) circular polarisation and the equator represents the various inclinations of linear polarisation. Off the equator and the poles the remaining directions in the upper and lower hemispheres represent right- and left-handed elliptical polarisations, respectively. Opposing directions  $\pm \mathbf{n}$  correspond to orthogonal polarisations. See [BoWo1965] for details of the physical interpretation of the Poincaré sphere for polarised ray optics.

### 3.2.6 Poincaré's sphere and Hopf's fibration

The same map  $S^3 \mapsto S^2$  given by (3.2.16) from the  $n_0 = \text{const}$   $S^3$  to the Poincaré sphere  $S^2$  was later studied by Hopf [Ho1931], who found it to be a **fibration** of  $S^3$  over  $S^2$ . That is,  $S^3 \simeq S^2 \times S^1$  locally, where  $S^1$  is the fibre. A fibre bundle structure is defined descriptively, as follows.

**Definition 3.2.5 (Fibre bundle)** *In topology, a **fibre bundle** is a space which locally looks like a product of two spaces but may possess a different global structure. Every fibre bundle consists of a continuous surjective map  $\pi : E \mapsto B$ , where small regions in the total space  $E$  look like small regions in the product space  $B \times F$ , of the **base space**  $B$  with the **fibre space**  $F$  (Figure 3.1). Fibre bundles comprise a rich mathematical subject that is explained more completely in, e.g., [Is1999, La1999, Wa1983, Sp1979]. We shall confine our attention here to the one particular case leading to the Poincaré sphere.*



**Figure 3.1.** A fibre bundle  $E$  looks locally like the product space  $B \times F$ , of the base space  $B$  with the fibre space  $F$ . The map  $\pi : E \approx B \times F \mapsto B$  projects  $E$  onto the base space  $B$ .

## 106 3 : QUATERNIONS

**Remark 3.2.10** The *Hopf fibration*, or fibre bundle,  $S^3 \simeq S^2 \times S^1$  has spheres as its total space, base space and fibre, respectively. In terms of the Poincaré sphere one may think of the Hopf fibration locally as a sphere  $S^2$  which has a great circle  $S^1$  attached at every point. The phase on the great circles at opposite points are orthogonal (rotated by  $\pi/2$ , not  $\pi$ ); so passing once around the Poincaré sphere along a great circle rotates the  $S^1$  phase only by  $\pi$ , not  $2\pi$ . One must pass twice around a great circle on the Poincaré sphere to return to the original phase. Thus, the relation  $S^3 \simeq S^2 \times S^1$  only holds locally, not globally.  $\square$

**Remark 3.2.11** The conjugacy classes of  $S^3$  by unit quaternions yield the family of two-spheres  $S^2$  in formula (3.2.4) of Corollary 3.2.1. These also produce a version of the Hopf fibration  $S^3 \simeq S^2 \times S^1$ , obtained by identifying the Poincaré sphere (3.2.20) from the definitions (3.2.17).  $\square$

**Remark 3.2.12 (Hopf fibration/quaternionic conjugation)** Conjugating the pure unit quaternion along the  $z$ -axis  $[0, \hat{\mathbf{z}}]$  by the other unit quaternions yields the entire unit two-sphere  $S^2$ . This is to be expected from the complete set of unit vectors found by quaternionic conjugation in (3.2.2). However, it may be shown explicitly by computing the  $SU(2)$  multiplication for  $|a_1|^2 + |a_2|^2 = 1$ ,

$$\begin{bmatrix} a_1 & -a_2^* \\ a_2 & a_1^* \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} a_1^* & a_2^* \\ -a_2 & a_1 \end{bmatrix} = \begin{bmatrix} -in_3 & -in_1 + n_2 \\ -in_1 - n_2 & in_3 \end{bmatrix}.$$

This is the tilde map (3.2.14) once again and  $(n_1, n_2, n_3)$  are the components of the Hopf fibration [MaRa1994, El2007].  $\square$

In other words, cf. Equation (3.2.11),

$$-ig\sigma_3g^\dagger = -i\mathbf{n} \cdot \boldsymbol{\sigma},$$

in which  $g^\dagger = g^{-1} \in SU(2)$  and  $|\mathbf{n}|^2 = 1$ . This is the tilde map (3.2.14) once again and  $(n_1, n_2, n_3)$  are the spatial components of the Hopf fibration in (3.2.17).

## 3.2 QUATERNIONIC CONJUGATION 107

Thus, the isomorphism given in (3.1.8) and (3.2.14) between the unit quaternions and  $SU(2)$  expressed in terms of the Pauli spin matrices connects the quaternions to the mathematics of Poincaré's sphere  $\mathbb{C}^2 \mapsto S^2$ , Hopf's fibration  $S^3 \simeq S^2 \times S^1$  and the geometry of fibre bundles. This is a very deep network of connections that will amply reward the efforts of further study.

**Exercise.** Show that the Hopf fibration is a decomposition law for the group  $SU(2)$ .

Hint: Write the Hopf fibration in quaternionic form. ★

**Exercise.** Write the quaternionic version of unitary transformations of Hermitian matrices.

Hint: The Pauli spin matrices defined in (3.1.9) are Hermitian. To get started, you may want to take a look at Equation (3.2.12). ★

**Exercise.** Write the quaternionic version of orthogonal transformations of symmetric matrices as in Section 2.5.6.

★

**Remark 3.2.13** In quantum mechanics, the quantity corresponding to the Hermitian matrix  $\rho = \mathbf{a} \otimes \mathbf{a}^*$  in Equation (3.2.18) is called the *density matrix* and the Poincaré sphere in Equation (3.2.20) is called the *Bloch sphere*. □

### 3.2.7 Coquaternions

Not long after Hamilton discovered quaternions, James Cockle [Co1848] proposed an alternative field of numbers called *coquaternions*. A coquaternion  $c = [w, x, y, z]$  is defined by four real numbers, as follows:

$$c = w\mathbb{1} + xi + yj + zk, \quad \text{with } (w, x, y, z) \in \mathbb{R}^4,$$

where  $\mathbb{1}$  is the identity element and the other three coquaternion basis elements ( $i, j, k$ ) satisfy the multiplication rules,

$$\begin{aligned} ij = k = -ji, \quad jk = -i = -kj, \quad ki = j = -ik, \\ \mathbb{1}^2 = \mathbb{1}, \quad i^2 = -\mathbb{1}, \quad j^2 = \mathbb{1}, \quad k^2 = \mathbb{1}, \quad ijk = \mathbb{1}. \end{aligned}$$

The multiplication table of the coquaternion basis may be expressed as

	$\mathbb{1}$	$i$	$j$	$k$
$\mathbb{1}$	$\mathbb{1}$	$i$	$j$	$k$
$i$	$i$	$-\mathbb{1}$	$k$	$-j$
$j$	$j$	$-k$	$\mathbb{1}$	$-i$
$k$	$k$	$j$	$i$	$\mathbb{1}$

(3.2.21)

The multiplication rules for coquaternions may be represented in a basis of real  $2 \times 2$  matrices,

$$\begin{aligned} \mathbb{1} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad i = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \\ j &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad k = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned} \quad (3.2.22)$$

This means the coquaternion  $c = [w, x, y, z]$  may be represented as a *real matrix*

$$C = \begin{bmatrix} w + z & y + x \\ y - x & w - z \end{bmatrix}, \quad (3.2.23)$$



with determinant  $\det(C) = w^2 + x^2 - y^2 - z^2$ .

One defines the *conjugate* of the coquaternion  $\mathfrak{c} = [w, x, y, z]$  as  $\mathfrak{c}^* = [w, -x, -y, -z]$ ,

$$\mathfrak{c}^* = w\mathbb{1} - xi - yj - zk.$$

Thus, in a matrix representation, the conjugate coquaternion is expressed as

$$C^* = \begin{bmatrix} w - z & -y - x \\ -y + x & w + z \end{bmatrix}. \quad (3.2.24)$$

An *inner product* for coquaternions based on the matrix inner product may be defined as

$$(\mathfrak{c}_1, \mathfrak{c}_2) = \mathfrak{c}_1^* \cdot \mathfrak{c}_2 := \frac{1}{2} \text{tr}(C_1^* C_2) = w_1 w_2 + x_1 x_2 - y_1 y_2 - z_1 z_2.$$

The inner product of a coquaternion with its conjugate then defines its squared magnitude

$$|\mathfrak{c}|^2 = \mathfrak{c}^* \cdot \mathfrak{c} := \frac{1}{2} \text{tr}(C^* C) = w^2 + x^2 - y^2 - z^2 = \det(C),$$

which is indefinite in sign.

**Exercise.** Show that the unit coquaternions with  $|\mathfrak{c}|^2 = 1$  form a representation of the Lie group of  $2 \times 2$  symplectic matrices  $Sp(2, \mathbb{R})$ .

Hint: A symplectic matrix  $M \in Sp(2, \mathbb{R})$  satisfies

$$MJM^T = J \quad \text{for} \quad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (3.2.25)$$

★

**110** 3 : QUATERNIONS

**Remark 3.2.14** For a discussion of symplectic matrices in the context of the geometric mechanics approach to ray optics via Fermat's principle, see, e.g., [Ho2008].  $\square$

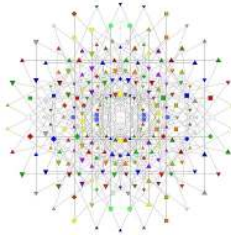
**Exercise.** Follow the developments in the earlier part of this chapter for quaternions far enough to define conjugacy classes for the action of unit coquaternions on vectors in  $\mathbb{R}^3$ .



**Exercise.** Compute the Euler–Rodrigues formula for the coquaternions.



**Remark 3.2.15** For a recent discussion of coquaternions in the study of complexified mechanics, see [BrGr2011].  $\square$



# 4

## ADJOINT AND COADJOINT ACTIONS

---

### Contents

4.1	Cayley–Klein dynamics for the rigid body	112
4.1.1	Cayley–Klein parameters, rigid-body dynamics	112
4.1.2	Body angular frequency	113
4.1.3	Cayley–Klein parameters	115
4.2	Actions of quaternions, Lie groups and Lie algebras	116
4.2.1	$AD$ , $Ad$ , $ad$ , $Ad^*$ and $ad^*$ actions of quaternions	117
4.2.2	$AD$ , $Ad$ , and $ad$ for Lie algebras and groups	118
4.3	Example: The Heisenberg Lie group	124
4.3.1	Definitions for the Heisenberg group	124
4.3.2	Adjoint actions: $AD$ , $Ad$ and $ad$	126
4.3.3	Coadjoint actions: $Ad^*$ and $ad^*$	127
4.3.4	Coadjoint motion and harmonic oscillations	129

---

## 4.1 Cayley–Klein dynamics for the rigid body

### 4.1.1 Cayley–Klein parameters, rigid-body dynamics

Recall that a time-dependent rotation by angle  $\theta(t)$  about direction  $\hat{\mathbf{n}}(t)$  is given by the unit quaternion (3.2.7) in Cayley–Klein parameters,

$$\hat{\mathbf{q}}(t) = [q_0(t), \mathbf{q}(t)] = \pm \left[ \cos \frac{\theta(t)}{2}, \sin \frac{\theta(t)}{2} \hat{\mathbf{n}}(t) \right]. \quad (4.1.1)$$

The *operation* of the unit quaternion  $\hat{\mathbf{q}}(t)$  on a vector  $\mathbf{X} = [0, \mathbf{X}]$  is given by quaternionic multiplication as

$$\begin{aligned} \hat{\mathbf{q}}(t)\mathbf{X} &= [q_0, \mathbf{q}][0, \mathbf{X}] = [-\mathbf{q} \cdot \mathbf{X}, q_0\mathbf{X} + \mathbf{q} \times \mathbf{X}] \\ &= \pm \left[ -\sin \frac{\theta(t)}{2} \hat{\mathbf{n}}(t) \cdot \mathbf{X}, \cos \frac{\theta(t)}{2} \mathbf{X} + \sin \frac{\theta(t)}{2} \hat{\mathbf{n}}(t) \times \mathbf{X} \right]. \end{aligned} \quad (4.1.2)$$

The corresponding time-dependent rotation is given by the Euler–Rodrigues formula (3.2.6) as

$$\mathbf{x}(t) = \hat{\mathbf{q}}(t)\mathbf{X}\hat{\mathbf{q}}^*(t) \quad \text{so that} \quad \mathbf{X} = \hat{\mathbf{q}}^*(t)\mathbf{x}(t)\hat{\mathbf{q}}(t) \quad (4.1.3)$$

in terms of the unit quaternion  $\hat{\mathbf{q}}(t)$ . Its time derivative is given by

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*\mathbf{x}\hat{\mathbf{q}}\hat{\mathbf{q}}^* + \hat{\mathbf{q}}\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*\mathbf{x}\hat{\mathbf{q}}\hat{\mathbf{q}}^* = \dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*\mathbf{x} + \mathbf{x}\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^* \\ &= \dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*\mathbf{x} + \mathbf{x}(\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*)^* = \dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*\mathbf{x} - \mathbf{x}(\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*) \\ &= \dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*\mathbf{x} - ((\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*)\mathbf{x})^* \\ &= 2\text{Im}((\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*)\mathbf{x}) \\ &= 2(\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*)\mathbf{x}. \end{aligned} \quad (4.1.4)$$

In quaternion components, this equation may be rewritten using

$$2\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^* = [0, \dot{\theta}\hat{\mathbf{n}} + \sin \theta \dot{\hat{\mathbf{n}}} + (1 - \cos \theta)\hat{\mathbf{n}} \times \dot{\hat{\mathbf{n}}}. \quad (4.1.5)$$

The quantity  $2\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^* = [0, 2(\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*)]$  is a pure quaternion whose vector component is denoted for the moment by enclosing it in parentheses as

$$2(\dot{\hat{\mathbf{q}}}\hat{\mathbf{q}}^*) = \dot{\theta}\hat{\mathbf{n}} + \sin \theta \dot{\hat{\mathbf{n}}} + (1 - \cos \theta)\hat{\mathbf{n}} \times \dot{\hat{\mathbf{n}}}.$$

## 4.1 CAYLEY–KLEIN DYNAMICS FOR THE RIGID BODY 113

As a consequence, the vector component of the quaternion equation (4.1.4) becomes

$$[0, \dot{\mathbf{x}}(t)] = 2[0, (\dot{\hat{q}}\hat{q}^*)][0, \mathbf{x}] = 2[0, (\dot{\hat{q}}\hat{q}^*) \times \mathbf{x}]. \quad (4.1.6)$$

### **Spatial angular frequency**

Upon recalling the isomorphism provided by the Euler–Rodrigues formula (3.2.6) for finite rotations,

$$\mathbf{x}(t) = O(t)\mathbf{X} = \hat{q}(t)\mathbf{X}\hat{q}^*(t), \quad (4.1.7)$$

the vector component of (4.1.6) yields a series of isomorphisms for the angular frequency,

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \dot{O}O^{-1}(t)\mathbf{x} \\ &= \hat{\omega}(t)\mathbf{x} \\ &= \boldsymbol{\omega}(t) \times \mathbf{x} \\ &= 2(\dot{\hat{q}}\hat{q}^*)(t) \times \mathbf{x}. \end{aligned} \quad (4.1.8)$$

Since the quaternion  $2\dot{\hat{q}}\hat{q}^*(t) = [0, 2(\dot{\hat{q}}\hat{q}^*)(t)]$  is equivalent to a vector in  $\mathbb{R}^3$ , we may simply use vector notation for it and equate the spatial angular frequencies as *vectors*. That is, we shall write

$$\boldsymbol{\omega}(t) = 2\dot{\hat{q}}\hat{q}^*(t), \quad (4.1.9)$$

and drop the parentheses  $(\cdot)$  when identifying pure quaternions with angular velocity vectors.

**Remark 4.1.1** Pure quaternions of the form  $\dot{\hat{q}}\hat{q}^*(t)$  may be identified with the tangent space of the *unit* quaternions at the identity.  $\square$

### **4.1.2 Body angular frequency**

The quaternion for the body angular frequency will have the corresponding vector expression,

$$\boldsymbol{\Omega}(t) = 2\hat{q}^*\dot{\hat{q}}(t). \quad (4.1.10)$$

## 114 4 : ADJOINT AND COADJOINT ACTIONS

Thus, only the vector parts enter the quaternionic descriptions of the spatial and body angular frequencies. The resulting isomorphisms are entirely sufficient to express the quaternionic versions of the rigid-body equations of motion in their Newtonian, Lagrangian and Hamiltonian forms, including the Lie–Poisson brackets. In particular, the kinetic energy for the rigid body is given by

$$K = \frac{1}{2} \boldsymbol{\Omega}(t) \cdot \mathbb{I} \boldsymbol{\Omega}(t) = 2 \left\langle \hat{\mathbf{q}}^* \dot{\hat{\mathbf{q}}}(t), \mathbb{I} \hat{\mathbf{q}}^* \dot{\hat{\mathbf{q}}}(t) \right\rangle. \quad (4.1.11)$$

So the quaternionic description of rigid-body dynamics reduces to the equivalent description in  $\mathbb{R}^3$ .

This equivalence in the two descriptions of rigid-body dynamics means that the relations for angular momentum, Hamilton’s principle and the Lie–Poisson brackets in terms of vector quantities all have identical expressions in the quaternionic picture. Likewise, the reconstruction of the Cayley–Klein parameters from the solution for the body angular velocity vector may be accomplished by integrating the linear quaternionic equation

$$\dot{\hat{\mathbf{q}}}(t) = \hat{\mathbf{q}} \boldsymbol{\Omega}(t) / 2, \quad (4.1.12)$$

or explicitly,

$$\frac{d}{dt} \left[ \cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}} \right] = \left[ \cos \frac{\theta}{2}, \sin \frac{\theta}{2} \hat{\mathbf{n}} \right] \left[ 0, \boldsymbol{\Omega}(t) / 2 \right]. \quad (4.1.13)$$

This is the linear reconstruction formula for the Cayley–Klein parameters.

**Remark 4.1.2** Expanding this linear equation for the Cayley–Klein parameters leads to a quaternionic equation for the Euler parameters that is linear in  $\boldsymbol{\Omega}(t)$ , but is nonlinear in  $\theta$  and  $\hat{\mathbf{n}}$ , namely

$$\left[ \dot{\theta}, \dot{\hat{\mathbf{n}}} \right] = \left[ 0, \hat{\mathbf{n}} \right] \left[ 0, \boldsymbol{\Omega}(t) + \boldsymbol{\Omega}(t) \times \hat{\mathbf{n}} \cot \frac{\theta}{2} \right]. \quad (4.1.14)$$

□

### 4.1.3 Cayley–Klein parameters

The series of isomorphisms in Equation (4.1.8) holds the key for rewriting Hamilton’s principle in Proposition 2.4.1 for Euler’s rigid-body equations in quaternionic form using Cayley–Klein parameters. The key step in proving Proposition 2.4.1 was deriving formula (2.4.7) for the variation of the body angular velocity. For this, one invokes equality of cross derivatives with respect to time  $t$  and variational parameter  $s$ . The hat map in that case then led to the key variational formula (2.4.7). The corresponding step for the quaternionic form of Hamilton’s principle in Cayley–Klein parameters also produces the key formula needed in this case.

**Proposition 4.1.1 (Cayley–Klein variational formula)** *The variation of the pure quaternion  $\Omega = 2\hat{q}^*\dot{\hat{q}}$  corresponding to body angular velocity in Cayley–Klein parameters satisfies the identity*

$$\Omega' - \dot{\Xi} = (\Omega \Xi - \Xi \Omega)/2 = \text{Im}(\Omega \Xi), \quad (4.1.15)$$

where  $\Xi := 2\hat{q}^*\hat{q}'$  and  $(\cdot)'$  denotes variation.

**Proof.** The body angular velocity is defined as  $\Omega = 2\hat{q}^*\dot{\hat{q}}$  in (4.1.10). Its variational derivative is found to be

$$\delta\Omega := \left. \frac{d}{ds}\Omega(s) \right|_{s=0} =: \Omega'. \quad (4.1.16)$$

Thus, the variation of  $\Omega$  may be expressed as

$$\Omega'/2 = (\hat{q}^*)'\dot{\hat{q}} + \hat{q}^*\dot{\hat{q}}'. \quad (4.1.17)$$

Now  $\epsilon = \hat{q}^*\hat{q}$  so that

$$\epsilon' = 0 = (\hat{q}^*)'\hat{q} + \hat{q}^*\hat{q}' \quad \text{and} \quad (\hat{q}^*)' = -\hat{q}^*\hat{q}'\hat{q}^*. \quad (4.1.18)$$

Hence, the variation of the angular frequency becomes

$$\begin{aligned} \delta\Omega/2 = \Omega'/2 &= -\hat{q}^*\hat{q}'\hat{q}^*\dot{\hat{q}} + \hat{q}^*\dot{\hat{q}}' \\ &= -\Xi\Omega/4 + \hat{q}^*\dot{\hat{q}}', \end{aligned} \quad (4.1.19)$$

## 116 4 : ADJOINT AND COADJOINT ACTIONS

where we have defined  $\Xi := 2\hat{q}^*\hat{q}'$ , which satisfies a similar relation,

$$\dot{\Xi}/2 = -\Omega \Xi/4 + \hat{q}^*\dot{\hat{q}}'. \quad (4.1.20)$$

Taking the difference of (4.1.19) and (4.1.20) yields

$$\Omega' - \dot{\Xi} = (\Omega \Xi - \Xi \Omega)/2 = \text{Im}(\Omega \Xi). \quad (4.1.21)$$

In quaternion components this formula becomes

$$\begin{aligned} [0, \Omega'] - [0, \dot{\Xi}] &= \frac{1}{2}([0, \Omega][0, \Xi] - [0, \Xi][0, \Omega]) \\ &= [0, \Omega \times \Xi], \end{aligned} \quad (4.1.22)$$

or, in the equivalent vector form,

$$\Omega' - \dot{\Xi} = \Omega \times \Xi. \quad (4.1.23)$$

This recovers the vector Equation (2.4.7), which was the key formula needed for writing Hamilton's principle in vector form, now reproduced in its pure quaternionic form for the Cayley–Klein parameters. ■

**Remark 4.1.3** Having expressed the key vector variational formula (2.4.7) in quaternionic form (4.1.21), the path for deriving Hamilton's principle for the rigid body in the quaternionic picture proceeds in parallel with the vector case. □

**Exercise.** State and prove Hamilton's principle for the rigid body in quaternionic form. ★

## 4.2 Actions of quaternions, Lie groups and Lie algebras

Quaternionic operations are isomorphic to the actions of Lie groups and Lie algebras. This isomorphism will allow us to develop Hamilton's principle for mechanics on Lie groups by following a path that parallels the one taken for quaternions.



### 4.2.1 AD, Ad, ad, Ad\* and ad\* actions of quaternions

We introduce the following notation for how the quaternions act among themselves and on the vectors in their left- and right-invariant tangent spaces at the identity. (These left- and right-invariant vectors are the body and space angular frequencies, respectively.)

- AD (conjugacy of quaternions),

$$\text{AD}_{\hat{q}} \tau := \hat{q} \tau \hat{q}^*,$$

- Ad (conjugacy of angular velocities),

$$\text{Ad}_{\hat{q}} \Omega = \hat{q} \Omega \hat{q}^* =: \omega,$$

- ad (commutator of angular velocities),

$$\text{ad}_{\Omega} \Xi = \text{Im}(\Omega \Xi) := (\Omega \Xi - \Xi \Omega)/2.$$

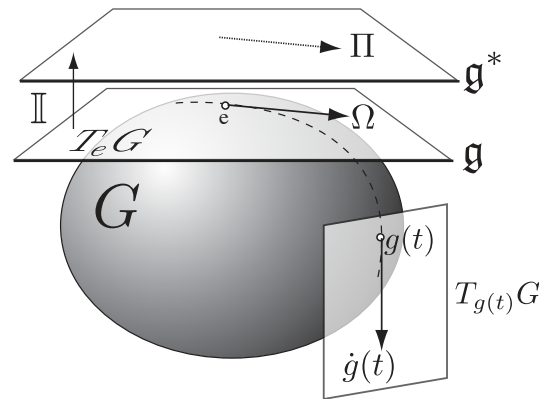
The pairing  $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \mapsto \mathbb{R}$  in formula (3.1.22) also allows one to define the corresponding dual operations. These are

- $\text{coAD} \langle \text{AD}_{\hat{q}}^* \mathfrak{s}, \tau \rangle = \langle \mathfrak{s}, \text{AD}_{\hat{q}} \tau \rangle,$
- $\text{coAd} \langle \text{Ad}_{\hat{q}}^* \Xi, \Omega \rangle = \langle \Xi, \text{Ad}_{\hat{q}} \Omega \rangle,$
- $\text{coad} \langle \text{ad}_{\Omega}^* \Upsilon, \Xi \rangle = \langle \Upsilon, \text{ad}_{\Omega} \Xi \rangle.$

**Exercise.** Prove that any pure quaternion is in the conjugacy class of  $[0, \hat{\mathbf{k}}]$  with  $\hat{\mathbf{k}} = (0, 0, 1)^T$  under the Ad action of a unit quaternion.

Hint: Compare with the formula in Remark 3.2.12. ★

## 118 4 : ADJOINT AND COADJOINT ACTIONS



**Figure 4.1.** The tangent space at the identity  $e$  of the group  $G$  is its Lie algebra  $\mathfrak{g}$ , a vector space represented here as a plane. The moment of inertia  $\mathbb{I}$  maps the vector  $\Omega \in \mathfrak{g}$  into the dual vector  $\Pi = \mathbb{I}\Omega \in \mathfrak{g}^*$ . The dual Lie algebra  $\mathfrak{g}^*$  is another vector space, also represented as a plane in the figure. A group orbit in  $G$  has tangent vector  $\dot{g}(t)$  at point  $g(t)$  which may be transported back to the identity by acting with  $g^{-1}(t) \in G$  from either the left as  $\Omega = g^{-1}(t)\dot{g}(t)$  or the right as  $\omega = \dot{g}(t)g^{-1}(t)$ .

#### 4.2.2 AD, Ad, and ad for Lie algebras and groups

The notation for the conjugacy relations among the quaternions in Section 4.2.1 follows the standard notation for the corresponding actions of a Lie group on itself, on its Lie algebra (its tangent space at the identity), the action of the Lie algebra on itself, and their dual actions. By the isomorphism between the quaternions and the matrix Lie group  $G = SU(2)$ , one may define these corresponding operations for other *matrix Lie groups*.

##### **ADjoint, Adjoint and adjoint for matrix Lie groups**

- AD (conjugacy classes of a matrix Lie group): The map  $I_g : G \rightarrow G$  given by  $I_g(h) \rightarrow ghg^{-1}$  for matrix Lie group elements  $g, h \in G$  is the *inner automorphism* associated with  $g$ . Orbits

of this action are called *conjugacy classes*.

$$\text{AD} : G \times G \rightarrow G : \quad \text{AD}_g h := ghg^{-1}.$$

- Differentiate  $I_g(h)$  with respect to  $h$  at  $h = e$  to produce the *Adjoint operation*,

$$\text{Ad} : G \times \mathfrak{g} \rightarrow \mathfrak{g} : \quad \text{Ad}_g \eta = T_e I_g \eta =: g\eta g^{-1},$$

with  $\eta = h'(0)$ .

- Differentiate  $\text{Ad}_g \eta$  with respect to  $g$  at  $g = e$  in the direction  $\xi$  to produce the *adjoint operation*,

$$\text{ad} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g} : \quad T_e(\text{Ad}_g \eta) \xi = [\xi, \eta] = \text{ad}_\xi \eta.$$

Explicitly, one computes the ad operation by differentiating the Ad operation directly as

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{g(t)} \eta &= \left. \frac{d}{dt} \right|_{t=0} \left( g(t)\eta g^{-1}(t) \right) \\ &= \dot{g}(0)\eta g^{-1}(0) - g(0)\eta g^{-1}(0)\dot{g}(0)g^{-1}(0) \\ &= \xi\eta - \eta\xi = [\xi, \eta] = \text{ad}_\xi \eta, \end{aligned} \quad (4.2.1)$$

where  $g(0) = Id$ ,  $\xi = \dot{g}(0)$  and the *Lie bracket*

$$[\xi, \eta] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g},$$

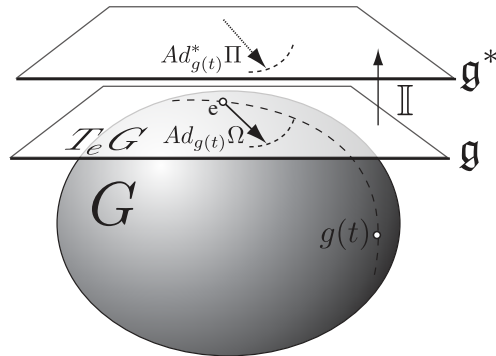
is the matrix commutator for a matrix Lie algebra.

**Remark 4.2.1 (Adjoint action)** Composition of the Adjoint action of  $G \times \mathfrak{g} \rightarrow \mathfrak{g}$  of a Lie group on its Lie algebra represents the group composition law as

$$\text{Ad}_g \text{Ad}_h \eta = g(h\eta h^{-1})g^{-1} = (gh)\eta(gh)^{-1} = \text{Ad}_{gh} \eta,$$

for any  $\eta \in \mathfrak{g}$ . □

## 120 4 : ADJOINT AND COADJOINT ACTIONS



**Figure 4.2.** The  $\text{Ad}$  and  $\text{Ad}^*$  operations of  $g(t)$  act, respectively, on the Lie algebra  $\text{Ad} : G \times \mathfrak{g} \rightarrow \mathfrak{g}$  and on its dual  $\text{Ad}^* : G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ .

**Exercise.** Verify that (note the minus sign)

$$\left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{g^{-1}(t)} \eta = -\text{ad}_\xi \eta,$$

for any fixed  $\eta \in \mathfrak{g}$ .



**Proposition 4.2.1 (Adjoint motion equation)** Let  $g(t)$  be a path in a Lie group  $G$  and  $\eta(t)$  be a path in its Lie algebra  $\mathfrak{g}$ . Then

$$\frac{d}{dt} \text{Ad}_{g(t)} \eta(t) = \text{Ad}_{g(t)} \left[ \frac{d\eta}{dt} + \text{ad}_{\xi(t)} \eta(t) \right],$$

where  $\xi(t) = g(t)^{-1} \dot{g}(t)$ .

**Proof.** By Equation (4.2.1), for a curve  $\eta(t) \in \mathfrak{g}$ ,

$$\begin{aligned} \frac{d}{dt}\Big|_{t=t_0} \text{Ad}_{g(t)} \eta(t) &= \frac{d}{dt}\Big|_{t=t_0} \left( g(t)\eta(t)g^{-1}(t) \right) \\ &= g(t_0) \left( \dot{\eta}(t_0) + g^{-1}(t_0)\dot{g}(t_0)\eta(t_0) \right. \\ &\quad \left. - \eta(t_0)g^{-1}(t_0)\dot{g}(t_0) \right) g^{-1}(t_0) \\ &= \left[ \text{Ad}_{g(t)} \left( \frac{d\eta}{dt} + \text{ad}_{\xi}\eta \right) \right]_{t=t_0}. \end{aligned} \quad (4.2.2)$$

■

**Exercise. (Inverse Adjoint motion relation)** Verify that

$$\frac{d}{dt} \text{Ad}_{g(t)^{-1}} \eta = -\text{ad}_{\xi} \text{Ad}_{g(t)^{-1}} \eta, \quad (4.2.3)$$

for any fixed  $\eta \in \mathfrak{g}$ . Note the placement of  $\text{Ad}_{g(t)^{-1}}$  and compare with Exercise on page 120. ★

**Compute the coAdjoint and coadjoint operations by taking duals**

The pairing

$$\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R} \quad (4.2.4)$$

(which is assumed to be nondegenerate) between a Lie algebra  $\mathfrak{g}$  and its dual vector space  $\mathfrak{g}^*$  allows one to define the following dual operations:

- The *coAdjoint operation* of a Lie group on the dual of its Lie algebra is defined by the pairing with the Ad operation,

$$\text{Ad}^* : G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^* : \langle \text{Ad}_g^* \mu, \eta \rangle := \langle \mu, \text{Ad}_g \eta \rangle, \quad (4.2.5)$$

for  $g \in G$ ,  $\mu \in \mathfrak{g}^*$  and  $\xi \in \mathfrak{g}$ .

## 122 4 : ADJOINT AND COADJOINT ACTIONS

- Likewise, the *coadjoint operation* is defined by the pairing with the ad operation,

$$\text{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \rightarrow \mathfrak{g}^* : \langle \text{ad}_\xi^* \mu, \eta \rangle := \langle \mu, \text{ad}_\xi \eta \rangle, \quad (4.2.6)$$

for  $\mu \in \mathfrak{g}^*$  and  $\xi, \eta \in \mathfrak{g}$ .

**Definition 4.2.1 (CoAdjoint action)** *The map*

$$\Phi^* : G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^* \quad \text{given by} \quad (g, \mu) \mapsto \text{Ad}_{g^{-1}}^* \mu \quad (4.2.7)$$

*defines the coAdjoint action of the Lie group  $G$  on its dual Lie algebra  $\mathfrak{g}^*$ .*

**Remark 4.2.2 (Coadjoint group action with  $g^{-1}$ )** Composition of coAdjoint operations with  $\Phi^*$  reverses the order in the group composition law as

$$\text{Ad}_g^* \text{Ad}_h^* = \text{Ad}_{hg}^*.$$

However, taking the inverse  $g^{-1}$  in Definition 4.2.1 of the coAdjoint action  $\Phi^*$  restores the order and thereby allows it to represent the group composition law when acting on the dual Lie algebra, for then

$$\text{Ad}_{g^{-1}}^* \text{Ad}_{h^{-1}}^* = \text{Ad}_{h^{-1}g^{-1}}^* = \text{Ad}_{(gh)^{-1}}^*. \quad (4.2.8)$$

(See [MaRa1994] for further discussion of this point.)  $\square$

The following proposition will be used later in the context of Euler–Poincaré reduction.

**Proposition 4.2.2 (Coadjoint motion relation)** *Let  $g(t)$  be a path in a Lie group  $G$  and  $\mu(t)$  be a path in  $\mathfrak{g}^*$ . The corresponding  $\text{Ad}^*$  operation satisfies*

$$\frac{d}{dt} \text{Ad}_{g(t)^{-1}}^* \mu(t) = \text{Ad}_{g(t)^{-1}}^* \left[ \frac{d\mu}{dt} - \text{ad}_{\xi(t)}^* \mu(t) \right], \quad (4.2.9)$$

where  $\xi(t) = g(t)^{-1} \dot{g}(t)$ .

**Proof.** The exercise on page 121 introduces the inverse Adjoint motion relation (4.2.3) for any fixed  $\eta \in \mathfrak{g}$ , repeated as

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

Relation (4.2.3) may be proven by the following computation,

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=t_0} \text{Ad}_{g(t)^{-1}} \eta &= \left. \frac{d}{dt} \right|_{t=t_0} \text{Ad}_{g(t)^{-1}g(t_0)} (\text{Ad}_{g(t_0)^{-1}} \eta) \\ &= -\text{ad}_{\xi(t_0)} (\text{Ad}_{g(t_0)^{-1}} \eta) , \end{aligned}$$

in which for the last step one recalls

$$\left. \frac{d}{dt} \right|_{t=t_0} g(t)^{-1}g(t_0) = (-g(t_0)^{-1}\dot{g}(t_0)g(t_0)^{-1})g(t_0) = -\xi(t_0) .$$

Relation (4.2.3) plays a key role in demonstrating relation (4.2.9) in the theorem, as follows. Using the pairing  $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$  between the Lie algebra and its dual, one computes

$$\begin{aligned} \left\langle \frac{d}{dt} \text{Ad}_{g(t)^{-1}}^* \mu(t), \eta \right\rangle &= \frac{d}{dt} \left\langle \text{Ad}_{g(t)^{-1}}^* \mu(t), \eta \right\rangle \\ &\text{by (4.2.5)} = \frac{d}{dt} \langle \mu(t), \text{Ad}_{g(t)^{-1}} \eta \rangle \\ &= \left\langle \frac{d\mu}{dt}, \text{Ad}_{g(t)^{-1}} \eta \right\rangle + \left\langle \mu(t), \frac{d}{dt} \text{Ad}_{g(t)^{-1}} \eta \right\rangle \\ &\text{by (4.2.3)} = \left\langle \frac{d\mu}{dt}, \text{Ad}_{g(t)^{-1}} \eta \right\rangle + \langle \mu(t), -\text{ad}_{\xi(t)} (\text{Ad}_{g(t)^{-1}} \eta) \rangle \\ &\text{by (4.2.6)} = \left\langle \frac{d\mu}{dt}, \text{Ad}_{g(t)^{-1}} \eta \right\rangle - \langle \text{ad}_{\xi(t)}^* \mu(t), \text{Ad}_{g(t)^{-1}} \eta \rangle \\ &\text{by (4.2.5)} = \left\langle \text{Ad}_{g(t)^{-1}}^* \frac{d\mu}{dt}, \eta \right\rangle - \langle \text{Ad}_{g(t)^{-1}}^* \text{ad}_{\xi(t)}^* \mu(t), \eta \rangle \\ &= \left\langle \text{Ad}_{g(t)^{-1}}^* \left[ \frac{d\mu}{dt} - \text{ad}_{\xi(t)}^* \mu(t) \right], \eta \right\rangle . \end{aligned}$$

This concludes the proof. ■

**Corollary 4.2.1** *The coadjoint orbit relation*

$$\mu(t) = \text{Ad}_{g(t)}^* \mu(0) \quad (4.2.10)$$

is the solution of the *coadjoint motion equation* for  $\mu(t)$ ,

$$\frac{d\mu}{dt} - \text{ad}_{\xi(t)}^* \mu(t) = 0. \quad (4.2.11)$$

**Proof.** Substituting Equation (4.2.11) into Equation (4.2.9) yields

$$\text{Ad}_{g(t)^{-1}}^* \mu(t) = \mu(0).$$

Operating on this equation with  $\text{Ad}_{g(t)}^*$  and recalling the composition rule for  $\text{Ad}^*$  from Remark 4.2.2 yields the result (4.2.10). ■

## 4.3 Example: The Heisenberg Lie group

### 4.3.1 Definitions for the Heisenberg group

The subset of the  $3 \times 3$  real matrices  $SL(3, \mathbb{R})$  given by the *upper triangular matrices*

$$\left\{ H = \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \mid a, b, c \in \mathbb{R} \right\} \quad (4.3.1)$$

defines a noncommutative group under matrix multiplication.

The  $3 \times 3$  matrix representation of this group acts on the *extended* planar vector  $(x, y, 1)^T$  as

$$\begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + ay + c \\ y + b \\ 1 \end{pmatrix}.$$



## 4.3 EXAMPLE: THE HEISENBERG LIE GROUP 125

The group  $H$  is called the Heisenberg group and it has three parameters. To begin studying its properties, consider the matrices in  $H$  given by

$$A = \begin{bmatrix} 1 & a_1 & a_3 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & b_1 & b_3 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.3.2)$$

The matrix product gives another element of  $H$ ,

$$AB = \begin{bmatrix} 1 & a_1 + b_1 & a_3 + b_3 + a_1 b_2 \\ 0 & 1 & a_2 + b_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.3.3)$$

and the inverses are

$$A^{-1} = \begin{bmatrix} 1 & -a_1 & a_1 a_2 - a_3 \\ 0 & 1 & -a_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & -b_1 & b_1 b_2 - b_3 \\ 0 & 1 & -b_2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.3.4)$$

We are dealing with a matrix (Lie) group. The *group commutator* is defined by

$$[A, B] := ABA^{-1}B^{-1} = \begin{bmatrix} 1 & 0 & a_1 b_2 - b_1 a_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.3.5)$$

Hence, the *commutator subgroup*  $\Gamma_1(H) = [H, H]$  has the form

$$\Gamma_1(H) = \{[A, B] : A, B \in H\} = \left\{ \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; k \in \mathbb{R} \right\}. \quad (4.3.6)$$

An element  $C$  of the commutator subgroup  $\Gamma_1(H)$  is of the form

$$C = \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \Gamma_1(H), \quad (4.3.7)$$

## 126 4 : ADJOINT AND COADJOINT ACTIONS

and we have the products

$$AC = \begin{bmatrix} 1 & a_1 & a_3 + k \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix} = CA. \quad (4.3.8)$$

Consequently,  $[A, C] = AC(CA)^{-1} = AC(AC)^{-1} = I_3$ . Hence, the subgroup of second commutators  $\Gamma_2(H) = [\Gamma_1(H), H]$  commutes with the rest of the group, which is thus *nilpotent of second order*.

### 4.3.2 Adjoint actions: AD, Ad and ad

Using the inverses in Equation (4.3.4) we compute the *group automorphism*

$$\text{AD}_B A = BAB^{-1} = \begin{bmatrix} 1 & a_1 & a_3 - a_1 b_2 + b_1 a_2 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.3.9)$$

Linearising the group automorphism  $\text{AD}_B A$  in  $A$  at the identity yields the Ad operation,

$$\begin{aligned} \text{Ad}_B \xi &= B \xi|_{\text{Id}} B^{-1} = \begin{bmatrix} 1 & b_1 & b_3 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -b_1 & b_1 b_2 - b_3 \\ 0 & 1 & -b_2 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & \xi_1 & \xi_3 + b_1 \xi_2 - b_2 \xi_1 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (4.3.10)$$

This is the Ad operation of the Heisenberg group  $H$  on its Lie algebra  $\mathfrak{h}(\mathbb{R}) \simeq \mathbb{R}^3$ :

$$\text{Ad} : H(\mathbb{R}) \times \mathfrak{h}(\mathbb{R}) \rightarrow \mathfrak{h}(\mathbb{R}). \quad (4.3.11)$$

One defines the right-invariant tangent vector,

$$\xi = \dot{A}A^{-1} = \begin{bmatrix} 0 & \dot{a}_1 & \dot{a}_3 - a_2 \dot{a}_1 \\ 0 & 0 & \dot{a}_2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{h}, \quad (4.3.12)$$

## 4.3 EXAMPLE: THE HEISENBERG LIE GROUP 127

and the left-invariant tangent vector,

$$\Xi = A^{-1}\dot{A} = \begin{bmatrix} 0 & \dot{a}_1 & \dot{a}_3 - a_1\dot{a}_2 \\ 0 & 0 & \dot{a}_2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \Xi_1 & \Xi_3 \\ 0 & 0 & \Xi_2 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{h}. \quad (4.3.13)$$

Next, we linearise  $\text{Ad}_B\xi$  in  $B$  around the identity to find the ad operation of the Heisenberg Lie algebra  $\mathfrak{h}$  on itself,

$$\text{ad} : \mathfrak{h} \times \mathfrak{h} \rightarrow \mathfrak{h}. \quad (4.3.14)$$

This is given explicitly by

$$\text{ad}_\eta\xi = [\eta, \xi] := \eta\xi - \xi\eta = \begin{bmatrix} 0 & 0 & \eta_1\xi_2 - \xi_1\eta_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (4.3.15)$$

Under the equivalence  $\mathfrak{h} \simeq \mathbb{R}^3$  provided by

$$\begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \mapsto \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} := \boldsymbol{\xi} \quad (4.3.16)$$

we may identify the Lie bracket with the projection onto the third component of the vector cross product:

$$[\eta, \xi] \mapsto \begin{bmatrix} 0 \\ 0 \\ \hat{\mathbf{3}} \cdot \boldsymbol{\eta} \times \boldsymbol{\xi} \end{bmatrix}. \quad (4.3.17)$$

### 4.3.3 Coadjoint actions: $\text{Ad}^*$ and $\text{ad}^*$

The inner product on the Heisenberg Lie algebra  $\mathfrak{h} \times \mathfrak{h} \rightarrow \mathbb{R}$  is defined by the matrix trace pairing

$$\langle \eta, \xi \rangle = \text{Tr}(\eta^T \xi) = \boldsymbol{\eta} \cdot \boldsymbol{\xi}. \quad (4.3.18)$$

## 128 4 : ADJOINT AND COADJOINT ACTIONS

Thus, elements of the dual Lie algebra  $\mathfrak{h}^*(\mathbb{R})$  may be represented as *lower triangular matrices*,

$$\mu = \begin{bmatrix} 0 & 0 & 0 \\ \mu_1 & 0 & 0 \\ \mu_3 & \mu_2 & 0 \end{bmatrix} \in \mathfrak{h}^*(\mathbb{R}). \quad (4.3.19)$$

The  $\text{Ad}^*$  operation of the Heisenberg group  $H(\mathbb{R})$  on its dual Lie algebra  $\mathfrak{h}^* \simeq \mathbb{R}^3$  is defined in terms of the matrix pairing by

$$\langle \text{Ad}_B^* \mu, \xi \rangle := \langle \mu, \text{Ad}_B \xi \rangle. \quad (4.3.20)$$

Explicitly, one may compute

$$\begin{aligned} \langle \mu, \text{Ad}_B \xi \rangle &= \text{Tr} \left( \begin{bmatrix} 0 & 0 & 0 \\ \mu_1 & 0 & 0 \\ \mu_3 & \mu_2 & 0 \end{bmatrix} \begin{bmatrix} 0 & \xi_1 & \xi_3 + b_1 \xi_2 - b_2 \xi_1 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \right) \\ &= \boldsymbol{\mu} \cdot \boldsymbol{\xi} + \mu_3(b_1 \xi_2 - b_2 \xi_1) \end{aligned} \quad (4.3.21)$$

$$\begin{aligned} &= \text{Tr} \left( \begin{bmatrix} 0 & 0 & 0 \\ \mu_1 - b_2 \mu_3 & 0 & 0 \\ \mu_3 & \mu_2 + b_1 \mu_3 & 0 \end{bmatrix} \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \right) \\ &= \langle \text{Ad}_B^* \mu, \xi \rangle. \end{aligned} \quad (4.3.22)$$

Thus, we have the formula for  $\text{Ad}_B^* \mu$ :

$$\text{Ad}_B^* \mu = \begin{bmatrix} 0 & 0 & 0 \\ \mu_1 - b_2 \mu_3 & 0 & 0 \\ \mu_3 & \mu_2 + b_1 \mu_3 & 0 \end{bmatrix}. \quad (4.3.23)$$

Likewise, the  $\text{ad}^*$  operation of the Heisenberg Lie algebra  $\mathfrak{h}$  on its dual  $\mathfrak{h}^*$  is defined in terms of the matrix pairing by

$$\langle \text{ad}_\eta^* \mu, \xi \rangle := \langle \mu, \text{ad}_\eta \xi \rangle \quad (4.3.24)$$

## 4.3 EXAMPLE: THE HEISENBERG LIE GROUP 129

$$\begin{aligned}
\langle \mu, \text{ad}_\eta \xi \rangle &= \text{Tr} \left( \begin{bmatrix} 0 & 0 & 0 \\ \mu_1 & 0 & 0 \\ \mu_3 & \mu_2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \eta_1 \xi_2 - \xi_1 \eta_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
&= \mu_3(\eta_1 \xi_2 - \eta_2 \xi_1) \tag{4.3.25}
\end{aligned}$$

$$\begin{aligned}
&= \text{Tr} \left( \begin{bmatrix} 0 & 0 & 0 \\ -\eta_2 \mu_3 & 0 & 0 \\ 0 & \eta_1 \mu_3 & 0 \end{bmatrix} \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
&= \langle \text{ad}_\eta^* \mu, \xi \rangle. \tag{4.3.26}
\end{aligned}$$

Thus, we have the formula for  $\text{ad}_\eta^* \mu$ :

$$\text{ad}_\eta^* \mu = \begin{bmatrix} 0 & 0 & 0 \\ -\eta_2 \mu_3 & 0 & 0 \\ 0 & \eta_1 \mu_3 & 0 \end{bmatrix}. \tag{4.3.27}$$

#### 4.3.4 Coadjoint motion and harmonic oscillations

According to Proposition 4.2.2, the coadjoint motion relation arises by differentiating along the coadjoint orbit. Let  $A(t)$  be a path in the Heisenberg Lie group  $H$  and  $\mu(t)$  be a path in  $\mathfrak{h}^*$ . Then we compute

$$\frac{d}{dt} \left( \text{Ad}_{A(t)}^* \mu(t) \right) = \text{Ad}_{A(t)}^* \left[ \frac{d\mu}{dt} - \text{ad}_{\eta(t)}^* \mu(t) \right], \tag{4.3.28}$$

where  $\eta(t) = A(t)^{-1} \dot{A}(t)$ .

With  $\eta = A^{-1} \dot{A}$ , Corollary 4.2.11 provides the differential equation for the coadjoint orbit,

$$\dot{\mu}(t) = \text{Ad}_{A(t)}^* \mu(0).$$

The desired differential equation is the *coadjoint motion equation*

$$\dot{\mu} = \text{ad}_\eta^* \mu,$$

## 130 4 : ADJOINT AND COADJOINT ACTIONS

which may be written for the Heisenberg Lie group  $H$  as

$$\dot{\mu} = \begin{bmatrix} 0 & 0 & 0 \\ \dot{\mu}_1 & 0 & 0 \\ \dot{\mu}_3 & \dot{\mu}_2 & 0 \end{bmatrix} = \text{ad}_{\eta}^* \mu = \begin{bmatrix} 0 & 0 & 0 \\ -\eta_2 \mu_3 & 0 & 0 \\ 0 & \eta_1 \mu_3 & 0 \end{bmatrix}. \quad (4.3.29)$$

That is,

$$\frac{d}{dt}(\mu_1, \mu_2, \mu_3) = (-\eta_2 \mu_3, \eta_1 \mu_3, 0). \quad (4.3.30)$$

Thus, the coadjoint motion equation for the Heisenberg group preserves the level sets of  $\mu_3$ .

If we define the linear map  $\mathfrak{h} \rightarrow \mathfrak{h}^* : (\mu_1, \mu_2) = (I_1 \eta_1, I_2 \eta_2)$  then the coadjoint motion equations become

$$\begin{aligned} \dot{\mu}_1 &= -\mu_3 \mu_2 / I_2, \\ \dot{\mu}_2 &= \mu_3 \mu_1 / I_1, \\ \dot{\mu}_3 &= 0. \end{aligned} \quad (4.3.31)$$

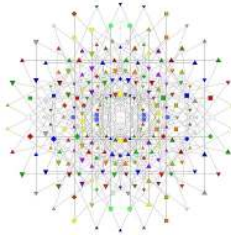
Upon taking another time derivative, this set reduces to the equations

$$\ddot{\mu}_k = -\frac{\mu_3^2}{I_1 I_2} \mu_k, \quad \text{for } k = 1, 2. \quad (4.3.32)$$

These are the equations for a planar isotropic harmonic oscillator on a level set of  $\mu_3$ .

This calculation has proved the following.

**Proposition 4.3.1** *Planar isotropic harmonic oscillations describe coadjoint orbits on the Heisenberg Lie group. The coadjoint orbits are  $(\mu_1, \mu_2)$  ellipses on level sets of  $\mu_3$ .*



# 5

## THE SPECIAL ORTHOGONAL GROUP $SO(3)$

---

### Contents

5.1	Adjoint and coadjoint actions of $SO(3)$	132
5.1.1	Ad and ad operations for the hat map	132
5.1.2	AD, Ad and ad actions of $SO(3)$	133
5.1.3	Dual Lie algebra isomorphism	135

---

## 5.1 Adjoint and coadjoint actions of $SO(3)$

Recall that the Lie group  $SO(3)$  of special orthogonal matrices is defined by

$$SO(3) := \{A \mid A \in 3 \times 3 \text{ orthogonal matrices, } \det(A) = 1\}.$$

The action of the matrix Lie group  $SO(3)$  on vectors in  $\mathbb{R}^3$  by left multiplication represents rotations in three dimensions. Its Lie algebra  $so(3)$  comprises the  $3 \times 3$  skew-symmetric matrices. Elements of  $so(3)$  represent angular velocities, and elements of its dual space  $so(3)^*$  under the matrix trace pairing represent angular momenta.

### 5.1.1 Ad and ad operations for the hat map

As shown in Theorem 2.1.1, the Lie algebra  $(so(3), [\cdot, \cdot])$  with matrix commutator bracket  $[\cdot, \cdot]$  maps to the Lie algebra  $(\mathbb{R}^3, \times)$  with vector product  $\times$ , by the linear isomorphism

$$\mathbf{u} := (u^1, u^2, u^3) \in \mathbb{R}^3 \mapsto \hat{\mathbf{u}} := \begin{bmatrix} 0 & -u^3 & u^2 \\ u^3 & 0 & -u^1 \\ -u^2 & u^1 & 0 \end{bmatrix} \in so(3).$$

In matrix and vector components, the linear isomorphism is

$$\hat{u}_{ij} := -\epsilon_{ijk} u^k.$$

Equivalently, this isomorphism is given by

$$\hat{\mathbf{u}}\mathbf{v} = \mathbf{u} \times \mathbf{v} \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathbb{R}^3.$$

This is the hat map  $\hat{\cdot} : (so(3), [\cdot, \cdot]) \rightarrow (\mathbb{R}^3, \times)$  defined earlier in (2.1.10) and (2.1.11) using

$$\hat{\mathbf{u}} = \mathbf{u} \cdot \hat{\mathbf{J}} = u^a \hat{J}_a,$$

which holds for the  $so(3)$  basis set (2.1.4) of skew-symmetric  $3 \times 3$  matrices  $\hat{J}_a$ , with  $a = 1, 2, 3$ .



5.1 ADJOINT AND COADJOINT ACTIONS OF  $SO(3)$  133

**Exercise.** Verify the following formulas for  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ :

$$\begin{aligned} (\mathbf{u} \times \mathbf{v})^\wedge &= \hat{u}\hat{v} - \hat{v}\hat{u} =: [\hat{u}, \hat{v}], \\ [\hat{u}, \hat{v}]\mathbf{w} &= (\mathbf{u} \times \mathbf{v}) \times \mathbf{w}, \\ ((\mathbf{u} \times \mathbf{v}) \times \mathbf{w})^\wedge &= [[\hat{u}, \hat{v}], \hat{w}], \\ \mathbf{u} \cdot \mathbf{v} &= -\frac{1}{2} \text{trace}(\hat{u}\hat{v}) \\ &=: \langle \hat{u}, \hat{v} \rangle, \end{aligned}$$

in which the dot product of vectors is also the natural pairing of  $3 \times 3$  skew-symmetric matrices. ★

**Exercise. (Jacobi identity under the hat map)** Verify that the Jacobi identity for the cross product of vectors in  $\mathbb{R}^3$  is equivalent to the Jacobi identity for the commutator product of  $3 \times 3$  skew matrices by proving the following identity satisfied by the hat map,

$$\begin{aligned} ((\mathbf{u} \times \mathbf{v}) \times \mathbf{w} + (\mathbf{v} \times \mathbf{w}) \times \mathbf{u} + (\mathbf{w} \times \mathbf{u}) \times \mathbf{v})^\wedge \\ = 0 = [[\hat{u}, \hat{v}], \hat{w}] + [[\hat{v}, \hat{w}], \hat{u}] + [[\hat{w}, \hat{u}], \hat{v}]. \end{aligned}$$

★

5.1.2 AD, Ad and ad actions of  $SO(3)$ 

- AD action of  $SO(3)$  on itself: The AD action for  $SO(3)$  is conjugation by matrix multiplication

$$I_A(B) = ABA^{-1}.$$

- Ad action of  $SO(3)$  on its Lie algebra  $so(3)$ : The corresponding adjoint action of  $SO(3)$  on  $so(3)$  may be obtained as follows.

## 134 5 : SPECIAL ORTHOGONAL GROUP

Differentiating  $B(t)$  at  $B(0) = Id$  gives

$$\text{Ad}_A \hat{v} = \left. \frac{d}{dt} \right|_{t=0} AB(t)A^{-1} = A\hat{v}A^{-1}, \quad \text{with } \hat{v} = B'(0).$$

One calculates the pairing with a vector  $\mathbf{w} \in \mathbb{R}^3$  as

$$\text{Ad}_A \hat{v}(\mathbf{w}) = A\hat{v}(A^{-1}\mathbf{w}) = A(\mathbf{v} \times A^{-1}\mathbf{w}) = A\mathbf{v} \times \mathbf{w} = (A\mathbf{v})^\wedge \mathbf{w},$$

where we have used the relation

$$A(\mathbf{u} \times \mathbf{v}) = A\mathbf{u} \times A\mathbf{v},$$

which holds for any  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$  and  $A \in SO(3)$ . Consequently,

$$\text{Ad}_A \hat{v} = (A\mathbf{v})^\wedge.$$

Identifying  $so(3) \simeq \mathbb{R}^3$  then gives

$$\text{Ad}_A \mathbf{v} = A\mathbf{v}.$$

So (speaking prose all our lives) the adjoint (Ad) action of the Lie group  $SO(3)$  on its Lie algebra  $so(3)$  may be identified with multiplication of a matrix in  $SO(3)$  times a vector in  $\mathbb{R}^3$ .

- ad action of  $so(3)$  on itself: Differentiating again gives the ad action of the Lie algebra  $so(3)$  on itself:

$$[\hat{u}, \hat{v}] = \text{ad}_{\hat{u}} \hat{v} = \left. \frac{d}{dt} \right|_{t=0} \left( e^{t\hat{u}} \mathbf{v} \right)^\wedge = (\hat{u}\mathbf{v})^\wedge = (\mathbf{u} \times \mathbf{v})^\wedge.$$

So the ad action of the Lie algebra  $so(3)$  on itself is by the matrix commutator of skew-symmetric matrices, which the hat map (isomorphism) identifies with the vector cross product.

- Infinitesimal generator: Likewise, the *infinitesimal generator* corresponding to  $\mathbf{u} \in \mathbb{R}^3$  has the expression

$$\mathbf{u}_{\mathbb{R}^3}(\mathbf{x}) := \left. \frac{d}{dt} \right|_{t=0} e^{t\hat{u}} \mathbf{x} = \hat{u} \mathbf{x} = \mathbf{u} \times \mathbf{x}.$$

### 5.1.3 Dual Lie algebra isomorphism $\check{\cdot} : so(3)^* \rightarrow \mathbb{R}^3$

**Proposition 5.1.1 (Coadjoint actions)** *The dual  $so(3)^*$  is identified with  $\mathbb{R}^3$  by the isomorphism*

$$\mathbf{\Pi} \in \mathbb{R}^3 \mapsto \check{\Pi} \in so(3)^* : \langle \check{\Pi}, \widehat{\mathbf{u}} \rangle := \mathbf{\Pi} \cdot \mathbf{u} \text{ for any } \mathbf{u} \in \mathbb{R}^3.$$

*In terms of this isomorphism, the coAdjoint action of  $SO(3)$  on  $so(3)^*$  is given by*

$$\text{Ad}_{A^{-1}}^* \check{\Pi} = (A\mathbf{\Pi})^\check{\cdot}, \quad (5.1.1)$$

*and the coadjoint action of  $so(3)$  on  $so(3)^*$  is given by*

$$\text{ad}_{\widehat{\mathbf{u}}}^* \check{\Pi} = (\mathbf{\Pi} \times \mathbf{u})^\check{\cdot}. \quad (5.1.2)$$

**Proof.**

- Computing the coAdjoint  $\text{Ad}^*$  action of  $SO(3)$  on  $so(3)^*$ : One computes from the definition,

$$\begin{aligned} \langle \text{Ad}_{A^{-1}}^* \check{\Pi}, \widehat{\mathbf{u}} \rangle &= \langle \check{\Pi}, \text{Ad}_{A^{-1}} \widehat{\mathbf{u}} \rangle = \langle \check{\Pi}, (A^{-1}\mathbf{u})^\wedge \rangle = \mathbf{\Pi} \cdot A^T \mathbf{u} \\ &= A\mathbf{\Pi} \cdot \mathbf{u} = \langle (A\mathbf{\Pi})^\check{\cdot}, \widehat{\mathbf{u}} \rangle. \end{aligned}$$

That is, the coAdjoint action of  $SO(3)$  on  $so(3)^*$  has the expression in (5.1.1),

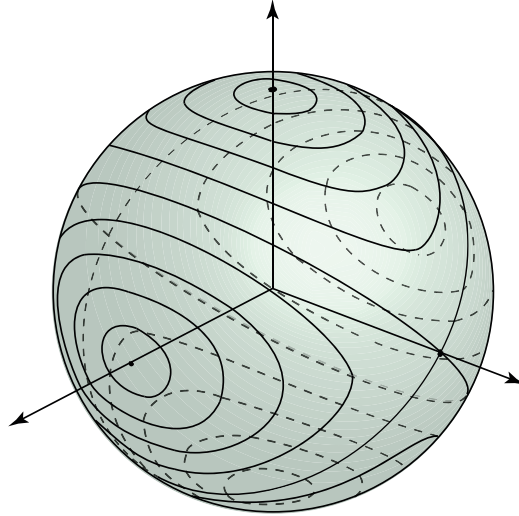
$$\text{Ad}_{A^{-1}}^* \check{\Pi} = (A\mathbf{\Pi})^\check{\cdot}.$$

- Computing the  $\text{ad}^*$  action of  $so(3)$  on its dual  $so(3)^*$ : Let  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$  and note that

$$\begin{aligned} \langle \text{ad}_{\widehat{\mathbf{u}}}^* \check{\Pi}, \widehat{\mathbf{v}} \rangle &= \langle \check{\Pi}, [\widehat{\mathbf{u}}, \widehat{\mathbf{v}}] \rangle = \langle \check{\Pi}, (\mathbf{u} \times \mathbf{v})^\wedge \rangle = \mathbf{\Pi} \cdot (\mathbf{u} \times \mathbf{v}) \\ &= (\mathbf{\Pi} \times \mathbf{u}) \cdot \mathbf{v} = \langle (\mathbf{\Pi} \times \mathbf{u})^\check{\cdot}, \widehat{\mathbf{v}} \rangle, \end{aligned}$$

which shows that  $\text{ad}_{\widehat{\mathbf{u}}}^* \check{\Pi} = (\mathbf{\Pi} \times \mathbf{u})^\check{\cdot}$ , thereby proving (5.1.2). ■

## 136 5 : SPECIAL ORTHOGONAL GROUP



**Figure 5.1.** A coAdjoint orbit of the action of  $SO(3)$  on  $so(3)^*$  is a sphere of radius  $|\mathbf{\Pi}|$ . The curves on this coAdjoint orbit are its intersections with the level sets of kinetic energy of a rigid body.

**Remark 5.1.1**

- The coAdjoint orbit

$$\mathcal{O} = \{A\mathbf{\Pi} \mid A \in SO(3)\} \subset \mathbb{R}^3$$

of  $SO(3)$  through  $\mathbf{\Pi} \in \mathbb{R}^3$  is a sphere  $S^2$  of radius  $|\mathbf{\Pi}|$  centred at the origin (Figure 5.1).

- The set  $\{\mathbf{\Pi} \times \mathbf{u} \mid \mathbf{u} \in \mathbb{R}^3\} = T_{\mathbf{\Pi}}\mathcal{O}$  is the plane perpendicular to  $\mathbf{\Pi}$ , i.e., the tangent space to the sphere  $S^2$  of radius  $|\mathbf{\Pi}|$  centred at the origin.

□

5.1 ADJOINT AND COADJOINT ACTIONS OF  $SO(3)$  137

**Exercise.** What are the analogues of the hat map

$$\hat{\cdot} : (so(3), \text{ad}) \rightarrow (\mathbb{R}^3, \times),$$

and its dual

$$\check{\cdot} : (so(3)^*, \text{ad}^*) \rightarrow (\mathbb{R}^3, \times),$$

- for the three-dimensional Lie algebras  $sp(2, \mathbb{R})$ ,  $so(2, 1)$ ,  $su(1, 1)$  and  $sl(2, \mathbb{R})$ ?
- for the six-dimensional Lie algebra  $so(4)$ ? ★

**Exercise.** Compute formula (4.2.9) from Proposition 4.2.2 for the matrix Lie group  $SO(3)$ . ★

**Answer.** Let  $A(t)$  be a path in the matrix Lie group  $SO(3)$  and  $\check{\Pi}(t)$  be a path in  $\mathfrak{g}^*$ . Then compute

$$\frac{d}{dt} \text{Ad}_{A(t)^{-1}}^* \check{\Pi}(t) = \text{Ad}_{A(t)^{-1}}^* \left[ \frac{d\check{\Pi}}{dt} - \text{ad}_{\hat{\Omega}}^* \check{\Pi}(t) \right], \quad (5.1.3)$$

where  $\hat{\Omega}(t) = A(t)^{-1} \dot{A}(t) \in so(3)$  is the body angular velocity.

$$\begin{aligned} \left\langle \frac{d}{dt} \text{Ad}_{A(t)^{-1}}^* \check{\Pi}(t), \hat{u} \right\rangle &= \frac{d}{dt} \left\langle \check{\Pi}(t), \text{Ad}_{A(t)^{-1}} \hat{u} \right\rangle \\ &= \frac{d}{dt} \left\langle \check{\Pi}(t), (A(t)^{-1} \mathbf{u})^\wedge \right\rangle \\ &= \frac{d}{dt} \left( \mathbf{\Pi}(t) \cdot A(t)^{-1} \mathbf{u} \right) \end{aligned}$$

## 138 5 : SPECIAL ORTHOGONAL GROUP

$$\begin{aligned}
&= \left( \dot{\mathbf{\Pi}} \cdot A(t)^{-1} \mathbf{u} \right) - \left( \mathbf{\Pi} \cdot A(t)^{-1} \dot{A}(t)^{-1} \mathbf{u} \right) \\
&= \left( \dot{\mathbf{\Pi}} \cdot A(t)^{-1} \mathbf{u} \right) - \left( \mathbf{\Pi} \cdot \widehat{\Omega} A(t)^{-1} \mathbf{u} \right) \\
&= \left( \dot{\mathbf{\Pi}} \cdot A(t)^{-1} \mathbf{u} \right) - \left( \mathbf{\Pi} \cdot \mathbf{\Omega} \times A(t)^{-1} \mathbf{u} \right) \\
&= \left( (\dot{\mathbf{\Pi}} + \mathbf{\Omega} \times \mathbf{\Pi}) \cdot A(t)^{-1} \mathbf{u} \right) \\
&= \left( A(t) (\dot{\mathbf{\Pi}} + \mathbf{\Omega} \times \mathbf{\Pi}) \right) \cdot \mathbf{u} \\
&= \left\langle \text{Ad}_{A(t)^{-1}}^* \left[ \frac{d\check{\mathbf{\Pi}}}{dt} - \text{ad}_{\check{\mathbf{\Omega}}}^* \check{\mathbf{\Pi}}(t) \right], \widehat{\mathbf{u}} \right\rangle.
\end{aligned}$$

▲

**Remark 5.1.2 (Spatial angular momentum conservation)** This computation provides a geometrical proof that the equation of motion for the *body angular momentum*  $\mathbf{\Pi}$ ,

$$\frac{d\mathbf{\Pi}}{dt} + \mathbf{\Omega} \times \mathbf{\Pi} = 0,$$

implies conservation of the *spatial angular momentum*  $\boldsymbol{\pi}$  given by

$$\boldsymbol{\pi}(t) = A(t)\mathbf{\Pi}(t). \quad (5.1.4)$$

Namely, the quantity  $\mathbf{\Pi}(t) \cdot A(t)^{-1} \mathbf{u}$  is conserved for any fixed vector  $\mathbf{u} \in \mathbb{R}^3$ , and

$$\mathbf{\Pi}(t) \cdot A(t)^{-1} \mathbf{u} = A(t)\mathbf{\Pi}(t) \cdot \mathbf{u} = \boldsymbol{\pi}(t) \cdot \mathbf{u}.$$

This conservation law makes perfect sense in the light of Lemma 2.1.4 for the rigid body, which summarises the two equivalent sets of equations in the spatial and body frames as

$$\frac{d\boldsymbol{\pi}}{dt} = 0 \quad \text{and} \quad \frac{d\mathbf{\Pi}}{dt} + (\mathbb{I}^{-1}\mathbf{\Pi}) \times \mathbf{\Pi} = 0. \quad (5.1.5)$$

5.1 ADJOINT AND COADJOINT ACTIONS OF  $SO(3)$  139

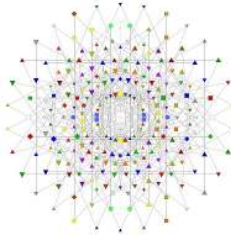
Since  $d\pi/dt = 0$ , we may invert the relation (5.1.4) as

$$\mathbf{\Pi}(t) = A(t)^{-1}\boldsymbol{\pi} = A(t)^{-1}\mathbf{\Pi}(0) \quad (5.1.6)$$

and interpret  $\boldsymbol{\pi}$  as the initial value of  $\mathbf{\Pi}(t)$  by setting  $A(0)^{-1} = Id$ . This interpretation is also consistent with the representation of rigid-body motion as a time-dependent curve  $A(t)$  in the Lie group  $SO(3)$ .  $\square$







# 6

## ADJOINT AND COADJOINT SEMIDIRECT-PRODUCT GROUP ACTIONS

---

### Contents

6.1	Special Euclidean group $SE(3)$	142
6.2	Adjoint operations for $SE(3)$	144
6.3	Adjoint actions of $SE(3)$ 's Lie algebra	148
6.3.1	The ad action of $se(3)$ on itself	148
6.3.2	The $\text{ad}^*$ action of $se(3)$ on its dual $se(3)^*$	149
6.3.3	Left versus right	151
6.4	Special Euclidean group $SE(2)$	153
6.5	Semidirect-product group $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$	156
6.5.1	Definitions for $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$	156
6.5.2	AD, Ad, and ad actions	158
6.5.3	$\text{Ad}^*$ and $\text{ad}^*$ actions	160
6.5.4	Coadjoint motion relation	162
6.6	Galilean group	164
6.6.1	Definitions for $G(3)$	164
6.6.2	AD, Ad, and ad actions of $G(3)$	165
6.7	Iterated semidirect products	167

---

## 6.1 Special Euclidean group $SE(3)$

As a set, the special Euclidean group in three dimensions  $SE(3)$  is the Cartesian product  $S = SO(3) \times \mathbb{R}^3$ . The group  $SE(3)$  acts on  $x \in \mathbb{R}^3$  by rotations and translations:  $x \rightarrow Rx + v$ , with  $R \in SO(3)$  and  $v \in \mathbb{R}^3$ . This group action may be represented by multiplication from the *left* of  $4 \times 4$  block matrices of the form

$$E(R, v) \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} Rx + v \\ 1 \end{bmatrix}.$$

Group multiplication in  $SE(3)$  may also be represented by  $4 \times 4$  *matrix multiplication*, as

$$\begin{aligned} E(\tilde{R}, \tilde{v})E(R, v) &= \begin{pmatrix} \tilde{R} & \tilde{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \tilde{R}R & \tilde{R}v + \tilde{v} \\ 0 & 1 \end{pmatrix} = E(\tilde{R}R, \tilde{R}v + \tilde{v}). \end{aligned}$$

One may abbreviate this multiplication by merely writing the top row as

$$(\tilde{R}, \tilde{v})(R, v) = (\tilde{R}R, \tilde{R}v + \tilde{v}).$$

The inverse of a group element  $(R, v)^{-1}$  is naturally identified with the matrix inverse as

$$E(R, v)^{-1} = \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}v \\ 0 & 1 \end{pmatrix},$$

so that

$$E(R, v)^{-1}E(R, v) = \begin{pmatrix} R^{-1} & -R^{-1}v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

or in the abbreviated top-row notation

$$(R, v)^{-1} = (R^{-1}, -R^{-1}v).$$

**Remark 6.1.1** More generally, these formulas for  $SE(3)$  are identified as the left action of a *semidirect-product Lie group*,

$$SE(3) \simeq SO(3) \ltimes \mathbb{R}^3.$$

Semidirect-product Lie groups were defined in Chapter 1. For more details, see, e.g., [MaRa1994]. Semidirect-product group multiplication is defined as follows for the case treated here, in which the normal subgroup is a vector space.  $\square$

**Definition 6.1.1 (Semidirect-product group action)** Suppose a Lie group  $G$  acts from the left by linear maps on a vector space  $V$ . (This will also induce a left action of  $G$  on the dual space  $V^*$ .) The *semidirect-product group*  $S = G \ltimes V$  is the Cartesian product of sets  $S = G \times V$  whose group multiplication is defined by

$$(g_1, v_1)(g_2, v_2) = (g_1g_2, v_1 + g_1v_2), \quad (6.1.1)$$

where the action of  $g \in G$  on  $v \in V$  is denoted simply as  $gv$  (on the left). The identity element is  $(e, 0)$  where  $e$  is the identity in  $G$ . The inverse of an element is given by

$$(g, v)^{-1} = (g^{-1}, -g^{-1}v). \quad (6.1.2)$$

**Remark 6.1.2** These formulas show that  $SE(3)$  is a semidirect-product Lie group.  $\square$

## 6.2 Adjoint operations for $SE(3)$

### AD operation

The AD operation  $AD : SE(3) \times SE(3) \mapsto SE(3)$  is conveniently expressed in the top-row notation as

$$\begin{aligned}
 AD_{(R, v)}(\tilde{R}, \tilde{v}) &= (R, v)(\tilde{R}, \tilde{v})(R, v)^{-1} \\
 &= (R, v)(\tilde{R}, \tilde{v})(R^{-1}, -R^{-1}v) \\
 &= (R, v)(\tilde{R}R^{-1}, \tilde{v} - \tilde{R}R^{-1}v) \\
 &= (R\tilde{R}R^{-1}, v + R\tilde{v} - R\tilde{R}R^{-1}v). \quad (6.2.1)
 \end{aligned}$$

**Remark 6.2.1** This formula has its counterpart for general semidirect products in which the normal subgroup is a vector space,

$$\begin{aligned}
 AD_{(g_1, v_1)}(g_2, v_2) &= (g_1, v_1)(g_2, v_2)(g_1, v_1)^{-1} \\
 &= (g_1g_2g_1^{-1}, v_1 + g_1v_2 - g_1g_2g_1^{-1}v_1).
 \end{aligned}$$

□

### Ad operation

Taking time derivatives of quantities adorned with the tilde ( $\tilde{\cdot}$ ) in formula (6.2.1) for  $AD_{(R, v)}(\tilde{R}(t), \tilde{v}(t))$  and evaluating at the identity  $t = 0$  yields

$$Ad_{(R, v)}(\dot{\tilde{R}}(0), \dot{\tilde{v}}(0)) = (Ad_R\dot{\tilde{R}}(0), -Ad_R\dot{\tilde{R}}(0)v + R\dot{\tilde{v}}(0)).$$

Setting  $\dot{\tilde{R}}(0) = \tilde{\xi}$  and  $\dot{\tilde{v}}(0) = \tilde{\alpha}$  defines the Ad action of  $SE(3)$  on its

Lie algebra with elements  $(\tilde{\xi}, \tilde{\alpha}) \in se(3)$  as  $\text{Ad} : SE(3) \times se(3) \rightarrow se(3)$ ,

$$\begin{aligned} \text{Ad}_{(R,v)}(\tilde{\xi}, \tilde{\alpha}) &= (\text{Ad}_R \tilde{\xi}, -\text{Ad}_R \tilde{\xi} v + R \tilde{\alpha}) \\ &= (R \tilde{\xi} R^{-1}, -R \tilde{\xi} R^{-1} v + R \tilde{\alpha}). \end{aligned} \quad (6.2.2)$$

**Remark 6.2.2** In vector form under the hat map this becomes

$$\text{Ad}_{(R,\mathbf{v})}(\tilde{\xi}, \tilde{\alpha}) = (R \tilde{\xi}, -R \tilde{\xi} \times \mathbf{v} + R \tilde{\alpha}). \quad (6.2.3)$$

□

**Remark 6.2.3** The Ad operation for left actions of a semidirect-product group  $S$  is given by

$$\text{Ad}_{(g,v)}(\xi, \alpha) = (g \xi g^{-1}, g \alpha - g \xi g^{-1} v).$$

By Equation (6.1.2) one then finds the Adjoint action of the inverse

$$\text{Ad}_{(g,v)^{-1}}(\xi, \alpha) = \text{Ad}_{(g^{-1}, -g^{-1}v)}(\xi, \alpha) = (g^{-1} \xi g, g^{-1} \alpha + g^{-1} \xi v),$$

where the left action of the Lie algebra  $\mathfrak{g}$  on  $V$  is denoted by concatenation, as in  $\xi v$ . □

### Ad\* operation

The pairing  $\langle \cdot, \cdot \rangle : se(3)^* \times se(3) \mapsto \mathbb{R}$  is obtained by identifying  $SE(3) \simeq SO(3) \times \mathbb{R}^3$  and taking the sum,

$$\langle (\mu, \beta), (\xi, \alpha) \rangle \equiv \frac{1}{2} \text{tr}(\mu^T \xi) + \beta \cdot \alpha, \quad (6.2.4)$$

with  $\mu \in so(3)^*$ ,  $\xi \in so(3)$  represented as skew-symmetric  $3 \times 3$  matrices and  $\beta, \alpha \in \mathbb{R}^3$  represented in usual vector notation. Thus, one computes the Ad\* operation as

$$\begin{aligned} \langle \text{Ad}_{(R,v)^{-1}}^*(\mu, \beta), (\xi, \alpha) \rangle &= \langle (\mu, \beta), \text{Ad}_{(R,v)^{-1}}(\xi, \alpha) \rangle \\ &= \langle (\mu, \beta), (R^{-1} \xi R, R^{-1} \alpha + R^{-1} \xi v) \rangle \\ &= \langle \mu, R^{-1} \xi R \rangle + \langle \beta, R^{-1} \alpha + R^{-1} \xi v \rangle \end{aligned}$$

## 146 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

$$\begin{aligned}
&= \langle R\mu R^{-1}, \xi \rangle + \langle R\beta, \alpha + \xi v \rangle \\
&= \langle R\mu R^{-1}, \xi \rangle + \langle R\beta, \alpha \rangle \\
&\quad + \langle \text{skew}(v \otimes R\beta), \xi \rangle \\
&= \langle R\mu R^{-1} + \text{skew}(v \otimes R\beta), \xi \rangle \\
&\quad + \langle R\beta, \alpha \rangle,
\end{aligned}$$

where  $\text{skew}(v \otimes R\beta)$  is the skew-symmetric part of  $v \otimes R\beta$ , which arises upon taking the trace of the product  $v \otimes R\beta$  with the skew-symmetric  $3 \times 3$  matrix  $\xi$ . One also uses the induced pairing

$$\langle \beta, R^{-1}\alpha \rangle = \langle R\beta, \alpha \rangle,$$

since  $\beta^T R^{-1}\alpha = (R\beta)^T \alpha$ .

**Remark 6.2.4** This computation expresses the  $\text{Ad}^*$  operation  $\text{Ad}^* : SE(3) \times se(3)^* \mapsto se(3)^*$  in its  $SO(3)$  and  $\mathbb{R}^3$  components explicitly as

$$\text{Ad}_{(R,v)^{-1}}^*(\mu, \beta) = \left( R\mu R^{-1} + \text{skew}(v \otimes R\beta), R\beta \right). \quad (6.2.5)$$

In vector form under the dual of the hat map the previous formula becomes

$$\text{Ad}_{(R,v)^{-1}}^*(\boldsymbol{\mu}, \boldsymbol{\beta}) = (R\boldsymbol{\mu} + \mathbf{v} \times R\boldsymbol{\beta}, R\boldsymbol{\beta}). \quad (6.2.6)$$

□

**Remark 6.2.5 (Semidirect-product  $\text{Ad}$ ,  $\text{Ad}^*$  actions)** Upon denoting the various group and algebra actions by concatenation from the left, the Adjoint and coAdjoint actions for semidirect products may be expressed in slightly simpler form as (see, e.g., [MaRa1994])

$$(g, v)(\xi, \alpha) = (g\xi, g\alpha - (g\xi)v) \quad (6.2.7)$$

and

$$(g, v)(\mu, \beta) = (g\mu + v \diamond (g\beta), g\beta), \quad (6.2.8)$$

where  $(g, v) \in S = G \times V$ ,  $(\xi, \alpha) \in \mathfrak{s} = \mathfrak{g} \times V$ ,  $(\mu, \beta) \in \mathfrak{s}^* = \mathfrak{g}^* \times V^*$ ,  $g\xi = \text{Ad}_g \xi$ ,  $g\mu = \text{Ad}_{g^{-1}}^* \mu$ ,  $g\beta$  denotes the induced left action of  $g$  on  $\beta$  (the left action of  $G$  on  $V$  induces a left action of  $G$  on  $V^*$  – the inverse of the transpose of the action on  $V$ ).

□

**Definition 6.2.1 (The diamond operation  $\diamond$ )** The diamond operation  $\diamond$  that appears in Equation (6.2.8) is defined by

$$\langle v \diamond (g\beta), \xi \rangle = - \langle (g\beta) \diamond v, \xi \rangle = \langle (g\beta), \xi v \rangle. \quad (6.2.9)$$

That is, the diamond operation minus the dual of the (left) Lie algebra action. In the present case for  $SE(3)$  one has

$$v \diamond (R\beta) = \text{skew}(v \otimes R\beta) = (v \times R\beta)^\wedge,$$

by the hat map, as in Equation (6.2.6).

**Exercise. (Coadjoint action  $\text{Ad}_{g^{-1}}^*$  of semidirect-product groups)** Show that the map  $\text{Ad}_{(g,v)^{-1}}^*$  preserves the action

$$(g_1, v_1)(g_2, v_2) = (g_1 g_2, v_1 + g_1 v_2)$$

of a semidirect-product group by computing the composition,

$$\text{Ad}_{(g_1, v_1)^{-1}}^* \text{Ad}_{(g_2, v_2)^{-1}}^* = \text{Ad}_{(g_1 g_2, v_1 + g_1 v_2)^{-1}}^* .$$

★

## 6.3 Adjoint actions of $SE(3)$ 's Lie algebra

### 6.3.1 The ad action of $se(3)$ on itself

To express the operation  $\text{ad}^* : se(3)^* \times se(3) \rightarrow se(3)^*$ , one begins by computing its corresponding ad operation,  $\text{ad} : se(3) \times se(3) \rightarrow se(3)$ . This is done by taking time derivatives of *unadorned* quantities of  $\text{Ad}_{(R(t), v(t))}(\tilde{\xi}, \tilde{\alpha})$  in Equation (6.2.2) evaluated at the identity to find

$$\begin{aligned} \text{ad}_{(\dot{R}(0), \dot{v}(0))}(\tilde{\xi}, \tilde{\alpha}) &= \left( \dot{R}\tilde{\xi}R^{-1} - R\tilde{\xi}R^{-1}\dot{R}R^{-1}, \right. \\ &\quad \left. - \dot{R}\tilde{\xi}R^{-1}v + R\tilde{\xi}R^{-1}\dot{R}R^{-1}v - R\tilde{\xi}R^{-1}\dot{v} + \dot{R}\tilde{\alpha} \right) \Big|_{\text{Id}}. \end{aligned}$$

As before, one sets  $\dot{R}(0) = \xi$ ,  $\dot{v}(0) = \alpha$ ,  $R(0) = \text{Id}$  and  $v(0) = 0$ . In this notation, the ad operation  $\text{ad}_{(\xi, \alpha)}$  for the right-invariant Lie algebra action of  $se(3)$  may thus be rewritten as

$$\begin{aligned} \text{ad}_{(\xi, \alpha)}(\tilde{\xi}, \tilde{\alpha}) &= (\xi\tilde{\xi} - \tilde{\xi}\xi, -(\xi\tilde{\xi} - \tilde{\xi}\xi)v - \tilde{\xi}(\xi v + \alpha) + \xi\tilde{\alpha}) \Big|_{\text{Id}} \\ &= ([\xi, \tilde{\xi}], -\xi\tilde{\xi}v + \xi\tilde{\alpha} - \tilde{\xi}\alpha) \Big|_{\text{Id}} \\ &= (\text{ad}_\xi \tilde{\xi}, \xi\tilde{\alpha} - \tilde{\xi}\alpha), \end{aligned}$$

where the last step uses  $v(0) = 0$ . The result is just the matrix commutator,

$$\text{ad}_{(\xi, \alpha)}(\tilde{\xi}, \tilde{\alpha}) = \left[ \begin{pmatrix} \xi & \alpha \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \tilde{\xi} & \tilde{\alpha} \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} [\xi, \tilde{\xi}] & \xi\tilde{\alpha} - \tilde{\xi}\alpha \\ 0 & 0 \end{pmatrix}.$$

**Remark 6.3.1 (The semidirect-product Lie bracket)** The (left) Lie algebra of the semidirect-product Lie group  $S$  is the



semidirect-product Lie algebra,  $\mathfrak{s} = \mathfrak{g} \ltimes V$ , whose Lie bracket is expressed as

$$[(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], \xi_1 v_2 - \xi_2 v_1), \quad (6.3.1)$$

where the induced action of  $\mathfrak{g}$  on  $V$  is denoted by concatenation, as in  $\xi_1 v_2$ .  $\square$

**Remark 6.3.2** In vector notation, using the hat map  $\hat{\cdot} : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$  given by  $(\hat{v})_{ij} = -\epsilon_{ijk} v_k$  so that  $\hat{v} \mathbf{w} = \mathbf{v} \times \mathbf{w}$  and  $[\hat{v}, \hat{w}] = (\mathbf{v} \times \mathbf{w})^\wedge$ , one has  $[\xi, \tilde{\xi}] = (\xi \times \tilde{\xi})^\wedge$  and finds the correspondence (isomorphism)

$$\begin{aligned} \text{ad}_{(\xi, \alpha)}(\tilde{\xi}, \tilde{\alpha}) &= [(\xi, \alpha), (\tilde{\xi}, \tilde{\alpha})] \\ &= ([\xi, \tilde{\xi}], \xi \tilde{\alpha} - \tilde{\xi} \alpha) \\ &= \left( (\xi \times \tilde{\xi})^\wedge, (\xi \times \tilde{\alpha} - \tilde{\xi} \times \alpha) \right). \end{aligned}$$

This expression will be useful in interpreting the  $\text{ad}$  and  $\text{ad}^*$  actions as motion on  $\mathbb{R}^3$ .  $\square$

### 6.3.2 The $\text{ad}^*$ action of $se(3)$ on its dual $se(3)^*$

One computes the  $\text{ad}^*$  action of  $se(3)$  on its dual  $se(3)^*$  by using the pairing,

$$\begin{aligned} \langle \text{ad}_{(\xi, \alpha)}^*(\mu, \beta), (\tilde{\xi}, \tilde{\alpha}) \rangle &= \langle (\mu, \beta), \text{ad}_{(\xi, \alpha)}(\tilde{\xi}, \tilde{\alpha}) \rangle \\ &= \langle (\mu, \beta), (\text{ad}_\xi \tilde{\xi}, \xi \tilde{\alpha} - \tilde{\xi} \alpha) \rangle \quad (6.3.2) \\ &= \langle \mu, \text{ad}_\xi \tilde{\xi} \rangle + \langle \beta, \xi \tilde{\alpha} \rangle - \langle \beta, \tilde{\xi} \alpha \rangle \end{aligned}$$

## 150 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

$$\begin{aligned}
&= \langle \text{ad}_\xi^* \mu, \tilde{\xi} \rangle + \langle -\xi\beta, \tilde{\alpha} \rangle + \langle \beta \diamond \alpha, \tilde{\xi} \rangle \\
&= \langle (\text{ad}_\xi^* \mu + \beta \diamond \alpha, -\xi\beta), (\tilde{\xi}, \tilde{\alpha}) \rangle.
\end{aligned}$$

**Remark 6.3.3** Again the diamond operation

$$\diamond : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathfrak{so}(3)^* \simeq \mathbb{R}^3$$

arises, as defined in Equation (6.2.9) by the dual Lie algebra actions,

$$\langle \beta \diamond \alpha, \tilde{\xi} \rangle = -\langle \beta, \tilde{\xi}\alpha \rangle.$$

In vector notation, this becomes

$$-\langle \beta, \tilde{\xi}\alpha \rangle = -\beta \cdot \tilde{\xi} \times \alpha = \beta \times \alpha \cdot \tilde{\xi} = \langle \beta \diamond \alpha, \tilde{\xi} \rangle.$$

Thus, the diamond operation for  $\mathfrak{se}(3)$  is simply the cross product of vectors in  $\mathbb{R}^3$ .  $\square$

Under the hat map, the pairing  $\langle \cdot, \cdot \rangle : \mathfrak{se}(3)^* \times \mathfrak{se}(3) \rightarrow \mathbb{R}$  in (6.2.4) transforms into the dot product of vectors in  $\mathbb{R}^3$ ,

$$\langle (\mu, \beta), (\xi, \alpha) \rangle = \mu \cdot \xi + \beta \cdot \alpha.$$

Thus, the  $\text{ad}^*$  action of  $\mathfrak{se}(3)$  may be expressed in terms of vector operations, as

$$\begin{aligned}
&\langle \text{ad}_{(\xi, \alpha)}^*(\mu, \beta), (\tilde{\xi}, \tilde{\alpha}) \rangle \\
&= \langle (\mu, \beta), \text{ad}_{(\xi, \alpha)}(\tilde{\xi}, \tilde{\alpha}) \rangle \\
&= \mu \cdot (\xi \times \tilde{\xi}) + \beta \cdot (\xi \times \tilde{\alpha} - \tilde{\xi} \times \alpha) \\
&= (\mu \times \xi - \alpha \times \beta) \cdot \tilde{\xi} - \xi \times \beta \cdot \tilde{\alpha}. \quad (6.3.3)
\end{aligned}$$

**Exercise.** Find the conservation laws for the equation of coadjoint motion

$$\frac{d}{dt}(\boldsymbol{\mu}, \boldsymbol{\beta}) = \text{ad}_{(\boldsymbol{\xi}, \boldsymbol{\alpha})}^*(\boldsymbol{\mu}, \boldsymbol{\beta})$$

on  $se(3)^*$ , the dual of the special Euclidean Lie algebra in three dimensions, when  $(\boldsymbol{\mu}, \boldsymbol{\beta})$  are linearly related to  $(\boldsymbol{\xi}, \boldsymbol{\alpha})$  by  $(\boldsymbol{\mu}, \boldsymbol{\beta}) = (\mathbf{I}\boldsymbol{\xi}, \mathbf{K}\boldsymbol{\alpha})$ , for symmetric matrices  $(\mathbf{I}, \mathbf{K})$ . ★

**Summary.** The adjoint and coadjoint actions for  $SE(3) \simeq SO(3) \circledast \mathbb{R}^3$  are

$$\begin{aligned} \text{AD}_{(R,v)}(\tilde{R}, \tilde{v}) &= (R\tilde{R}R^{-1}, v + R\tilde{v} - R\tilde{R}R^{-1}v), \\ \text{Ad}_{(R,v)^{-1}}(\boldsymbol{\xi}, \boldsymbol{\alpha}) &= (R^{-1}\boldsymbol{\xi}R, R^{-1}\boldsymbol{\alpha} + R^{-1}\boldsymbol{\xi}v), \\ \text{ad}_{(\boldsymbol{\xi}, \boldsymbol{\alpha})}(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\alpha}}) &= \left( (\boldsymbol{\xi} \times \tilde{\boldsymbol{\xi}})^\wedge, (\boldsymbol{\xi} \times \tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\xi}} \times \boldsymbol{\alpha}) \right), \\ \text{Ad}_{(R,v)^{-1}}^*(\boldsymbol{\mu}, \boldsymbol{\beta}) &= (R\boldsymbol{\mu} + \mathbf{v} \times R\boldsymbol{\beta}, R\boldsymbol{\beta}), \\ \text{ad}_{(\boldsymbol{\xi}, \boldsymbol{\alpha})}^*(\boldsymbol{\mu}, \boldsymbol{\beta}) &= \left( \boldsymbol{\mu} \times \boldsymbol{\xi} - \boldsymbol{\alpha} \times \boldsymbol{\beta}, -\boldsymbol{\xi} \times \boldsymbol{\beta} \right). \end{aligned}$$

### 6.3.3 Left versus right

When working with various models of continuum mechanics and plasmas it is convenient to work with *right* representations of  $G$  on the vector space  $V$  (as in, for example, [HoMaRa1998]). We shall denote the semidirect product by the same symbol,  $S = G \circledast V$ , the action of  $G$  on  $V$  being denoted by  $vg$ . The formulas change under these conventions as follows. Group multiplication (the analogue of (6.1.1)) is given by

$$(g_1, v_1)(g_2, v_2) = (g_1g_2, v_2 + v_1g_2), \quad (6.3.4)$$

## 152 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

and the Lie algebra bracket on  $\mathfrak{s} = \mathfrak{g} \oplus V$  (the analogue of (6.3.1)) has the expression

$$[(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], v_1\xi_2 - v_2\xi_1), \quad (6.3.5)$$

where we denote the induced action of  $\mathfrak{g}$  on  $V$  by concatenation, as in  $v_1\xi_2$ . The adjoint and coadjoint actions have the formulas (analogues of (6.2.7) and (6.2.8))

$$(g, v)(\xi, u) = (g\xi, (u + v\xi)g^{-1}), \quad (6.3.6)$$

$$(g, v)(\mu, a) = (g\mu + (vg^{-1}) \diamond (ag^{-1}), ag^{-1}), \quad (6.3.7)$$

where, as usual,  $g\xi = \text{Ad}_g\xi$ ,  $g\mu = \text{Ad}_{g^{-1}}^*\mu$ ,  $ag$  denotes the inverse of the dual isomorphism defined by  $g \in G$  (so that  $g \mapsto ag$  is a *right* action). Note that the adjoint and coadjoint actions are *left* actions. In this case, the  $\mathfrak{g}$ -actions on  $\mathfrak{g}^*$  and  $V^*$  are defined as before to be minus the dual map given by the  $\mathfrak{g}$ -actions on  $\mathfrak{g}$  and  $V$  and are denoted, respectively, by  $\xi\mu$  (because it is a left action) and  $a\xi$  (because it is a right action).

**Left-invariant tangent vectors**

The left-invariant tangent vectors to  $(R, v)$  at the identity  $(\text{Id}, 0)$  are given by

$$(\xi, \alpha) = (R, v)^{-1}(\dot{R}, \dot{v}), \quad (6.3.8)$$

or, in matrix form,

$$\begin{aligned} \begin{pmatrix} \xi & \alpha \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} R^{-1} & -R^{-1}v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \dot{R} & \dot{v} \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} R^{-1}\dot{R} & R^{-1}\dot{v} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \xi & R^{-1}\dot{v} \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

This gives the *reconstruction formula* for left-invariant tangent vectors,

$$(\dot{R}, \dot{v}) = (R, v)(\xi, \alpha) = (R\xi, R\alpha), \quad (6.3.9)$$

which may also be expressed in matrix form as

$$\begin{pmatrix} \dot{R} & \dot{v} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \xi & \alpha \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R\xi & R\alpha \\ 0 & 0 \end{pmatrix}.$$

**Exercise. (Right-invariant tangent vectors)** The right-invariant tangent vectors to  $(R, v)$  at the identity  $(\text{Id}, 0)$  are given by

$$(\xi, \alpha) = (\dot{R}, \dot{v})(R, v)^{-1} = (\dot{R}R^{-1}, -\dot{R}R^{-1}v + \dot{v}). \quad (6.3.10)$$

Show that the reconstruction formula for right-invariant tangent vectors is given by

$$(\dot{R}, \dot{v}) = (\xi, \alpha)(R, v) = (\xi R, \xi v + \alpha). \quad (6.3.11) \quad \star$$

**Remark 6.3.4 (Right- vs left-invariant reconstructions)** The reconstruction formulas for right-invariant (6.3.11) and left-invariant (6.3.9) tangent vectors in  $SE(3)$  are completely different. In particular, right-invariant reconstruction involves translations, while left-invariant reconstruction does not.  $\square$

## 6.4 Special Euclidean group $SE(2)$

The special Euclidean group of the plane  $SE(2) \simeq SO(2) \ltimes \mathbb{R}^2$  has coordinates

$$(R_\theta, v) = \begin{pmatrix} R_\theta & v \\ 0 & 1 \end{pmatrix},$$

where  $v \in \mathbb{R}^2$  is a vector in the plane and  $R_\theta$  is the rotation matrix

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

## 154 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

**Exercise.** Calculate the inverse of  $(R_\theta, v)$  and show that the Lie algebra  $se(2)$  of  $SE(2)$  consists of the  $3 \times 3$  block matrices of the form

$$\begin{pmatrix} -\xi\mathbb{J} & \alpha \\ 0 & 0 \end{pmatrix}, \quad \text{where } \mathbb{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

(The skew-symmetric  $2 \times 2$  matrix  $\mathbb{J} = -\mathbb{J}^T = -\mathbb{J}^{-1}$  represents rotation by  $-\pi/2$ .) ★

**Exercise.** Identify the Lie algebra  $se(2)$  with  $\mathbb{R}^3$  via the isomorphism

$$\begin{pmatrix} -\xi\mathbb{J} & \alpha \\ 0 & 0 \end{pmatrix} \in se(2) \mapsto (\xi, \alpha) \in \mathbb{R}^3,$$

and compute the expression for the Lie algebra bracket as

$$\begin{aligned} [(\xi, \alpha_1, \alpha_2), (\tilde{\xi}, \tilde{\alpha}_1, \tilde{\alpha}_2)] &= (0, -\xi\tilde{\alpha}_2 + \tilde{\xi}\alpha_2, \xi\tilde{\alpha}_1 - \tilde{\xi}\alpha_1) \\ &= (0, -\xi\mathbb{J}\tilde{\alpha} + \tilde{\xi}\mathbb{J}\alpha), \end{aligned}$$

where  $\alpha = (\alpha_1, \alpha_2)$  and  $\tilde{\alpha} = (\tilde{\alpha}_1, \tilde{\alpha}_2)$ . ★

**Exercise.** Check using  $R_\theta\mathbb{J} = \mathbb{J}R_\theta$  that the adjoint action for  $SE(2)$  of

$$(R_\theta, v) = \begin{pmatrix} R_\theta & v \\ 0 & 1 \end{pmatrix} \quad \text{on} \quad (\xi, \alpha) = \begin{pmatrix} -\xi\mathbb{J} & \alpha \\ 0 & 0 \end{pmatrix}$$

is given by

$$(R_\theta, v)(\xi, \alpha)(R_\theta, v)^{-1} = \begin{pmatrix} -\xi\mathbb{J} & \xi\mathbb{J}v + R_\theta\alpha \\ 0 & 0 \end{pmatrix},$$

or in coordinates

$$\text{Ad}_{(R_\theta, v)}(\xi, \alpha) = (\xi, \xi \mathbb{J}v + R_\theta \alpha). \quad (6.4.1)$$

★

The pairing  $\langle \cdot, \cdot \rangle : se(2)^* \times se(2) \mapsto \mathbb{R}$  is obtained by identifying  $SE(2) \simeq SO(2) \times \mathbb{R}^2$  and taking the sum,

$$\langle (\mu, \beta), (\xi, \alpha) \rangle \equiv \mu\xi + \beta \cdot \alpha, \quad (6.4.2)$$

with  $\mu \in so(2)^*$ ,  $\xi \in so(2)$  represented as skew-symmetric  $2 \times 2$  matrices and  $\beta, \alpha \in \mathbb{R}^2$  represented as planar vectors. Elements of the dual Lie algebra  $se(2)^*$  may be written as block matrices of the form

$$(\mu, \beta) = \begin{pmatrix} \frac{\mu}{2} \mathbb{J} & 0 \\ \beta & 0 \end{pmatrix},$$

since

$$\text{tr} \left[ \begin{pmatrix} \frac{\mu}{2} \mathbb{J} & 0 \\ \beta & 0 \end{pmatrix} \begin{pmatrix} -\xi \mathbb{J} & \alpha \\ 0 & 0 \end{pmatrix} \right] = \mu\xi + \beta \cdot \alpha,$$

via the nondegenerate pairing provided by the trace of the matrix product. Thus, we may identify the dual Lie algebra  $se(2)^*$  with  $\mathbb{R}^3$  via the isomorphism

$$\begin{pmatrix} \frac{\mu}{2} \mathbb{J} & 0 \\ \beta & 0 \end{pmatrix} \in se(2)^* \mapsto (\mu, \beta) \in \mathbb{R}^3,$$

so that in these coordinates the pairing (6.4.2) between  $se(2)^*$  and  $se(2)$  becomes the usual dot product in  $\mathbb{R}^3$ .

**Exercise.** Check that the coadjoint action of  $SE(2)$  on  $se(2)^*$  is given by

$$\text{Ad}_{(R_\theta, v)}^*(\mu, \beta) = (\mu - R_\theta \beta \cdot \mathbb{J}v, R_\theta \beta). \quad (6.4.3)$$

★

**Exercise.** Show that the coadjoint orbits for  $SE(2)$  are the cylinders  $T^*S_\alpha^1 = \{(\mu, \beta) : |\beta| = \text{constant}\}$ , together with points on the  $\mu$ -axis. ★

**Exercise.** What are the Casimirs for  $SE(2)$ ? ★

## 6.5 Semidirect-product group $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$

As a further example, we compute the adjoint and coadjoint actions for the semidirect-product group  $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$ . The change to  $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$  from  $SE(2) \cong SO(2) \circledast \mathbb{R}^2$  incorporates area-preserving dilations and both left and right  $SO(2)$  rotations into  $SE(2)$ , as well as the translations  $\mathbb{R}^2$ . These additional degrees of freedom may be recognised in the polar decomposition of  $R \in SL(2, \mathbb{R})$  into  $R = O_1 SO_2$ , in which  $S$  is a  $2 \times 2$  diagonal matrix of unit determinant, and  $O_1$  and  $O_2$  are the left and right  $SO(2)$  rotations, respectively. After defining the matrix representations of the Lie group and its Lie algebra, we derive its AD, Ad and ad actions. We then define the dual Lie algebra and derive its Ad\* and ad\* actions. Finally, we compute its coadjoint motion equations.

### 6.5.1 Definitions for $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$

We consider the semidirect-product Lie group  $G = SL(2, \mathbb{R}) \circledast \mathbb{R}^2$  in which  $SL(2, \mathbb{R})$  acts on  $\mathbb{R}^2$  by matrix multiplication from the left. The group composition rule is

$$(\tilde{R}, \tilde{v})(R, v) = (\tilde{R}R, \tilde{R}v + \tilde{v}), \quad (6.5.1)$$



6.5 SEMIDIRECT-PRODUCT GROUP  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  157

which can be represented by multiplication of  $3 \times 3$  matrices. That is, the action of  $G$  on  $\mathbb{R}^3$  has a matrix representation, given by

$$(R, v) \mapsto \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix}, \quad (6.5.2)$$

for  $R \in SL(2, \mathbb{R})$ , so that  $\det R = 1$  and  $v \in \mathbb{R}^2$ . The matrix multiplication

$$\begin{pmatrix} \tilde{R} & \tilde{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \tilde{R}R & \tilde{R}v + \tilde{v} \\ 0 & 1 \end{pmatrix} \quad (6.5.3)$$

agrees with (6.5.1). The inverse is given by

$$(\tilde{R}, \tilde{v})^{-1} = (\tilde{R}^{-1}, -\tilde{R}^{-1}\tilde{v}) \quad (6.5.4)$$

and the identity element is  $(\mathbb{1}, 0)$ , where  $\mathbb{1}$  is the  $2 \times 2$  identity element of  $SL(2, \mathbb{R})$ .

The  $3 \times 3$  matrix representation of this group in (6.5.2) acts on the *extended* vector  $(r, 1)^T$  as

$$\begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r \\ 1 \end{pmatrix} = \begin{pmatrix} Rr + v \\ 1 \end{pmatrix}.$$

The Lie group  $G = SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  has five parameters. These may be identified by considering their action on an ellipse whose centre is initially at the origin of coordinates in the plane. In the polar decomposition of the matrix  $R = O_2 S O_1 \in SL(2, \mathbb{R})$ , the orthogonal matrix  $O_1 \in SO(2)$  rotates the planar reference coordinates about the centre at the origin into the principal axes of the fixed ellipse. The diagonal matrix  $S$  with  $\det S = 1$  then stretches the principal axes while preserving the area of the ellipse. Next, the orthogonal matrix  $O_2 \in SO(2)$  rigidly rotates the rescaled ellipse about its centre. And finally, the vector  $v \in \mathbb{R}^2$  translates the centre of the rescaled rotated ellipse to a new location in the plane.

### 6.5.2 AD, Ad, and ad actions

A matrix calculation represents the AD action  $G \times G \rightarrow G$  as

$$\begin{aligned} \text{AD}_{(\tilde{R}, \tilde{v})}(R, v) &= (\tilde{R}, \tilde{v})(R, v)(\tilde{R}^{-1}, -\tilde{R}^{-1}\tilde{v}) \\ &= (\text{AD}_{\tilde{R}} R, -\text{AD}_{\tilde{R}} R\tilde{v} + \tilde{R}v + \tilde{v}) \\ &= \begin{pmatrix} \tilde{R}R\tilde{R}^{-1} & -\tilde{R}R\tilde{R}^{-1}\tilde{v} + \tilde{R}v + \tilde{v} \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Next, the Ad action  $G \times \mathfrak{g} \rightarrow \mathfrak{g}$  may be computed. By taking derivatives of the matrix representation (6.5.2) for the Lie group  $G = SL(2, \mathbb{R}) \otimes \mathbb{R}^2$  at the identity, one defines the basis

$$\begin{aligned} X &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\ H_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \end{aligned} \quad (6.5.5)$$

for the matrix representation of the Lie algebra  $\mathfrak{g} = sl(2, \mathbb{R}) \oplus \mathbb{R}^2$ . The traceless property of the  $2 \times 2$   $sl(2, \mathbb{R})$  matrices follows from taking the derivative at the identity of the unit determinant condition for  $SL(2, \mathbb{R})$ . A matrix representation of the Lie algebra

$$\xi = (A, h) \in \mathfrak{g} = sl(2, \mathbb{R}) \oplus \mathbb{R}^2$$

is given in this basis by  $A = xX + yY + zZ$  and  $h = h_1H_1 + h_2H_2$ , with  $(x, y, z) \in \mathbb{R}^3$  and  $(h_1, h_2) \in \mathbb{R}^2$ , so that

$$A = \begin{pmatrix} x & y+z \\ y-z & -x \end{pmatrix} \in sl(2, \mathbb{R}) \quad \text{and} \quad h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \quad (6.5.6)$$

Consequently, one finds the Ad action in matrix form,

$$\begin{aligned} \text{Ad}_{(\tilde{R}, \tilde{v})}(A, h) &= \left. \frac{d}{dt} \right|_{t=0} \text{AD}_{(\tilde{R}, \tilde{v})}(e^{tA}, th) \\ &= \begin{pmatrix} \tilde{R}A\tilde{R}^{-1} & -\tilde{R}A\tilde{R}^{-1}\tilde{v} + \tilde{R}h \\ 0 & 0 \end{pmatrix}. \end{aligned} \quad (6.5.7)$$

6.5 SEMIDIRECT-PRODUCT GROUP  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  159

The matrix Lie group representation (6.5.2) implies the following  $3 \times 3$  matrix Lie algebra representation, denoted by  $M$ ,

$$(A, h) \mapsto M(x, y, z, h_1, h_2) = \begin{pmatrix} x & y+z & h_1 \\ y-z & -x & h_2 \\ 0 & 0 & 0 \end{pmatrix}. \quad (6.5.8)$$

This may be abbreviated as

$$(A, h) \mapsto M(A, h) = \begin{pmatrix} A & h \\ 0 & 0 \end{pmatrix}. \quad (6.5.9)$$

In this matrix representation, one obtains

$$\text{Ad}_{(\tilde{R}, \tilde{v})}(A, h) = \begin{pmatrix} \tilde{R} & \tilde{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & h \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{R}^{-1} & -\tilde{R}^{-1}\tilde{v} \\ 0 & 1 \end{pmatrix}, \quad (6.5.10)$$

which recovers the result in formula (6.5.7).

The ad action may now be computed by a matrix commutation,

$$\begin{aligned} \text{ad}_{(\tilde{A}, \tilde{h})}(A, h) &= \left[ \begin{pmatrix} \tilde{A} & \tilde{h} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} A & h \\ 0 & 0 \end{pmatrix} \right] \\ &= \begin{pmatrix} [\tilde{A}, A] & \tilde{A}\tilde{h} - A\tilde{h} \\ 0 & 0 \end{pmatrix} \\ &=: \left( [\tilde{A}, A], \tilde{A}\tilde{h} - A\tilde{h} \right). \end{aligned} \quad (6.5.11)$$

This is the standard form of the ad action of a semidirect-product Lie algebra.

## 160 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

**6.5.3 Ad\* and ad\* actions**

A  $3 \times 3$  matrix representation of the Lie algebra  $\mathfrak{g} = sl(2, \mathbb{R}) \oplus \mathbb{R}^2$  is defined by the basis

$$\begin{aligned} X &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & Y &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & Z &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ H_1 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & H_2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (6.5.12)$$

The corresponding basis elements in  $\mathfrak{g}^*$  dual to these are their transposes, denoted by  $X^*, Y^*, Z^*, H_1^*, H_2^*$ . An element of  $\mathfrak{g}^*$  can be written as a row component vector with respect to this basis. Namely,

$$aX^* + bY^* + cZ^* + k_1H_1^* + k_2H_2^* =: \begin{pmatrix} a & b & c & k_1 & k_2 \end{pmatrix}. \quad (6.5.13)$$

The matrix representation of the dual Lie algebra  $\mathfrak{g}^*$  is defined via the following map,

$$aX^* + bY^* + cZ^* + k_1H_1^* + k_2H_2^* \mapsto M^*(a, b, c, k_1, k_2), \quad (6.5.14)$$

with the traceless matrix representation

$$M^*(a, b, c, k_1, k_2) = \begin{pmatrix} a & b - c & 0 \\ b + c & -a & 0 \\ k_1 & k_2 & 0 \end{pmatrix}. \quad (6.5.15)$$

To compute the Ad\* action, denote

$$D = \begin{pmatrix} a & b - c \\ b + c & -a \end{pmatrix}, \quad k = \begin{pmatrix} k_1 & k_2 \end{pmatrix}. \quad (6.5.16)$$

Then consider  $(D, k) \in \mathfrak{g}^*$  defined as the element  $aX^* + \dots + k_2H_2^*$ , which in the matrix representation (6.5.15) corresponds to

$$(D, k) \mapsto \begin{pmatrix} D & 0 \\ k & 0 \end{pmatrix}. \quad (6.5.17)$$

6.5 SEMIDIRECT-PRODUCT GROUP  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  161

Let  $A$  and  $h$  be as in (6.5.6). Then, upon denoting the map  $\mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$  as the pairing  $\langle \cdot, \cdot \rangle_{\mathfrak{g}^* \times \mathfrak{g}}$ , one finds

$$\begin{aligned} \langle \text{Ad}_{(R,v)}^*(D, k), (A, h) \rangle_{\mathfrak{g}^* \times \mathfrak{g}} &= \langle (D, k), \text{Ad}_{(R,v)}(A, h) \rangle \\ &= \frac{1}{2} \text{Tr} \left( \begin{pmatrix} D & 0 \\ k & 0 \end{pmatrix} g \begin{pmatrix} A & h \\ 0 & 0 \end{pmatrix} g^{-1} \right) \\ &= \frac{1}{2} \text{Tr} \left( g^{-1} \begin{pmatrix} D & 0 \\ k & 0 \end{pmatrix} g \begin{pmatrix} A & h \\ 0 & 0 \end{pmatrix} \right) \end{aligned}$$

where, as in (6.5.2),

$$g = \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix}. \quad (6.5.18)$$

A side calculation yields

$$g^{-1} \begin{pmatrix} D & 0 \\ k & 0 \end{pmatrix} g = \begin{pmatrix} R^{-1}(D - vk)R & * \\ kR & * \end{pmatrix}, \quad (6.5.19)$$

in which the entries at positions  $*$  are immaterial, since taking the trace with the last factor in the pairing above will eliminate them. One may also make the upper left part of this matrix traceless, by subtracting the product of its trace times half the  $2 \times 2$  identity,  $\mathbb{1}$ . This is allowed, since the trace of this product with the traceless matrix  $A$  will not contribute in the pairing.

Hence, one arrives at the matrix representation of the  $\text{Ad}^*$  action,

$$\text{Ad}_{(R,v)}^*(D, k) = \begin{pmatrix} R^{-1}(D - vk)R + \frac{1}{2}(\mathbf{v} \cdot \mathbf{k})\mathbb{1} & 0 \\ kR & 0 \end{pmatrix}, \quad (6.5.20)$$

in which  $\text{Tr}(R^{-1}DR) = 0$ , so the trace of the  $2 \times 2$  upper left submatrix vanishes. The corresponding result for  $(R, v)^{-1}$  is

$$\text{Ad}_{(R,v)^{-1}}^*(D, k) = \begin{pmatrix} RDR^{-1} + (vk)R^{-1} - \frac{1}{2}\text{Tr}((vk)R^{-1}) & 0 \\ kR^{-1} & 0 \end{pmatrix}.$$

## 162 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

The next step is to compute  $\text{ad}^*$  from

$$\begin{aligned} \langle \text{ad}_{(A,h)}^*(D, k), (\tilde{A}, \tilde{h}) \rangle &= \langle (D, k), \text{ad}_{(A,h)}(\tilde{A}, \tilde{h}) \rangle \quad (6.5.21) \\ &= \frac{1}{2} \text{Tr} \left( \begin{pmatrix} D & 0 \\ k & 0 \end{pmatrix} \left[ \begin{pmatrix} A & h \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \tilde{A} & \tilde{h} \\ 0 & 0 \end{pmatrix} \right] \right). \end{aligned}$$

This leads to

$$\text{ad}_{(A,h)}^* \begin{pmatrix} D & 0 \\ k & 0 \end{pmatrix} = \begin{pmatrix} [D, A] - hk + \frac{1}{2}(\mathbf{h} \cdot \mathbf{k})\mathbb{1} & 0 \\ kA & 0 \end{pmatrix}. \quad (6.5.22)$$

The commutator  $[D, A]$  in this equation may be computed in the  $2 \times 2$  matrix basis  $(X, Y, Z)$  for  $sl(2, \mathbb{R})$ , since in that basis we have  $A = xX + yY + zZ$  and  $D = aX + bY - cZ$ , the latter because  $(X^T, Y^T, Z^T) = (X, Y, -Z)$ .

**Summary.** The adjoint and coadjoint actions for  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  are

$$\begin{aligned} \text{Ad}_{(\tilde{R}, \tilde{v})}(R, v) &= \left( \tilde{R}R\tilde{R}^{-1}, -\tilde{R}R\tilde{R}^{-1}\tilde{v} + \tilde{R}v + \tilde{v} \right), \\ \text{Ad}_{(\tilde{R}, \tilde{v})}(A, h) &= \left( \tilde{R}A\tilde{R}^{-1}, -\tilde{R}A\tilde{R}^{-1}\tilde{v} + \tilde{R}h \right), \\ \text{ad}_{(\tilde{A}, \tilde{h})}(A, h) &= \left( [\tilde{A}, A], \tilde{A}h - A\tilde{h} \right), \\ \text{Ad}_{(R,v)}^*(D, k) &= \left( R^{-1}(D - vk)R + \frac{1}{2}(\mathbf{v} \cdot \mathbf{k})\mathbb{1}, kR \right), \\ \text{ad}_{(A,h)}^*(D, k) &= \left( [D, A] - hk + \frac{1}{2}(\mathbf{h} \cdot \mathbf{k})\mathbb{1}, kA \right). \end{aligned}$$

### 6.5.4 Coadjoint motion relation

According to Proposition 4.2.2, the coadjoint motion relation arises from differentiating the coadjoint orbit relation. Let  $g(t) = (R(t), v(t))$  be a path in the Lie group  $G = SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  and let  $\mu(t) = (D(t), l(t))$  be a path in the dual Lie algebra  $\mathfrak{g}^*$ . Then Proposition 4.2.2 supplies the coadjoint motion relation

$$\frac{d}{dt} \left( \text{Ad}_{g(t)}^* \mu \right) = \text{Ad}_{g(t)}^* \left[ \frac{d\mu}{dt} - \text{ad}_{\xi(t)}^* \mu(t) \right], \quad (6.5.23)$$

where  $\xi(t) = g(t)^{-1} \dot{g}(t) = (A(t), h(t))$ .

6.5 SEMIDIRECT-PRODUCT GROUP  $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$  163

Corollary 4.2.11 provides the differential equation for the coadjoint orbit relation

$$\mu(t) = \text{Ad}_{g(t)}^* \mu(0).$$

The desired differential equation is the *coadjoint motion equation*

$$\frac{d\mu}{dt} = \text{ad}_{\xi(t)}^* \mu(t).$$

This equation is expressed for the Lie group  $G = SL(2, \mathbb{R}) \circledast \mathbb{R}^2$  as

$$\frac{d}{dt} (D, k) = \left( [D, A] - hk + \frac{1}{2}(\mathbf{h} \cdot \mathbf{k})\mathbb{1}, kA \right), \quad (6.5.24)$$

with  $(D, k) \in \mathfrak{g}^*$  defined in (6.5.17) and  $(A, h) \in \mathfrak{g}$  defined in (6.5.6). When  $(A, h)$  is known as a smooth invertible function of  $(D, k)$ , e.g.,  $(A, h) = (\partial H / \partial D, \partial H / \partial k)$  for a Hamiltonian  $H(D, k)$ , then this equation becomes a Lie–Poisson Hamiltonian system for the dynamics of a rotating, stretching, circulating and translating ellipse, whose area is preserved by the motion. See [Ho1991] for the corresponding discussion of coadjoint motion on  $GL(2, \mathbb{R})$ .

**Exercise.** Recompute the ad action and  $\text{ad}^*$  action for the semidirect-product group  $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$  after introducing the polar decomposition  $R = O_1 S O_2$ , in which  $S$  is a diagonal matrix and  $O_1$  and  $O_2$  are  $SO(2)$  rotations. Identify two types of centrifugal terms in the coadjoint orbit equations. ★

**Exercise.** Compare the ad and  $\text{ad}^*$  actions for the semidirect-product groups  $SL(2, \mathbb{R}) \circledast \mathbb{R}^2$  and  $Sp(2) \circledast \mathbb{R}^2$ . ★

**Exercise.** Compute the ad action and  $\text{ad}^*$  action for the Galilean group  $SE(3) \circledast \mathbb{R}^4$ . ★

**Exercise.** Write the semidirect-product action  $SU(n) \ltimes H_n$  of  $n$ -dimensional unitary transformations on the vector space of  $n \times n$  Hermitian matrices. Compute its  $\text{ad}$  action and  $\text{ad}^*$  action.

How is this related to  $O(n) \ltimes S_n$ , the semidirect-product group of  $n$ -dimensional orthogonal transformations  $O(n)$  acting on the vector space of  $n \times n$  symmetric matrices  $S_n$ ?

★

## 6.6 Galilean group

### 6.6.1 Definitions for $G(3)$

As discussed in Chapter 1, the Galilean group in three dimensions  $G(3)$  has ten parameters,

$$(O \in SO(3), \mathbf{r}_0 \in \mathbb{R}^3, \mathbf{v}_0 \in \mathbb{R}^3, t_0 \in \mathbb{R}).$$

The Galilean group is a semidirect-product Lie group, which may be written as

$$G(3) = SE(3) \ltimes \mathbb{R}^4 = (SO(3) \ltimes \mathbb{R}^3) \ltimes \mathbb{R}^4. \quad (6.6.1)$$

That is, the subgroup of Euclidean motions, which comprises rotations and Galilean velocity boosts  $(O, \mathbf{v}_0) \in SE(3)$ , acts homogeneously on the subgroups of space and time translations  $(\mathbf{r}_0, t_0) \in \mathbb{R}^4$  which commute with each other.



### Matrix representation of $G(3)$

The formula for group composition  $G(3) \times G(3) \rightarrow G(3)$  may be represented by matrix multiplication as in Equation (1.5.3),

$$\begin{aligned} \tilde{g}g &= \begin{pmatrix} \tilde{O} & \tilde{\mathbf{v}} & \tilde{\mathbf{r}} \\ \mathbf{0} & 1 & \tilde{t} \\ \mathbf{0} & 0 & 1 \end{pmatrix} \begin{pmatrix} O & \mathbf{v} & \mathbf{r} \\ \mathbf{0} & 1 & t \\ \mathbf{0} & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \tilde{O}O & \tilde{O}\mathbf{v} + \tilde{\mathbf{v}} & \tilde{O}\mathbf{r} + \tilde{\mathbf{v}}t + \tilde{\mathbf{r}} \\ \mathbf{0} & 1 & \tilde{t} + t \\ \mathbf{0} & 0 & 1 \end{pmatrix}. \end{aligned} \quad (6.6.2)$$

This may also be expressed succinctly in row notation as

$$\begin{aligned} \tilde{g}g &= (\tilde{O}, \tilde{\mathbf{v}}, \tilde{\mathbf{r}}, \tilde{t})(O, \mathbf{v}, \mathbf{r}, t) \\ &= (\tilde{O}O, \tilde{O}\mathbf{v} + \tilde{\mathbf{v}}, \tilde{O}\mathbf{r} + \tilde{\mathbf{v}}t + \tilde{\mathbf{r}}, \tilde{t} + t). \end{aligned} \quad (6.6.3)$$

The inverse operation is given in matrix form as

$$\begin{aligned} \tilde{g}^{-1} &= \begin{pmatrix} \tilde{O}^{-1} & -\tilde{O}^{-1}\tilde{\mathbf{v}} & -\tilde{O}^{-1}(\tilde{\mathbf{r}} - \tilde{t}\tilde{\mathbf{v}}) \\ \mathbf{0} & 1 & -\tilde{t} \\ \mathbf{0} & 0 & 1 \end{pmatrix} \\ &= (\tilde{O}^{-1}, -\tilde{O}^{-1}\tilde{\mathbf{v}}, -\tilde{O}^{-1}(\tilde{\mathbf{r}} - \tilde{t}\tilde{\mathbf{v}}), -\tilde{t}). \end{aligned} \quad (6.6.4)$$

#### 6.6.2 AD, Ad, and ad actions of $G(3)$

A matrix calculation represents the AD action  $G(3) \times G(3) \rightarrow G(3)$  as

$$\begin{aligned} \text{AD}_{\tilde{g}}g &= \text{AD}_{(\tilde{O}, \tilde{\mathbf{v}}, \tilde{\mathbf{r}}, \tilde{t})}(O, \mathbf{v}, \mathbf{r}, t) \\ &= (\tilde{O}, \tilde{\mathbf{v}}, \tilde{\mathbf{r}}, \tilde{t})(O, \mathbf{v}, \mathbf{r}, t)(\tilde{O}^{-1}, -\tilde{O}^{-1}\tilde{\mathbf{v}}, -\tilde{O}^{-1}(\tilde{\mathbf{r}} - \tilde{t}\tilde{\mathbf{v}}), -\tilde{t}) \\ &= \left( \text{AD}_{\tilde{O}}O, -(\text{AD}_{\tilde{O}}O)\tilde{\mathbf{v}} + \tilde{O}\mathbf{v} + \tilde{\mathbf{v}}, \right. \\ &\quad \left. -(\text{AD}_{\tilde{O}}O)(\mathbf{r} - \mathbf{v}t) - (\tilde{O}\mathbf{v} + \tilde{\mathbf{v}})\tilde{t} + \tilde{O}\mathbf{r} + \tilde{\mathbf{v}}t + \tilde{\mathbf{r}}, t \right). \end{aligned}$$

Next, the Ad action  $G(3) \times \mathfrak{g}(3) \rightarrow \mathfrak{g}(3)$  may be computed. By taking derivatives of the matrix representation (1.5.3) for the Galilean

## 166 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

group  $G(3)$  at the identity, one finds

$$\Xi := g^{-1}(s)g'(s)|_{s=0} = \begin{pmatrix} \hat{\Xi} & \mathbf{v} & \mathbf{r} - \mathbf{v}t \\ 0 & 0 & t \\ 0 & 0 & 0 \end{pmatrix} =: (\hat{\Xi}, \mathbf{v}, \mathbf{r} - \mathbf{v}t, t),$$

in terms of the  $3 \times 3$  skew-symmetric matrix  $\hat{\Xi} = O^{-1}(s)O'(s)|_{s=0}$  and the Galilean shift parameters  $\mathbf{v}$ ,  $\mathbf{r}$ ,  $t$ . Consequently, one finds the Ad action in matrix form,

$$\begin{aligned} \text{Ad}_{\hat{g}}\Xi &= \text{Ad}_{(\tilde{O}, \tilde{\mathbf{v}}, \tilde{\mathbf{r}}, \tilde{t})}(\hat{\Xi}, \mathbf{v}, \mathbf{r} - \mathbf{v}t, t) && (6.6.5) \\ &= \begin{pmatrix} \text{Ad}_{\tilde{O}}\hat{\Xi} & -(\text{Ad}_{\tilde{O}}\hat{\Xi})\tilde{\mathbf{v}} + \tilde{O}\mathbf{v} & -(\text{Ad}_{\tilde{O}}\hat{\Xi})(\tilde{\mathbf{r}} - \tilde{\mathbf{v}}\tilde{t}) - \tilde{O}\mathbf{v}\tilde{t} + \tilde{O}(\mathbf{r} - \mathbf{v}t) + \tilde{\mathbf{v}}t \\ 0 & 0 & t \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Taking the derivative of the tilde variables and evaluating at the identity recovers the commutation relation asserted in Section 1.7,

$$\begin{aligned} \text{ad}_{\tilde{\Xi}}\Xi &= [(\tilde{\Xi}, \tilde{\mathbf{v}}, \tilde{\mathbf{r}} - \tilde{\mathbf{v}}\tilde{t}, \tilde{t}), (\hat{\Xi}, \mathbf{v}, \mathbf{r} - \mathbf{v}t, t)] \\ &= \left( [\tilde{\Xi}, \hat{\Xi}], \tilde{\Xi}\mathbf{v} - \hat{\Xi}\tilde{\mathbf{v}}, \tilde{\Xi}(\mathbf{r}, \mathbf{v}, t) - \hat{\Xi}(\tilde{\mathbf{r}}, \tilde{\mathbf{v}}, \tilde{t}), 0 \right), \end{aligned}$$

where

$$\begin{aligned} &\tilde{\Xi}(\mathbf{r}, \mathbf{v}, t) - \hat{\Xi}(\tilde{\mathbf{r}}, \tilde{\mathbf{v}}, \tilde{t}) \\ &:= \left( \tilde{\Xi}(\mathbf{r} - \mathbf{v}t) + \tilde{\mathbf{v}}t \right) - \left( \hat{\Xi}(\tilde{\mathbf{r}} - \tilde{\mathbf{v}}\tilde{t}) + \mathbf{v}\tilde{t} \right). \end{aligned}$$

The Galilean group's  $\text{Ad}^*$  and  $\text{ad}^*$  actions may now be obtained by using the matrix transpose pairing to define the dual Lie algebra.

**Exercise.** Compute the coadjoint actions of  $G(3)$ , the Galilean group in three dimensions. ★

## 6.7 Iterated semidirect products

### ad action

Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ . Consider the Lie group obtained by iterating the semidirect-product action of  $G$  on itself  $n$  times,

$$G_1 \mathbb{S} (\dots (G_{n-2} \mathbb{S} (G_{n-1} \mathbb{S} G_n)) \dots). \quad (6.7.1)$$

Its Lie algebra elements are denoted as

$$(u_1, u_2, \dots, u_n) \in \mathfrak{g}^{\times n} := \mathfrak{g}_1 \times (\dots (\mathfrak{g}_{n-2} \times (\mathfrak{g}_{n-1} \times \mathfrak{g}_n)) \dots).$$

These possess the iterated semidirect product Lie algebra action

$$(u_1, u_2, \dots, u_n) \in \mathfrak{g}^{\mathbb{S}n} := \mathfrak{g}_1 \mathbb{S} (\dots (\mathfrak{g}_{n-2} \mathbb{S} (\mathfrak{g}_{n-1} \mathbb{S} \mathfrak{g}_n)) \dots)$$

given by

$$\begin{aligned} & \text{ad}_{(u_1, u_2, \dots, u_n)}(v_1, v_2, \dots, v_n) \\ &= \left( \text{ad}_{u_1} v_1, \text{ad}_{u_1+u_2} v_2 + \text{ad}_{u_2} v_1, \text{ad}_{u_1+u_2+u_3} v_3 + \text{ad}_{u_3} (v_1 + v_2), \right. \\ & \quad \left. \dots, \sum_{k=1}^n \text{ad}_{u_k} v_n + \sum_{k=1}^{n-1} \text{ad}_{u_n} v_k \right) \end{aligned} \quad (6.7.2)$$

in which the level  $n = 1$  and  $n = 2$  formulas are already familiar.

**Exercise.** Prove formula (6.7.2) by induction, in which the  $m$ -th step is given by

$$\dots, \text{ad}_{u_m} v_m + \text{ad}_{u_m} \left( \sum_{k=1}^{m-1} v_k \right) - \text{ad}_{v_m} \left( \sum_{k=1}^{m-1} u_k \right), \dots$$

This is how infinitesimal transformations compose under iteration of semidirect-product Lie algebra action. ★

## 168 6 : SEMIDIRECT-PRODUCT GROUP ACTIONS

**ad\* action**

One calculates  $\text{ad}^*$  for the iterated semidirect product by using (6.7.2) in the pairing  $\langle \cdot, \cdot \rangle : (\mathfrak{g}^{\otimes n})^* \times \mathfrak{g}^{\otimes n} \rightarrow \mathbb{R}$

$$\begin{aligned} & \langle (\mu_1, \mu_2, \dots, \mu_n), \text{ad}_{(u_1, u_2, \dots, u_n)}(v_1, v_2, \dots, v_n) \rangle \\ &= \langle \text{ad}^*_{(u_1, u_2, \dots, u_n)}(\mu_1, \mu_2, \dots, \mu_n), (v_1, v_2, \dots, v_n) \rangle \end{aligned}$$

to find the coadjoint action

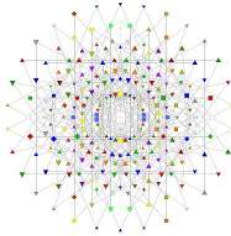
$$\begin{aligned} & \text{ad}^*_{(u_1, u_2, \dots, u_n)}(\mu_1, \mu_2, \dots, \mu_n) \tag{6.7.3} \\ &= \left( \sum_{k=1}^n \text{ad}^*_{u_k} \mu_k, \dots, \sum_{k=1}^m \text{ad}^*_{u_k} \mu_m + \sum_{k=m+1}^n \text{ad}^*_{u_k} \mu_k, \dots, \sum_{k=1}^n \text{ad}^*_{u_k} \mu_n \right). \end{aligned}$$

**Exercise.** Derive the system of Euler–Poincaré equations for coadjoint motion on  $(\mathfrak{g}^{\otimes n})^*$ .

Show that its Lie–Poisson structure may be diagonalised by taking a linear combination of the variables. ★

**Exercise.** Use the hat map to write the system of Euler–Poincaré equations for coadjoint motion on  $(\mathfrak{so}(3)^{\otimes n})^*$  in  $\mathbb{R}^3$  vector form.

Also write its diagonal Lie–Poisson structure. ★



# 7

## EULER–POINCARÉ AND LIE–POISSON EQUATIONS ON $SE(3)$

---

### Contents

7.1	Euler–Poincaré equations for left-invariant Lagrangians under $SE(3)$	170
7.1.1	Legendre transform from $se(3)$ to $se(3)^*$	172
7.1.2	Lie–Poisson bracket on $se(3)^*$	172
7.1.3	Coadjoint motion on $se(3)^*$	173
7.2	Kirchhoff equations on $se(3)^*$	176
7.2.1	Looks can be deceiving: The heavy top	178

---

## 7.1 Euler–Poincaré equations for left-invariant Lagrangians under $SE(3)$

The matrix Euler–Poincaré equation for left-invariant Lagrangians was introduced in Proposition 2.4.3, in the context of Manakov’s [Man1976] commutator formulation of the Euler rigid-body equations. As we shall see, Euler–Poincaré evolution is naturally expressed as coadjoint motion.

The variational derivatives of a given left-invariant Lagrangian  $\ell(\xi, \alpha) : se(3) \rightarrow \mathbb{R}$  are expressed as

$$(\mu, \beta) = \left( \frac{\delta \ell}{\delta \xi}, \frac{\delta \ell}{\delta \alpha} \right) \in se(3)^*. \quad (7.1.1)$$

The commutator form of the matrix Euler–Poincaré Equation (2.4.27) is

$$\frac{dM}{dt} = [M, \Omega] \quad \text{with} \quad M = \frac{\delta l}{\delta \Omega}, \quad (7.1.2)$$

for any Lagrangian  $l(\Omega)$ , where  $\Omega = g^{-1}\dot{g} \in \mathfrak{g}$  and  $\mathfrak{g}$  is the left-invariant matrix Lie algebra of any matrix Lie group  $G$ . Recall that  $M = \delta l / \delta \Omega \in \mathfrak{g}^*$ , where the dual  $\mathfrak{g}^*$  is defined in terms of the matrix trace pairing. Hence, we may rewrite the Euler–Poincaré Equation (7.1.2) in terms of the left action of the Lie algebra on its dual. This happens by the  $\text{ad}^*$  action; so the Euler–Poincaré equation becomes

$$\frac{dM}{dt} = \text{ad}_{\Omega}^* M \quad \text{with} \quad M = \frac{\delta l}{\delta \Omega}. \quad (7.1.3)$$

**Proposition 7.1.1** *The corresponding Euler–Poincaré equation for the special Euclidean group in three dimensions,  $SE(3)$ , describes coadjoint motion on  $se(3)^*$ . Namely,*

$$\begin{aligned} \left\langle \left( \frac{d\mu}{dt}, \frac{d\beta}{dt} \right), (\tilde{\xi}, \tilde{\alpha}) \right\rangle &= \langle \text{ad}_{(\xi, \alpha)}^* (\mu, \beta), (\tilde{\xi}, \tilde{\alpha}) \rangle & (7.1.4) \\ &= \langle (\text{ad}_{\xi}^* \mu + \beta \diamond \alpha, -\xi \beta), (\tilde{\xi}, \tilde{\alpha}) \rangle. \end{aligned}$$

**Proof.** The proof follows immediately from the definition of the Euler–Poincaré equation for  $SE(3)$ . ■

**Definition 7.1.1 (Euler–Poincaré equation for  $SE(3)$ )** *The Euler–Poincaré equation for the special Euclidean group in three dimensions,  $SE(3)$ , is*

$$\left( \frac{d\mu}{dt}, \frac{d\beta}{dt} \right) = \text{ad}_{(\xi, \alpha)}^*(\mu, \beta). \quad (7.1.5)$$

In the  $so(3)^*$  and  $\mathbb{R}^3$  components of  $se(3)^*$  this is

$$\frac{d\mu}{dt} = \text{ad}_{\xi}^* \mu + \beta \diamond \alpha \quad \text{and} \quad \frac{d\beta}{dt} = -\xi \beta, \quad (7.1.6)$$

or in terms of variational quantities,

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_{\xi}^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta \alpha} \diamond \alpha \quad \text{and} \quad \frac{d}{dt} \frac{\delta l}{\delta \alpha} = -\xi \frac{\delta l}{\delta \alpha}. \quad (7.1.7)$$

**Remark 7.1.1 (Vector Euler–Poincaré equation for  $SE(3)$ )** In vector form, the EP equations become, cf. the vector form of  $\text{ad}^*$  for  $se(3)$  in (6.3.3),

$$\dot{\mu} = \mu \times \xi - \alpha \times \beta \quad \text{and} \quad \dot{\beta} = -\xi \times \beta. \quad (7.1.8)$$

That is, in terms of vector variational quantities,

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \frac{\delta l}{\delta \xi} \times \xi - \alpha \times \frac{\delta l}{\delta \alpha} \quad \text{and} \quad \frac{d}{dt} \frac{\delta l}{\delta \alpha} = -\xi \times \frac{\delta l}{\delta \alpha}. \quad (7.1.9)$$

These vector EP equations on  $se(3)^*$  are readily seen to conserve the quantities  $C_1 = \mu \cdot \beta$  and  $C_2 = |\beta|^2$ , corresponding to a reference direction  $\beta(0)$  and the projection of the angular momentum vector  $\mu$  in this direction at any time. The quantities  $C_1$  and  $C_2$  will turn out to be *Casimirs* for the heavy top. Level surfaces of the quantities  $C_1$  and  $C_2$  define the coadjoint orbits on which the motion takes place for *any* choice of Lagrangian. □

172 7 : EULER–POINCARÉ AND LIE–POISSON EQUATIONS ON  $SE(3)$ 

**Exercise.** What are the corresponding vector EP equations on  $se(2)^*$ ? ★

7.1.1 Legendre transform from  $se(3)$  to  $se(3)^*$ 

We Legendre-transform the reduced Lagrangian  $\ell(\xi, \alpha) : se(3) \rightarrow \mathbb{R}$  to the Hamiltonian

$$h(\mu, \beta) = \mu \cdot \xi + \beta \cdot \alpha - \ell(\xi, \alpha),$$

whose variations are given by

$$\begin{aligned} \delta h(\mu, \beta) &= \frac{\partial h}{\partial \mu} \cdot \delta \mu + \frac{\partial h}{\partial \beta} \cdot \delta \beta \\ &= \xi \cdot \delta \mu + \alpha \cdot \delta \beta + \left( \mu - \frac{\partial \ell}{\partial \xi} \right) \cdot \delta \xi + \left( \beta - \frac{\partial \ell}{\partial \alpha} \right) \cdot \delta \alpha. \end{aligned}$$

Consequently, the Hamiltonian has derivatives  $\partial h / \partial \mu = \xi$  and  $\partial h / \partial \beta = \alpha$ .

**Exercise.** What is the Legendre transform from  $se(2)$  to  $se(2)^*$ ? ★

7.1.2 Lie–Poisson bracket on  $se(3)^*$ 

Rearranging the time derivative of a smooth function  $f$  yields

$$\begin{aligned} \frac{df}{dt}(\mu, \beta) &= \frac{\partial f}{\partial \mu} \cdot \dot{\mu} + \frac{\partial f}{\partial \beta} \cdot \dot{\beta} \\ &= \frac{\partial f}{\partial \mu} \cdot \left( \mu \times \frac{\partial h}{\partial \mu} + \beta \times \frac{\partial h}{\partial \beta} \right) + \frac{\partial f}{\partial \beta} \cdot \beta \times \frac{\partial h}{\partial \mu} \end{aligned}$$



## 7.1 LEFT-INVARIANT LAGRANGIANS 173

$$\begin{aligned}
&= -\boldsymbol{\mu} \cdot \frac{\partial f}{\partial \boldsymbol{\mu}} \times \frac{\partial h}{\partial \boldsymbol{\mu}} - \boldsymbol{\beta} \cdot \left( \frac{\partial f}{\partial \boldsymbol{\beta}} \times \frac{\partial h}{\partial \boldsymbol{\mu}} - \frac{\partial h}{\partial \boldsymbol{\beta}} \times \frac{\partial f}{\partial \boldsymbol{\mu}} \right) \\
&=: \{f, h\}.
\end{aligned}$$

This is the *Lie–Poisson bracket* defined on  $se(3)^*$  and expressed in terms of vectors  $(\boldsymbol{\mu}, \boldsymbol{\beta}) \in \mathbb{R}^3 \times \mathbb{R}^3$ . By construction, it returns the equations of motion in Hamiltonian form. The Lie–Poisson bracket may be written in matrix form as

$$\begin{aligned}
\{f, h\} &= \partial_{(\boldsymbol{\mu}, \boldsymbol{\beta})} f^T \mathbb{J} \partial_{(\boldsymbol{\mu}, \boldsymbol{\beta})} h \\
&= \begin{bmatrix} \partial f / \partial \boldsymbol{\mu} \\ \partial f / \partial \boldsymbol{\beta} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{\mu} \times & \boldsymbol{\beta} \times \\ \boldsymbol{\beta} \times & 0 \end{bmatrix} \begin{bmatrix} \partial h / \partial \boldsymbol{\mu} \\ \partial h / \partial \boldsymbol{\beta} \end{bmatrix}. \quad (7.1.10)
\end{aligned}$$

The Hamiltonian matrix  $\mathbb{J}$  has null eigenvectors

$$[0, \boldsymbol{\beta}]^T = \partial_{(\boldsymbol{\mu}, \boldsymbol{\beta})} |\boldsymbol{\beta}|^2 \quad \text{and} \quad [\boldsymbol{\beta}, \boldsymbol{\mu}]^T = \partial_{(\boldsymbol{\mu}, \boldsymbol{\beta})} (\boldsymbol{\beta} \cdot \boldsymbol{\mu}).$$

Consequently the *distinguished functions*  $C_1 = |\boldsymbol{\beta}|^2$  and  $C_2 = \boldsymbol{\mu} \cdot \boldsymbol{\beta}$  whose derivatives are the null eigenvectors of the Hamiltonian matrix  $\mathbb{J}$  will *Poisson-commute* with any smooth function  $f(\boldsymbol{\mu}, \boldsymbol{\beta})$ . Such distinguished functions are called the *Casimirs* of the Lie–Poisson bracket in (7.1.10).

**Exercise.** What is the Lie–Poisson Hamiltonian formulation on  $se(2)^*$ ? ★

### 7.1.3 Coadjoint motion on $se(3)^*$

**Theorem 7.1.1 (Conservation of  $se(3)^*$  momentum)** *The Euler–Poincaré Equation (7.1.5) for  $SE(3)$  conserves the momentum  $\text{Ad}_{(R(t), v(t))}^* (\boldsymbol{\mu}(t), \boldsymbol{\beta}(t))$ . That is,*

$$\frac{d}{dt} \left( \text{Ad}_{(R(t), v(t))}^* (\boldsymbol{\mu}(t), \boldsymbol{\beta}(t)) \right) = 0. \quad (7.1.11)$$

174 7 : EULER–POINCARÉ AND LIE–POISSON EQUATIONS ON  $SE(3)$ 

**Proof.** By Equation (4.2.9) in Proposition 4.2.2 one may rewrite the Euler–Poincaré Equation (7.1.5) for  $SE(3)$  equivalently as

$$\begin{aligned} & \frac{d}{dt} \left( \text{Ad}_{(R(t), v(t))^{-1}}^* (\mu(t), \beta(t)) \right) \\ &= \text{Ad}_{(R(t), v(t))^{-1}}^* \left[ \left( \frac{d\mu}{dt}, \frac{d\beta}{dt} \right) - \text{ad}_{(\xi, \alpha)}^* (\mu(t), \beta(t)) \right] = 0, \end{aligned} \quad (7.1.12)$$

where the left-invariant tangent vectors

$$(\xi(t), \alpha(t)) = (R^{-1}\dot{R}(t), R^{-1}\dot{v}(t))$$

to  $(R, v)$  at the identity are given in Equation (6.3.8) and the operation  $\text{Ad}_{(R(t), v(t))^{-1}}^*$  is given for  $se(3)^*$  explicitly in formula (6.2.5). ■

**Remark 7.1.2 (First integral)** The evolution Equation (7.1.11) conserves the first integral

$$\text{Ad}_{(R(t), v(t))^{-1}}^* (\mu(t), \beta(t)) = (\mu(0), \beta(0)). \quad (7.1.13)$$

This is the analogue for coadjoint motion on  $se(3)^*$  of spatial angular momentum conservation for  $so(3)^*$ . By Equation (6.2.5) the independently conserved  $so(3)^*$  and  $\mathbb{R}^3$  components of the  $se(3)^*$  momentum are

$$\begin{aligned} & \left( R(t)\mu(t)R(t)^{-1} + \text{skew}(v(t) \otimes R(t)\beta(t)), R(t)\beta(t) \right) \\ &= (\mu(0), \beta(0)). \end{aligned} \quad (7.1.14)$$

This formula is particularly convenient, because  $\beta(t) = R(t)^{-1}\beta(0)$  implies that  $\beta(t)$  simply states how the spatial vector  $\beta(0)$  looks in the rotating frame. □

**Corollary 7.1.1** *The solution of the EP Equation (7.1.5) for  $SE(3)$  evolves by coadjoint motion on the dual Lie algebra  $se(3)^*$ .*

**Proof.** The first integral of the evolution Equation (7.1.11) yields

$$(\mu(t), \beta(t)) = \text{Ad}_{(R(t), v(t))}^* (\mu(0), \beta(0)). \quad (7.1.15)$$

This expresses the solution as coadjoint motion on  $se(3)^*$ . ■

**Remark 7.1.3** Equation (7.1.12) for conservation of  $se(3)^*$  momentum is written in vector notation as

$$\begin{aligned} (\boldsymbol{\mu}(0), \boldsymbol{\beta}(0)) &= \text{Ad}_{(R(t), \mathbf{v}(t))}^* (\boldsymbol{\mu}(t), \boldsymbol{\beta}(t)) \\ &= \left( R(t)\boldsymbol{\mu}(t) + \mathbf{v}(t) \times R(t)\boldsymbol{\beta}(t), R(t)\boldsymbol{\beta}(t) \right), \end{aligned} \quad (7.1.16)$$

which may be rearranged into the evolution operator,

$$\begin{aligned} (\boldsymbol{\mu}(t), \boldsymbol{\beta}(t)) &= \text{Ad}_{(R(t), \mathbf{v}(t))}^* (\boldsymbol{\mu}(0), \boldsymbol{\beta}(0)) \\ &= \left( R^{-1}(t)(\boldsymbol{\mu}(0) - \mathbf{v}(t) \times \boldsymbol{\beta}(0)), R^{-1}(t)\boldsymbol{\beta}(0) \right). \end{aligned} \quad (7.1.17)$$

This is the vector form of coadjoint motion on  $se(3)^*$ . It is also the solution of the Hamiltonian system of equations defined by the Lie-Poisson bracket (7.1.10).  $\square$

**Exercise.** Verify directly that this solution satisfies the vector EP equations on  $se(3)^*$  in (7.1.8) and conserves the quantities  $C_1 = \boldsymbol{\mu} \cdot \boldsymbol{\beta}$  and  $C_2 = |\boldsymbol{\beta}|^2$ .  $\star$

**Exercise.** Compute the coadjoint solution of the vector EP equations on  $se(2)^*$  that corresponds to (7.1.17) on  $se(3)^*$ . What quantities are conserved in this case?  $\star$

**Remark 7.1.4** Completing the solution requires the group parameters  $R(t)$  and  $v(t)$  to be reconstructed from  $\xi(t)$  and  $\alpha(t)$  according to  $\dot{R} = R\xi$  and  $\dot{v} = R\alpha$  in Equation (6.3.9). In vector notation, the latter equations become  $\dot{R} = R\xi$  and  $\dot{\mathbf{v}} = R\boldsymbol{\alpha}$ , where  $\mathbf{v}$  is linear displacement of a moving body,  $\boldsymbol{\alpha}$  is its linear velocity and  $\boldsymbol{\xi}$  is its angular velocity.  $\square$

**Theorem 7.1.2 (Kelvin–Noether)** *The Euler–Poincaré Equation (7.1.5) for  $SE(3)$  preserves the natural pairing*

$$\frac{d}{dt} \left\langle (\mu(t), \beta(t)), \text{Ad}_{(R(t), v(t))^{-1}}(\eta, \gamma) \right\rangle = 0, \quad (7.1.18)$$

for any fixed  $(\eta, \gamma) \in \mathfrak{se}(3)$ .

**Proof.** Verifying directly,

$$\begin{aligned} \frac{d}{dt} \left\langle (\mu(t), \beta(t)), \text{Ad}_{(R(t), v(t))^{-1}} \eta \right\rangle & \quad (7.1.19) \\ &= \left\langle \frac{d}{dt} \left( \text{Ad}_{(R(t), v(t))^{-1}}^* (\mu(t), \beta(t)) \right), \eta \right\rangle = 0, \end{aligned}$$

by Equation (7.1.12). ■

**Remark 7.1.5** Theorem 7.1.2 is associated with Kelvin and Noether because it arises from symmetry (Noether) and it happens to coincide with the Kelvin circulation theorem in the Euler–Poincaré formulation of ideal fluid motion [HoMaRa1998]. □

## 7.2 Kirchhoff equations on $se(3)^*$

Suppose the Lagrangian is chosen to be the sum of the kinetic energies of rotational and translational motion of an ellipsoidal underwater vehicle with coincident centres of gravity and buoyancy. In this case, the Lagrangian is given by the sum of the rotational and translational kinetic energies as [HoJeLe1998]

$$l(\xi, \alpha) = \frac{1}{2} \langle \xi, \mathbb{I} \xi \rangle + \frac{1}{2} \langle \alpha, M \alpha \rangle = \frac{1}{2} \boldsymbol{\xi} \cdot \mathbb{I} \boldsymbol{\xi} + \frac{1}{2} \boldsymbol{\alpha} \cdot M \boldsymbol{\alpha}. \quad (7.2.1)$$

The two  $3 \times 3$  symmetric matrices (metrics) represent the moment of inertia ( $\mathbb{I}$ ) and the body mass matrix ( $M$ ) of the ellipsoidal underwater vehicle. The corresponding angular and linear momenta are

given as the pair of vectors

$$(\boldsymbol{\mu}, \boldsymbol{\beta}) = \left( \frac{\delta \ell}{\delta \boldsymbol{\xi}}, \frac{\delta \ell}{\delta \boldsymbol{\alpha}} \right) = (\mathbb{I}\boldsymbol{\xi}, \mathbb{M}\boldsymbol{\alpha}).$$

Consequently, the evolution of an ellipsoidal underwater vehicle is governed by a vector Euler–Poincaré equation on  $se(3)^*$  of the form (7.1.8),

$$\begin{aligned} (\mathbb{I}\boldsymbol{\xi})' &= -\boldsymbol{\xi} \times \mathbb{I}\boldsymbol{\xi} - \boldsymbol{\alpha} \times \mathbb{M}\boldsymbol{\alpha}, \\ (\mathbb{M}\boldsymbol{\alpha})' &= -\boldsymbol{\xi} \times \mathbb{M}\boldsymbol{\alpha}. \end{aligned} \quad (7.2.2)$$

**Remark 7.2.1** Of these vectors,  $\boldsymbol{\beta} = \mathbb{M}\boldsymbol{\alpha}$  is linear momentum and  $\boldsymbol{\mu} = \mathbb{I}\boldsymbol{\xi}$  is angular momentum. The Lagrangian  $l(\boldsymbol{\xi}, \boldsymbol{\alpha})$  in (7.2.1) is the kinetic energy for the motion of an ellipsoidal underwater vehicle.  $\square$

The Hamiltonian for the ellipsoidal underwater vehicle is the quadratic form

$$h(\boldsymbol{\mu}, \boldsymbol{\beta}) = \frac{1}{2}\boldsymbol{\mu} \cdot \mathbb{I}^{-1}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\beta} \cdot \mathbb{M}^{-1}\boldsymbol{\beta}. \quad (7.2.3)$$

The Lie–Poisson equations corresponding to (7.1.8) are

$$\begin{bmatrix} \dot{\boldsymbol{\mu}} \\ \dot{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu} \times & \boldsymbol{\beta} \times \\ \boldsymbol{\beta} \times & 0 \end{bmatrix} \begin{bmatrix} \partial h / \partial \boldsymbol{\mu} \\ \partial h / \partial \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu} \times & \boldsymbol{\beta} \times \\ \boldsymbol{\beta} \times & 0 \end{bmatrix} \begin{bmatrix} \mathbb{I}^{-1}\boldsymbol{\mu} \\ \mathbb{M}^{-1}\boldsymbol{\beta} \end{bmatrix}. \quad (7.2.4)$$

These are the Lie–Poisson equations for geodesic motion on  $se(3)^*$  with respect to the metric given by the Hamiltonian in (7.2.5).

**Remark 7.2.2** The stability of the equilibrium solutions of (7.2.2) for ellipsoidal underwater vehicles may be investigated within the present Euler–Poincaré framework. See [HoJeLe1998] and [GaMi1995].  $\square$

**Exercise.** Compute the Lie–Poisson equations for geodesic motion on  $se(2)^*$  with respect to a quadratic Hamiltonian metric.  $\star$

178 7 : EULER–POINCARÉ AND LIE–POISSON EQUATIONS ON  $SE(3)$ **7.2.1 Looks can be deceiving: The heavy top**

The Hamiltonian for the heavy top is the sum of its kinetic and potential energies,

$$h(\boldsymbol{\mu}, \boldsymbol{\beta}) = \underbrace{\frac{1}{2} \boldsymbol{\mu} \cdot \mathbb{I}^{-1} \boldsymbol{\mu}}_{\text{kinetic}} + \underbrace{mg\boldsymbol{\chi} \cdot \boldsymbol{\beta}}_{\text{potential}}, \quad (7.2.5)$$

in which  $\boldsymbol{\mu}$  is the body angular momentum,  $\boldsymbol{\chi}$  is the distance in the body from its point of support to its centre of mass,  $mg$  is its weight and  $\boldsymbol{\beta} = O^{-1}(t)\hat{\mathbf{z}}$  is the vertical direction, as seen from the body. The derivatives of this Hamiltonian are

$$\frac{\partial h}{\partial \boldsymbol{\mu}} = \boldsymbol{\xi} \quad \text{and} \quad \frac{\partial h}{\partial \boldsymbol{\beta}} = mg\boldsymbol{\chi}.$$

The correct equations of motion for the heavy-top emerge from this Hamiltonian and the Lie–Poisson bracket on  $se(3)^*$ :

$$\begin{aligned} \dot{\boldsymbol{\mu}} &= \{ \boldsymbol{\mu}, h \} = \boldsymbol{\mu} \times \frac{\partial h}{\partial \boldsymbol{\mu}} + \boldsymbol{\beta} \times \frac{\partial h}{\partial \boldsymbol{\beta}} \\ &= \boldsymbol{\mu} \times \boldsymbol{\xi} + \boldsymbol{\beta} \times mg\boldsymbol{\chi}, \end{aligned} \quad (7.2.6)$$

$$\dot{\boldsymbol{\beta}} = \{ \boldsymbol{\beta}, h \} = \boldsymbol{\beta} \times \frac{\partial h}{\partial \boldsymbol{\mu}} = \boldsymbol{\beta} \times \boldsymbol{\xi}. \quad (7.2.7)$$

It may seem surprising that the heavy-top dynamics would emerge from the  $SE(3)$  Lie–Poisson bracket, because the heavy top has no linear velocity or linear momentum. There is a story behind how this happened, which will be discussed in the next chapter.

**Exercise.** Does the Hamiltonian in (7.2.5) follow from the Legendre transformation of a Lagrangian  $\ell(\boldsymbol{\xi}, \boldsymbol{\alpha})$ ? If so, prove it. If not, what goes wrong? ★

**Exercise.** What are the equations corresponding to (7.2.6) and (7.2.7) for  $se(2)^*$ ? To what physical system do these equations correspond? ★

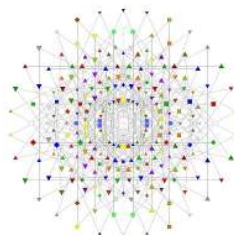
**Exercise.** Is the dynamical system governing  $(M, \mathbb{S})$  in Equations (2.5.25) and (2.5.26) Hamiltonian? Prove it. How is that system analogous to a heavy top? ★

**Exercise.** Write the Lie–Poisson Hamiltonian formulations for motion on the dual of each of the Lie algebras for

- the Heisenberg Lie group; and
- the semidirect-product Lie group  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$ . ★







# 8

## HEAVY-TOP EQUATIONS

---

### Contents

8.1	Introduction and definitions	182
8.2	Heavy-top action principle	183
8.3	Lie–Poisson brackets	184
8.3.1	Lie–Poisson brackets and momentum maps	185
8.3.2	Lie–Poisson brackets for the heavy top	186
8.4	Clebsch action principle	187
8.5	Kaluza–Klein construction	188

---

## 8.1 Introduction and definitions

A top is a rigid body of mass  $m$  rotating with a fixed point of support in a constant gravitational field of acceleration  $-g\hat{\mathbf{z}}$  pointing vertically downward. The orientation of the body relative to the vertical axis  $\hat{\mathbf{z}}$  is defined by the unit vector  $\mathbf{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$  for a curve  $\mathbf{R}(t) \in SO(3)$ . According to its definition, the unit vector  $\mathbf{\Gamma}$  represents the motion of the vertical direction as seen from the rotating body. Consequently, it satisfies the auxiliary motion equation,

$$\dot{\mathbf{\Gamma}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}(t)\mathbf{\Gamma} = -\widehat{\mathbf{\Omega}}(t)\mathbf{\Gamma} = \mathbf{\Gamma} \times \mathbf{\Omega}. \quad (8.1.1)$$

Here the rotation matrix  $\mathbf{R}(t) \in SO(3)$ , the skew matrix  $\widehat{\mathbf{\Omega}} = \mathbf{R}^{-1}\dot{\mathbf{R}} \in so(3)$  and the body angular frequency vector  $\mathbf{\Omega} \in \mathbb{R}^3$  are related by the hat map,  $\mathbf{\Omega} = (\mathbf{R}^{-1}\dot{\mathbf{R}})^\wedge$ , where

$$\text{hat map, } \wedge : (so(3), [\cdot, \cdot]) \rightarrow (\mathbb{R}^3, \times),$$

with  $\widehat{\mathbf{\Omega}}\mathbf{v} = \mathbf{\Omega} \times \mathbf{v}$  for any  $\mathbf{v} \in \mathbb{R}^3$ .

The motion of a top is determined from Euler's equations in vector form,

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega} + mg\mathbf{\Gamma} \times \boldsymbol{\chi}, \quad (8.1.2)$$

$$\dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}, \quad (8.1.3)$$

where  $\mathbf{\Omega}, \mathbf{\Gamma}, \boldsymbol{\chi} \in \mathbb{R}^3$  are vectors in the rotating body frame. Here

- $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$  is the body angular velocity vector.
- $\mathbb{I} = \text{diag}(I_1, I_2, I_3)$  is the moment of inertia tensor, diagonalised in the body principal axes.
- $\mathbf{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$  represents the motion of the unit vector along the vertical axis, as seen from the body.
- $\boldsymbol{\chi}$  is the constant vector in the body from the point of support to the body's centre of mass.
- $m$  is the total mass of the body and  $g$  is the constant acceleration of gravity.

## 8.2 Heavy-top action principle

**Proposition 8.2.1** *The heavy-top motion equation (8.1.2) is equivalent to the heavy-top action principle  $\delta S_{\text{red}} = 0$  for a reduced action,*

$$S_{\text{red}} = \int_a^b l(\mathbf{\Omega}, \mathbf{\Gamma}) dt = \int_a^b \frac{1}{2} \langle \mathbb{I}\mathbf{\Omega}, \mathbf{\Omega} \rangle - \langle mg\boldsymbol{\chi}, \mathbf{\Gamma} \rangle dt, \quad (8.2.1)$$

where variations of vectors  $\mathbf{\Omega}$  and  $\mathbf{\Gamma}$  are restricted to be of the form

$$\delta\mathbf{\Omega} = \dot{\boldsymbol{\Sigma}} + \mathbf{\Omega} \times \boldsymbol{\Sigma} \quad \text{and} \quad \delta\mathbf{\Gamma} = \mathbf{\Gamma} \times \boldsymbol{\Sigma}, \quad (8.2.2)$$

arising from variations of the corresponding definitions  $\hat{\boldsymbol{\Omega}} = \mathbf{R}^{-1}\dot{\mathbf{R}}$  and  $\mathbf{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$  in which  $\hat{\boldsymbol{\Sigma}}(t) = \mathbf{R}^{-1}\delta\mathbf{R}$  is a curve in  $\mathbb{R}^3$  that vanishes at the endpoints in time.

**Proof.** Since  $\mathbb{I}$  is symmetric and  $\boldsymbol{\chi}$  is constant, one finds the variation,

$$\begin{aligned} \delta \int_a^b l(\mathbf{\Omega}, \mathbf{\Gamma}) dt &= \int_a^b \langle \mathbb{I}\mathbf{\Omega}, \delta\mathbf{\Omega} \rangle - \langle mg\boldsymbol{\chi}, \delta\mathbf{\Gamma} \rangle dt \\ &= \int_a^b \langle \mathbb{I}\mathbf{\Omega}, \dot{\boldsymbol{\Sigma}} + \mathbf{\Omega} \times \boldsymbol{\Sigma} \rangle - \langle mg\boldsymbol{\chi}, \mathbf{\Gamma} \times \boldsymbol{\Sigma} \rangle dt \\ &= \int_a^b \left\langle -\frac{d}{dt}\mathbb{I}\mathbf{\Omega}, \boldsymbol{\Sigma} \right\rangle + \langle \mathbb{I}\mathbf{\Omega}, \mathbf{\Omega} \times \boldsymbol{\Sigma} \rangle - \langle mg\boldsymbol{\chi}, \mathbf{\Gamma} \times \boldsymbol{\Sigma} \rangle dt \\ &= \int_a^b \left\langle -\frac{d}{dt}\mathbb{I}\mathbf{\Omega} + \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega} + mg\boldsymbol{\chi} \times \mathbf{\Gamma}, \boldsymbol{\Sigma} \right\rangle dt, \end{aligned}$$

upon integrating by parts and using the endpoint conditions,  $\boldsymbol{\Sigma}(b) = \boldsymbol{\Sigma}(a) = 0$ . Since  $\boldsymbol{\Sigma}$  is otherwise arbitrary, (8.2.1) is equivalent to

$$-\frac{d}{dt}\mathbb{I}\mathbf{\Omega} + \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega} + mg\boldsymbol{\chi} \times \mathbf{\Gamma} = 0,$$

which is Euler's motion equation for the heavy top (8.1.2). This motion equation is completed by the auxiliary equation  $\dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}$  in (8.1.3) arising from the definition of  $\mathbf{\Gamma}$ . ■

## 184 8 : HEAVY-TOP EQUATIONS

The Legendre transformation for  $l(\Omega, \Gamma)$  gives the body angular momentum

$$\mathbf{\Pi} = \frac{\partial l}{\partial \Omega} = \mathbb{I}\Omega.$$

The well-known energy Hamiltonian for the heavy top then emerges as

$$h(\mathbf{\Pi}, \Gamma) = \mathbf{\Pi} \cdot \Omega - l(\Omega, \Gamma) = \frac{1}{2} \langle \mathbf{\Pi}, \mathbb{I}^{-1} \mathbf{\Pi} \rangle + \langle mg \chi, \Gamma \rangle, \quad (8.2.3)$$

which is the sum of the kinetic and potential energies of the top.

### The Lie–Poisson equations

Let  $f, h : \mathfrak{g}^* \rightarrow \mathbb{R}$  be real-valued functions on the dual space  $\mathfrak{g}^*$ . Denoting elements of  $\mathfrak{g}^*$  by  $\mu$ , the functional derivative of  $f$  at  $\mu$  is defined as the unique element  $\delta f / \delta \mu$  of  $\mathfrak{g}$  defined by

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [f(\mu + \varepsilon \delta \mu) - f(\mu)] = \left\langle \delta \mu, \frac{\delta f}{\delta \mu} \right\rangle, \quad (8.2.4)$$

for all  $\delta \mu \in \mathfrak{g}^*$ , where  $\langle \cdot, \cdot \rangle$  denotes the pairing between  $\mathfrak{g}^*$  and  $\mathfrak{g}$ .

## 8.3 Lie–Poisson brackets

**Definition 8.3.1 (Lie–Poisson equations)** The  $(\pm)$  Lie–Poisson brackets are defined by

$$\{f, h\}_{\pm}(\mu) = \pm \left\langle \mu, \left[ \frac{\delta f}{\delta \mu}, \frac{\delta h}{\delta \mu} \right] \right\rangle = \mp \left\langle \mu, \text{ad}_{\delta h / \delta \mu}^* \frac{\delta f}{\delta \mu} \right\rangle. \quad (8.3.1)$$

The corresponding Lie–Poisson equations, determined by  $\dot{f} = \{f, h\}$ , read

$$\dot{\mu} = \{\mu, h\} = \mp \text{ad}_{\delta h / \delta \mu}^* \mu, \quad (8.3.2)$$

where one defines the  $\text{ad}^*$  operation in terms of the pairing  $\langle \cdot, \cdot \rangle$ , by

$$\{f, h\} = \left\langle \mu, \text{ad}_{\delta h / \delta \mu} \frac{\delta f}{\delta \mu} \right\rangle = \left\langle \text{ad}_{\delta h / \delta \mu}^* \mu, \frac{\delta f}{\delta \mu} \right\rangle.$$

**Remark 8.3.1** The Lie–Poisson setting of mechanics is a special case of the general theory of systems on Poisson manifolds, for which there is now extensive theoretical development. (See [MaRa1994] for a start on this literature.)  $\square$

### 8.3.1 Lie–Poisson brackets and momentum maps

An important feature of the rigid-body bracket carries over to general Lie algebras. Namely, *Lie–Poisson brackets on  $\mathfrak{g}^*$  arise from canonical brackets on the cotangent bundle (phase space)  $T^*G$  associated with a Lie group  $G$  which has  $\mathfrak{g}$  as its associated Lie algebra. Thus, the process by which the Lie–Poisson brackets arise is the momentum map*

$$T^*G \mapsto \mathfrak{g}^*.$$

For example, a rigid body is free to rotate about its centre of mass and  $G$  is the (proper) rotation group  $SO(3)$ . The choice of  $T^*G$  as the primitive phase space is made according to the classical procedures of mechanics described earlier. For the description using Lagrangian mechanics, one forms the velocity phase space  $TG$ . The Hamiltonian description on  $T^*G$  is then obtained by standard procedures: Legendre transforms, etc.

The passage from  $T^*G$  to the space of  $\Pi$ 's (body angular momentum space) is determined by *left* translation on the group. This mapping is an example of a *momentum map*; that is, a mapping whose components are the “Noether quantities” associated with a symmetry group. That the map from  $T^*G$  to  $\mathfrak{g}^*$  is a Poisson map is a *general fact about momentum maps*. The Hamiltonian point of view of all this is a standard subject reviewed in Chapter 11.

**Remark 8.3.2 (Lie–Poisson description of the heavy top)** As it turns out, the underlying Lie algebra for the Lie–Poisson description of

## 186 8 : HEAVY-TOP EQUATIONS

the heavy top consists of the Lie algebra  $se(3, \mathbb{R})$  of infinitesimal Euclidean motions in  $\mathbb{R}^3$ . This is a bit surprising, because heavy-top motion itself does *not* actually arise through spatial translations by the Euclidean group; in fact, the body has a fixed point! Instead, the Lie algebra  $se(3, \mathbb{R})$  arises for another reason associated with the breaking of the  $SO(3)$  isotropy by the presence of the gravitational field. This symmetry breaking introduces a semidirect-product Lie–Poisson structure which happens to coincide with the dual of the Lie algebra  $se(3, \mathbb{R})$  in the case of the heavy top.  $\square$

### 8.3.2 Lie–Poisson brackets for the heavy top

The Lie algebra of the special Euclidean group in three dimensions is  $se(3) = \mathbb{R}^3 \times \mathbb{R}^3$  with the Lie bracket

$$[(\boldsymbol{\xi}, \mathbf{u}), (\boldsymbol{\eta}, \mathbf{v})] = (\boldsymbol{\xi} \times \boldsymbol{\eta}, \boldsymbol{\xi} \times \mathbf{v} - \boldsymbol{\eta} \times \mathbf{u}). \quad (8.3.3)$$

We identify the dual space with pairs  $(\boldsymbol{\Pi}, \boldsymbol{\Gamma})$ ; the corresponding  $(-)$  Lie–Poisson bracket called the *heavy-top bracket* is

$$\{f, h\}(\boldsymbol{\Pi}, \boldsymbol{\Gamma}) = -\boldsymbol{\Pi} \cdot \frac{\partial f}{\partial \boldsymbol{\Pi}} \times \frac{\partial h}{\partial \boldsymbol{\Pi}} - \boldsymbol{\Gamma} \cdot \left( \frac{\partial f}{\partial \boldsymbol{\Pi}} \times \frac{\partial h}{\partial \boldsymbol{\Gamma}} - \frac{\partial h}{\partial \boldsymbol{\Pi}} \times \frac{\partial f}{\partial \boldsymbol{\Gamma}} \right).$$

This Lie–Poisson bracket and the Hamiltonian (8.2.3) recover Equations (8.1.2) and (8.1.3) for the heavy top, as

$$\begin{aligned} \dot{\boldsymbol{\Pi}} = \{\boldsymbol{\Pi}, h\} &= \boldsymbol{\Pi} \times \frac{\partial h}{\partial \boldsymbol{\Pi}} + \boldsymbol{\Gamma} \times \frac{\partial h}{\partial \boldsymbol{\Gamma}} \\ &= \boldsymbol{\Pi} \times \mathbb{I}^{-1} \boldsymbol{\Pi} + \boldsymbol{\Gamma} \times mg \boldsymbol{\chi}, \end{aligned}$$

$$\dot{\boldsymbol{\Gamma}} = \{\boldsymbol{\Gamma}, h\} = \boldsymbol{\Gamma} \times \frac{\partial h}{\partial \boldsymbol{\Pi}} = \boldsymbol{\Gamma} \times \mathbb{I}^{-1} \boldsymbol{\Pi}.$$

**Remark 8.3.3 (Semidirect products and symmetry breaking)** The Lie algebra of the Euclidean group has a structure which is a special case of what is called a *semidirect product*. Here, it is the semidirect-product action  $so(3) \ltimes \mathbb{R}^3$  of the Lie algebra of rotations  $so(3)$  acting on the infinitesimal translations  $\mathbb{R}^3$ , which happens to coincide with  $se(3, \mathbb{R})$ .

In general, the Lie bracket for semidirect-product action  $\mathfrak{g} \ltimes V$  of a Lie algebra  $\mathfrak{g}$  on a vector space  $V$  is given by, cf. Equation (6.3.1),

$$\left[ (X, a), (\bar{X}, \bar{a}) \right] = \left( [X, \bar{X}], \bar{X}(a) - X(\bar{a}) \right),$$

in which  $X, \bar{X} \in \mathfrak{g}$  and  $a, \bar{a} \in V$ . Here, the action of the Lie algebra on the vector space is denoted, e.g.,  $X(\bar{a})$ . Usually, this action would be the Lie derivative.  $\square$

Lie–Poisson brackets defined on the dual spaces of semidirect-product Lie algebras tend to occur under rather general circumstances when the symmetry in  $T^*G$  is broken, e.g., reduced to an isotropy subgroup of a set of parameters. In particular, there are similarities in structure between the Poisson bracket for compressible flow and that for the heavy top. In the latter case, the vertical direction of gravity breaks the isotropy of  $\mathbb{R}^3$  from  $SO(3)$  to  $SO(2)$ . The general theory for semidirect products is reviewed in a variety of places, including [MaRaWe1984a, MaRaWe1984b].

Many interesting examples of Lie–Poisson brackets on semidirect products exist for fluid dynamics. These semidirect-product Lie–Poisson Hamiltonian theories range from simple fluids, to charged fluid plasmas, to magnetised fluids, to multiphase fluids, to super fluids, to Yang–Mills fluids, relativistic or not, and to liquid crystals. Many of these theories are discussed from the Euler–Poincaré viewpoint in [HoMaRa1998] and [Ho2002].

## 8.4 Clebsch action principle

**Proposition 8.4.1 (Clebsch heavy-top action principle)** *The heavy-top Equations (8.1.2) and (8.1.3) follow from a Clebsch constrained action principle,  $\delta S = 0$ , with*

$$S = \int_a^b \frac{1}{2} \langle \mathbb{I}\Omega, \Omega \rangle - \langle mg\chi, \Gamma \rangle + \langle \Xi, \dot{\Gamma} + \Omega \times \Gamma \rangle dt. \quad (8.4.1)$$

**Remark 8.4.1** The last term in this action is the *Clebsch constraint* for the auxiliary equation satisfied by the unit vector  $\Gamma$ . From its

## 188 8 : HEAVY-TOP EQUATIONS

definition  $\Gamma = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$  and the definition of the body angular velocity  $\Omega = \mathbf{R}^{-1}(t)\dot{\mathbf{R}}$ , this unit vector must satisfy

$$\dot{\Gamma} = -\mathbf{R}^{-1}\dot{\mathbf{R}}(t)\Gamma = -\hat{\Omega}(t)\Gamma = -\Omega \times \Gamma.$$

(The third equality invokes the hat map.) According to the Clebsch construction, the Lagrange multiplier  $\Xi$  enforcing the auxiliary Equation (8.4.1) will become the momentum canonically conjugate to the auxiliary variable  $\Gamma$ .  $\square$

**Proof.** The stationary variations of the constrained action (8.4.1) yield the following three *Clebsch relations*, cf. Equations (2.5.22) for the rigid body,

$$\begin{aligned} \delta\Omega : \quad \mathbb{I}\Omega + \Gamma \times \Xi &= 0, \\ \delta\Xi : \quad \dot{\Gamma} + \Omega \times \Gamma &= 0, \\ \delta\Gamma : \quad \dot{\Xi} + \Omega \times \Xi + mg\chi &= 0. \end{aligned}$$

As we shall see in Chapter 11, the first Clebsch relation defines the momentum map  $T^*\mathbb{R}^3 \rightarrow so(3)^*$  for the body angular momentum  $\mathbb{I}\Omega$ . From the other two Clebsch relations, the equation of motion for the body angular momentum may be computed as

$$\begin{aligned} \mathbb{I}\dot{\Omega} &= -\dot{\Gamma} \times \Xi - \Gamma \times \dot{\Xi} \\ &= (\Omega \times \Gamma) \times \Xi + \Gamma \times (\Omega \times \Xi + mg\chi) \\ &= \Omega \times (\Gamma \times \Xi) + \Gamma \times mg\chi \\ &= -\Omega \times (\mathbb{I}\Omega) + mg\Gamma \times \chi, \end{aligned}$$

which recovers Euler's motion Equation (8.1.2) for the heavy top.  $\blacksquare$

## 8.5 Kaluza–Klein construction

The Lagrangian in the heavy-top action principle (8.2.1) may be transformed into quadratic form. This is accomplished by suspending the system in a higher-dimensional space via the *Kaluza–Klein construction*. This construction proceeds for the heavy top as



a slight modification of the well-known Kaluza–Klein construction for a charged particle in a prescribed magnetic field.

Let  $Q_{KK}$  be the manifold  $SO(3) \times \mathbb{R}^3$  with variables  $(\mathbf{R}, \mathbf{q})$ . On  $Q_{KK}$  introduce the *Kaluza–Klein Lagrangian*

$$L_{KK} : TQ_{KK} \simeq TSO(3) \times T\mathbb{R}^3 \mapsto \mathbb{R},$$

as

$$\begin{aligned} L_{KK}(\mathbf{R}, \mathbf{q}, \dot{\mathbf{R}}, \dot{\mathbf{q}}; \hat{\mathbf{z}}) &= L_{KK}(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \mathbf{q}, \dot{\mathbf{q}}) \\ &= \frac{1}{2} \langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega} \rangle + \frac{1}{2} |\boldsymbol{\Gamma} + \dot{\mathbf{q}}|^2, \end{aligned} \quad (8.5.1)$$

with  $\boldsymbol{\Omega} = (\mathbf{R}^{-1}\dot{\mathbf{R}})^\wedge$  and  $\boldsymbol{\Gamma} = \mathbf{R}^{-1}\hat{\mathbf{z}}$ . The Lagrangian  $L_{KK}$  is positive-definite in  $(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \dot{\mathbf{q}})$ ; so it may be regarded as a kinetic energy which defines a metric, the *Kaluza–Klein metric* on  $TQ_{KK}$ .

The Legendre transformation for  $L_{KK}$  gives the momenta

$$\boldsymbol{\Pi} = \mathbb{I}\boldsymbol{\Omega} \quad \text{and} \quad \mathbf{p} = \boldsymbol{\Gamma} + \dot{\mathbf{q}}. \quad (8.5.2)$$

Since  $L_{KK}$  does not depend on  $\mathbf{q}$ , the Euler–Lagrange equation

$$\frac{d}{dt} \frac{\partial L_{KK}}{\partial \dot{\mathbf{q}}} = \frac{\partial L_{KK}}{\partial \mathbf{q}} = 0$$

shows that  $\mathbf{p} = \partial L_{KK} / \partial \dot{\mathbf{q}}$  is conserved. The *constant vector*  $\mathbf{p}$  is now identified as the vector in the body,

$$\mathbf{p} = \boldsymbol{\Gamma} + \dot{\mathbf{q}} = -mg\boldsymbol{\chi}.$$

After this identification, the heavy-top action principle in Proposition 8.2.1 with the Kaluza–Klein Lagrangian returns Euler’s motion equation for the heavy top (8.1.2).

The Hamiltonian  $H_{KK}$  associated with  $L_{KK}$  by the Legendre transformation (8.5.2) is

$$\begin{aligned} H_{KK}(\boldsymbol{\Pi}, \boldsymbol{\Gamma}, \mathbf{q}, \mathbf{p}) &= \boldsymbol{\Pi} \cdot \boldsymbol{\Omega} + \mathbf{p} \cdot \dot{\mathbf{q}} - L_{KK}(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \mathbf{q}, \dot{\mathbf{q}}) \\ &= \frac{1}{2} \boldsymbol{\Pi} \cdot \mathbb{I}^{-1} \boldsymbol{\Pi} - \mathbf{p} \cdot \boldsymbol{\Gamma} + \frac{1}{2} |\mathbf{p}|^2 \\ &= \frac{1}{2} \boldsymbol{\Pi} \cdot \mathbb{I}^{-1} \boldsymbol{\Pi} + \frac{1}{2} |\mathbf{p} - \boldsymbol{\Gamma}|^2 - \frac{1}{2} |\boldsymbol{\Gamma}|^2. \end{aligned}$$

## 190 8 : HEAVY-TOP EQUATIONS

Recall that  $\Gamma$  is a unit vector. On the constant level set  $|\Gamma|^2 = 1$ , the Kaluza–Klein Hamiltonian  $H_{KK}$  is a positive quadratic function, shifted by a constant. Likewise, on the constant level set  $\mathbf{p} = -mg\chi$ , the Kaluza–Klein Hamiltonian  $H_{KK}$  is a function of only the variables  $(\mathbf{\Pi}, \Gamma)$  and is equal to the Hamiltonian (8.2.3) for the heavy top up to an additive constant. As a result we have the following.

**Proposition 8.5.1** *The Lie–Poisson equations for the Kaluza–Klein Hamiltonian  $H_{KK}$  recover Euler’s equations for the heavy top, (8.1.2) and (8.1.3).*

**Proof.** The Lie–Poisson bracket may be written in matrix form explicitly as

$$\{f, h\} = \begin{bmatrix} \partial f / \partial \mathbf{\Pi} \\ \partial f / \partial \Gamma \\ \partial f / \partial \mathbf{q} \\ \partial f / \partial \mathbf{p} \end{bmatrix}^T \begin{bmatrix} \mathbf{\Pi} \times & \Gamma \times & 0 & 0 \\ \Gamma \times & 0 & 0 & 0 \\ 0 & 0 & 0 & Id \\ 0 & 0 & -Id & 0 \end{bmatrix} \begin{bmatrix} \partial h / \partial \mathbf{\Pi} \\ \partial h / \partial \Gamma \\ \partial h / \partial \mathbf{q} \\ \partial h / \partial \mathbf{p} \end{bmatrix}. \quad (8.5.3)$$

Consequently, one obtains the following Hamiltonian equations for  $h = H_{KK}(\mathbf{\Pi}, \Gamma, \mathbf{q}, \mathbf{p})$ ,

$$\begin{bmatrix} \dot{\mathbf{\Pi}} \\ \dot{\Gamma} \\ \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi} \times & \Gamma \times & 0 & 0 \\ \Gamma \times & 0 & 0 & 0 \\ 0 & 0 & 0 & Id \\ 0 & 0 & -Id & 0 \end{bmatrix} \begin{bmatrix} \mathbf{\Omega} \\ -\mathbf{p} \\ 0 \\ \mathbf{p} - \Gamma \end{bmatrix}. \quad (8.5.4)$$

These recover the heavy-top Equations (8.1.2) and (8.1.3) upon evaluating  $\mathbf{p} = -mg\chi$ . ■

**Exercise.** In an attempt to mimic Manakov’s beautiful idea for showing the integrability of the rigid body on  $SO(n)$ , one might imagine writing the three-dimensional heavy-top Equations (8.1.2) and (8.1.3) by inserting a spectral parameter  $\lambda$  as

$$\frac{d}{dt}(\Gamma + \lambda \mathbf{\Pi} + \lambda^2 \mathbf{J}) = (\Gamma + \lambda \mathbf{\Pi} + \lambda^2 \mathbf{J}) \times (\mathbf{\Omega} + \lambda \mathbf{K}),$$

with constant vectors  $\mathbf{J}$  and  $\mathbf{K}$  in  $\mathbb{R}^3$ . Does this formulation provide enough constants of motion to show the integrability of the heavy-top equations for some values of  $\chi$  and  $\mathbb{I}$ ? If so, which types of tops may be shown to be integrable this way? ★

**Answer.** The polynomial equation above implies the following relations, at various powers of  $\lambda$ :

$$\begin{aligned}\lambda^3 : \mathbf{J} \times \mathbf{K} = \mathbf{0} &\implies \mathbf{J} \parallel \mathbf{K}, \implies \mathbf{J} = \alpha \mathbf{K}, \alpha = \text{const.} \\ \lambda^2 : \dot{\mathbf{J}} = \mathbf{0} = \mathbb{I} \times \mathbf{K} + \mathbf{J} \times \mathbf{\Omega}, &\implies (\mathbb{I} \mathbf{\Omega} - \alpha \mathbf{\Omega}) \times \mathbf{K} = \mathbf{0}. \\ \lambda^1 : \dot{\mathbb{I}} = \mathbb{I} \times \mathbf{\Omega} + \mathbf{\Gamma} \times \mathbf{K}, &\implies \mathbf{K} = mg \boldsymbol{\chi}. \\ \lambda^0 : \dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}. &\end{aligned}$$

These relationships hold, provided the moment of inertia  $\mathbb{I}$  is either proportional to the identity (Euler top), or has two equal entries that make it cylindrically symmetric about the vector  $\boldsymbol{\chi}$  (Lagrange top).

This system conserves each of the coefficients of the powers of  $\lambda$  in  $|\mathbf{\Gamma} + \lambda \mathbb{I} + \lambda^2 \mathbf{J}|^2$ . That is, besides the kinematic constant  $|\mathbf{J}|^2$ , it conserves

$$|\mathbf{\Gamma}|^2, \quad \mathbf{\Gamma} \cdot \mathbb{I}, \quad \frac{1}{2\alpha} |\mathbb{I}|^2 + mg \mathbf{\Gamma} \cdot \boldsymbol{\chi}, \quad \mathbb{I} \cdot \boldsymbol{\chi}.$$

The first two are the Casimirs of the Lie–Poisson bracket in (8.3.4), the third is the Hamiltonian and the last is the  $\boldsymbol{\chi}$ -component of the angular momentum, which is conserved when the moment of inertia  $\mathbb{I}$  is cylindrically symmetric about the vector  $\boldsymbol{\chi}$ . This symmetry holds for the Euler top and the Lagrange top, which are indeed known to be integrable. For in-depth discussions of this approach to heavy-top dynamics, see [Ra1982, RaVM1982]. ▲

## 192 8 : HEAVY-TOP EQUATIONS

**Exercise.** Manakov's approach for the heavy top in the vector notation of the previous exercise suggests a similar application to the  $n \times n$  matrix commutator equation

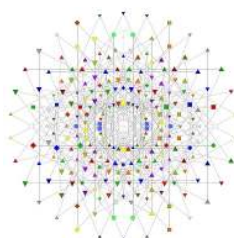
$$\frac{d}{dt}(\Gamma + \lambda\Pi + \lambda^2J) = [\Gamma + \lambda\Pi + \lambda^2J, \Omega + \lambda K]$$

with skew-symmetric  $(\Gamma, \Pi, \Omega, J, K)$  with constant  $(J, K)$ . Determine whether this approach could be used to extend Manakov's treatment of the rigid body in  $n$  dimensions to the  $n$ -dimensional versions of the Euler top and the Lagrange top. ★

**Exercise.** Extend the Manakov approach even further by computing the system of  $n \times n$  matrix equations for

$$\begin{aligned} \frac{d}{dt}(\Gamma + \lambda M + \lambda^2 N + \lambda^3 J) \\ = [\Gamma + \lambda M + \lambda^2 N + \lambda^3 J, \Omega + \lambda\omega + \lambda^2 K]. \end{aligned}$$

Is this extended matrix system Hamiltonian? If so, what is its Lie–Poisson bracket? ★



# 9

## THE EULER–POINCARÉ THEOREM

---

### Contents

9.1	Action principles on Lie algebras	194
9.2	Hamilton–Pontryagin principle	198
9.3	Clebsch approach to Euler–Poincaré	199
9.3.1	Defining the Lie derivative	201
9.3.2	Clebsch Euler–Poincaré principle	202
9.4	Lie–Poisson Hamiltonian formulation	206
9.4.1	Cotangent-lift momentum maps	207

---

## 9.1 Action principles on Lie algebras

Hamilton’s principle for stationary action was explained earlier for deriving Euler’s equations for rigid-body rotations in either their vector or quaternion forms. In the notation for the AD, Ad and ad actions of Lie groups and Lie algebras, Hamilton’s principle (that the equations of motion arise from stationarity of the action) for Lagrangians defined on Lie algebras may be expressed as follows. This is the Euler–Poincaré theorem [Po1901].

**Theorem 9.1.1 (Euler–Poincaré theorem)** *Stationarity*

$$\delta S(\xi) = \delta \int_a^b l(\xi) dt = 0 \quad (9.1.1)$$

*of an action*

$$S(\xi) = \int_a^b l(\xi) dt,$$

*whose Lagrangian is defined on the (left-invariant) Lie algebra  $\mathfrak{g}$  of a Lie group  $G$  by  $l(\xi) : \mathfrak{g} \mapsto \mathbb{R}$ , yields the **Euler–Poincaré equation** on  $\mathfrak{g}^*$ ,*

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \quad (9.1.2)$$

*for variations of the left-invariant Lie algebra element*

$$\xi = g^{-1} \dot{g}(t) \in \mathfrak{g}$$

*that are restricted to the form*

$$\delta \xi = \dot{\eta} + \text{ad}_\xi \eta, \quad (9.1.3)$$

*in which  $\eta(t) \in \mathfrak{g}$  is a curve in the Lie algebra  $\mathfrak{g}$  that vanishes at the endpoints in time.*

**Exercise.** What is the solution to the Euler–Poincaré Equation (9.1.2) in terms of  $\text{Ad}_{g(t)}^*$ ?

Hint: Take a look at the earlier equation (4.2.10). ★

**Remark 9.1.1** The earlier forms (4.1.21) and (4.1.23) of the variational formula for vectors and quaternions are now seen to apply more generally. Namely, such variations are defined for any Lie algebra. □

**Proof.** A direct computation proves Theorem 9.1.1. Later, we will explain the source of the constraint (9.1.3) on the form of the variations on the Lie algebra. One verifies the statement of the theorem by computing with a nondegenerate pairing  $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$ ,

$$\begin{aligned} 0 &= \delta \int_a^b l(\xi) dt = \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle dt \\ &= \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \dot{\eta} + \text{ad}_\xi \eta \right\rangle dt \\ &= \int_a^b \left\langle -\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \eta \right\rangle dt + \left\langle \frac{\delta l}{\delta \xi}, \eta \right\rangle \Big|_a^b, \end{aligned}$$

upon integrating by parts. The last term vanishes, by the endpoint conditions,  $\eta(b) = \eta(a) = 0$ .

Since  $\eta(t) \in \mathfrak{g}$  is otherwise arbitrary, (9.1.1) is equivalent to

$$-\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi} = 0,$$

which recovers the Euler–Poincaré Equation (9.1.2) in the statement of the theorem. ■

**Corollary 9.1.1 (Noether’s theorem for Euler–Poincaré)** *If  $\eta$  is an infinitesimal symmetry of the Lagrangian, then  $\langle \frac{\delta l}{\delta \xi}, \eta \rangle$  is its associated constant of the Euler–Poincaré motion.*

## 196 9 : THE EULER–POINCARÉ THEOREM

**Proof.** Consider the endpoint terms  $\langle \frac{\delta l}{\delta \xi}, \eta \rangle|_a^b$  arising in the variation  $\delta S$  in (9.1.1) and note that this implies for any time  $t \in [a, b]$  that

$$\left\langle \frac{\delta l}{\delta \xi(t)}, \eta(t) \right\rangle = \text{constant},$$

when the Euler–Poincaré Equations (9.1.2) are satisfied. ■

**Corollary 9.1.2 (Interpretation of Noether’s theorem)** *Noether’s theorem for the Euler–Poincaré stationary principle may be interpreted as conservation of the spatial momentum quantity*

$$\left( \text{Ad}_{g^{-1}(t)}^* \frac{\delta l}{\delta \xi(t)} \right) = \text{constant},$$

as a consequence of the Euler–Poincaré Equation (9.1.2).

**Proof.** Invoke left-invariance of the Lagrangian  $l(\xi)$  under  $g \rightarrow h_\epsilon g$  with  $h_\epsilon \in G$ . For this symmetry transformation, one has  $\delta g = \zeta g$  with  $\zeta = \frac{d}{d\epsilon}|_{\epsilon=0} h_\epsilon$ , so that

$$\eta = g^{-1} \delta g = \text{Ad}_{g^{-1}} \zeta \in \mathfrak{g}.$$

In particular, along a curve  $\eta(t)$  we have

$$\eta(t) = \text{Ad}_{g^{-1}(t)} \eta(0) \quad \text{on setting} \quad \zeta = \eta(0),$$

at any initial time  $t = 0$  (assuming of course that  $[0, t] \in [a, b]$ ). Consequently,

$$\left\langle \frac{\delta l}{\delta \xi(t)}, \eta(t) \right\rangle = \left\langle \frac{\delta l}{\delta \xi(0)}, \eta(0) \right\rangle = \left\langle \frac{\delta l}{\delta \xi(t)}, \text{Ad}_{g^{-1}(t)} \eta(0) \right\rangle.$$

For the nondegenerate pairing  $\langle \cdot, \cdot \rangle$ , this means that

$$\frac{\delta l}{\delta \xi(0)} = \left( \text{Ad}_{g^{-1}(t)}^* \frac{\delta l}{\delta \xi(t)} \right) = \text{constant}.$$

The constancy of this quantity under the Euler–Poincaré dynamics in (9.1.2) is verified, upon taking the time derivative and using the coadjoint motion relation (4.2.9) in Proposition 4.2.2. ■



**Remark 9.1.2** The form of the variation in (9.1.3) arises directly by

- (i) computing the variations of the left-invariant Lie algebra element  $\xi = g^{-1}\dot{g} \in \mathfrak{g}$  induced by taking variations  $\delta g$  in the group;
- (ii) taking the time derivative of the variation  $\eta = g^{-1}g' \in \mathfrak{g}$ ; and
- (iii) using the equality of cross derivatives ( $g'' = d^2g/dtds = g''$ ).

Namely, one computes, cf. Proposition (2.4.1) for the rigid body,

$$\xi' = (g^{-1}\dot{g})' = -g^{-1}g'g^{-1}\dot{g} + g^{-1}g'' = -\eta\xi + g^{-1}g'',$$

$$\dot{\eta} = (g^{-1}g')' = -g^{-1}\dot{g}g^{-1}g' + g^{-1}g'' = -\xi\eta + g^{-1}g''.$$

On taking the difference, the terms with cross derivatives cancel and one finds the variational formula (9.1.3),

$$\xi' - \dot{\eta} = [\xi, \eta] \quad \text{with} \quad [\xi, \eta] := \xi\eta - \eta\xi = \text{ad}_\xi \eta. \quad (9.1.4)$$

Thus, the same formal calculations as for vectors and quaternions also apply to Hamilton's principle on (matrix) Lie algebras.  $\square$

**Example 9.1.1 (Euler–Poincaré equation for  $SE(3)$ )** *The Euler–Poincaré Equation (9.1.2) for  $SE(3)$  is equivalent to*

$$\left( \frac{d}{dt} \frac{\delta l}{\delta \xi}, \frac{d}{dt} \frac{\delta l}{\delta \alpha} \right) = \text{ad}_{(\xi, \alpha)}^* \left( \frac{\delta l}{\delta \xi}, \frac{\delta l}{\delta \alpha} \right). \quad (9.1.5)$$

*This formula recovers the Euler–Poincaré Equation (7.1.7) for  $SE(3)$  upon using the definition of the  $\text{ad}^*$  operation for  $\mathfrak{se}(3)$  in Equation (6.3.2).*

**Remark 9.1.3** Corollary 9.1.2 is again the Kelvin–Noether Theorem 7.1.2, seen earlier for  $SE(3)$  and now proven for the Euler–Poincaré equations on an arbitrary Lie group.  $\square$

## 9.2 Hamilton–Pontryagin principle

Formula (9.1.4) for the variation of the vector  $\xi = g^{-1}\dot{g} \in \mathfrak{g}$  may be imposed as a constraint in Hamilton’s principle and thereby provide an immediate derivation of the Euler–Poincaré Equation (9.1.2). This constraint is incorporated into the following theorem [BoMa2009].

**Theorem 9.2.1 (Hamilton–Pontryagin principle)** *The Euler–Poincaré equation*

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} \quad (9.2.1)$$

*on the dual Lie algebra  $\mathfrak{g}^*$  is equivalent to the following implicit variational principle,*

$$\delta S(\xi, g, \dot{g}) = \delta \int_a^b l(\xi, g, \dot{g}) dt = 0, \quad (9.2.2)$$

*for a constrained action*

$$\begin{aligned} S(\xi, g, \dot{g}) &= \int_a^b l(\xi, g, \dot{g}) dt \\ &= \int_a^b \left[ l(\xi) + \langle \mu, (g^{-1}\dot{g} - \xi) \rangle \right] dt. \end{aligned} \quad (9.2.3)$$

**Proof.** The variations of  $S$  in formula (9.2.3) are given by

$$\delta S = \int_a^b \left\langle \frac{\delta l}{\delta \xi} - \mu, \delta \xi \right\rangle + \left\langle \delta \mu, (g^{-1}\dot{g} - \xi) \right\rangle + \left\langle \mu, \delta(g^{-1}\dot{g}) \right\rangle dt.$$

Substituting  $\delta(g^{-1}\dot{g})$  from (9.1.4) into the last term produces

$$\begin{aligned} \int_a^b \left\langle \mu, \delta(g^{-1}\dot{g}) \right\rangle dt &= \int_a^b \left\langle \mu, \dot{\eta} + \text{ad}_\xi \eta \right\rangle dt \\ &= \int_a^b \left\langle -\dot{\mu} + \text{ad}_\xi^* \mu, \eta \right\rangle dt + \left\langle \mu, \eta \right\rangle \Big|_a^b, \end{aligned}$$

## 9.3 CLEBSCH APPROACH TO EULER–POINCARÉ 199

where  $\eta = g^{-1}\delta g$  vanishes at the endpoints in time. Thus, stationarity  $\delta S = 0$  of the Hamilton–Pontryagin variational principle yields the following set of equations:

$$\frac{\delta l}{\delta \xi} = \mu, \quad g^{-1}\dot{g} = \xi, \quad \dot{\mu} = \text{ad}_\xi^* \mu. \quad (9.2.4)$$

■

**Remark 9.2.1 (Interpreting variational formulas (9.2.4))** The first formula in (9.2.4) is the fibre derivative needed in the Legendre transformation  $\mathfrak{g} \mapsto \mathfrak{g}^*$ , for passing to the Hamiltonian formulation. The second is the reconstruction formula for obtaining the solution curve  $g(t) \in G$  on the Lie group  $G$  given the solution  $\xi(t) = g^{-1}\dot{g} \in \mathfrak{g}$ . The third formula in (9.2.4) is the Euler–Poincaré equation on  $\mathfrak{g}^*$ . The interpretation of Noether’s theorem in Corollary 9.1.2 transfers to the Hamilton–Pontryagin variational principle as preservation of the quantity

$$\left( \text{Ad}_{g^{-1}(t)}^* \mu(t) \right) = \mu(0) = \text{constant},$$

under the Euler–Poincaré dynamics.

This Hamilton’s principle is said to be *implicit* because the definitions of the quantities describing the motion emerge only *after* the variations have been taken. See [YoMa2006] for discussions of recent developments in the theory of implicit variational principles.

□

**Exercise.** Compute the Euler–Poincaré equation on  $\mathfrak{g}^*$  when  $\xi(t) = \dot{g}g^{-1} \in \mathfrak{g}$  is *right-invariant*. ★

### 9.3 Clebsch approach to Euler–Poincaré

The Hamilton–Pontryagin (HP) Theorem 9.2.1 elegantly delivers the three key formulas in (9.2.4) needed for deriving the Lie–Poisson

## 200 9 : THE EULER–POINCARÉ THEOREM

Hamiltonian formulation of the Euler–Poincaré equation. Perhaps surprisingly, the HP theorem accomplishes this without invoking any properties of how the invariance group of the Lagrangian  $G$  acts on the configuration space  $M$ .

An alternative derivation of these formulas exists that uses the Clebsch approach and does invoke the action  $G \times M \rightarrow M$  of the Lie group on the configuration space,  $M$ , which is assumed to be a manifold. This alternative derivation is a bit more elaborate than the HP theorem. However, invoking the Lie group action on the configuration space provides additional valuable information. In particular, the alternative approach will yield information about the momentum map  $T^*M \mapsto \mathfrak{g}^*$  which explains precisely how the canonical phase space  $T^*M$  maps to the Poisson manifold of the dual Lie algebra  $\mathfrak{g}^*$ .

**Proposition 9.3.1 (Clebsch Euler–Poincaré principle)** *The Euler–Poincaré equation*

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} \quad (9.3.1)$$

on the dual Lie algebra  $\mathfrak{g}^*$  is equivalent to the following implicit variational principle,

$$\delta S(\xi, q, \dot{q}, p) = \delta \int_a^b l(\xi, q, \dot{q}, p) dt = 0, \quad (9.3.2)$$

for an action constrained by the reconstruction formula

$$\begin{aligned} S(\xi, q, \dot{q}, p) &= \int_a^b l(\xi, q, \dot{q}, p) dt \\ &= \int_a^b \left[ l(\xi) + \left\langle \left\langle p, \dot{q} + \mathcal{L}_\xi q \right\rangle \right\rangle \right] dt, \end{aligned} \quad (9.3.3)$$

in which the pairing  $\langle \cdot, \cdot \rangle : T^*M \times TM \mapsto \mathbb{R}$  maps an element of the cotangent space (a momentum covector) and an element from the tangent space (a velocity vector) to a real number. This is the natural pairing for an action integrand and it also occurs in the Legendre transformation.

**Remark 9.3.1** The Lagrange multiplier  $p$  in the second term of (9.3.3) imposes the constraint

$$\dot{q} + \mathcal{L}_\xi q = 0. \quad (9.3.4)$$

This is the formula for the evolution of the quantity  $q(t) = g^{-1}(t)q(0)$  under the left action of the Lie algebra element  $\xi \in \mathfrak{g}$  on it by the Lie derivative  $\mathcal{L}_\xi$  along  $\xi$ . (For right action by  $g$  so that  $q(t) = q(0)g(t)$ , the formula is  $\dot{q} - \mathcal{L}_\xi q = 0$ .)  $\square$

### 9.3.1 Defining the Lie derivative

One assumes the motion follows a trajectory  $q(t) \in M$  in the configuration space  $M$  given by  $q(t) = g(t)q(0)$ , where  $g(t) \in G$  is a time-dependent curve in the Lie group  $G$  which operates on the configuration space  $M$  by a flow  $\phi_t : G \times M \mapsto M$ . The flow property of the map  $\phi_t \circ \phi_s = \phi_{s+t}$  is guaranteed by the group composition law.

Just as for the free rotations, one defines the left-invariant and right-invariant velocity vectors. Namely, as for the body angular velocity,

$$\xi_L(t) = g^{-1}\dot{g}(t) \quad \text{is left-invariant under } g(t) \rightarrow hg(t),$$

and as for the spatial angular velocity,

$$\xi_R(t) = \dot{g}g^{-1}(t) \quad \text{is right-invariant under } g(t) \rightarrow g(t)h,$$

for any choice of matrix  $h \in G$ . This means neither of these velocities depends on the initial configuration.

#### **Right-invariant velocity vector**

The Lie derivative  $\mathcal{L}_\xi$  appearing in the reconstruction relation  $\dot{q} = -\mathcal{L}_\xi q$  in (9.3.4) is defined via the Lie group operation on the configuration space exactly as for free rotation. For example, one computes the tangent vectors to the motion induced by the group operation acting from the left as  $q(t) = g(t)q(0)$  by differentiating with respect to time  $t$ ,

$$\dot{q}(t) = \dot{g}(t)q(0) = \dot{g}g^{-1}(t)q(t) =: \mathcal{L}_{\xi_R}q(t),$$

where  $\xi_R = \dot{g}g^{-1}(t)$  is right-invariant. This is the analogue of the spatial angular velocity of a freely rotating rigid body.

## 202 9 : THE EULER–POINCARÉ THEOREM

**Left-invariant velocity vector**

Likewise, differentiating the right action  $q(t) = q(0)g(t)$  of the group on the configuration manifold yields

$$\dot{q}(t) = q(t)g^{-1}\dot{g}(t) =: \mathcal{L}_{\xi_L}q(t),$$

in which the quantity

$$\xi_L(t) = g^{-1}\dot{g}(t) = \text{Ad}_{g^{-1}(t)}\xi_R(t)$$

is the left-invariant tangent vector.

This analogy with free rotation dynamics should be a good guide for understanding the following manipulations, at least until we have a chance to illustrate the ideas with further examples.

**Exercise.** Compute the time derivatives and thus the forms of the right- and left-invariant velocity vectors for the group operations by the inverse  $q(t) = q(0)g^{-1}(t)$  and  $q(t) = g^{-1}(t)q(0)$ . Observe the equivalence (up to a sign) of these velocity vectors with the vectors  $\xi_R$  and  $\xi_L$ , respectively. Note that the reconstruction formula (9.3.4) arises from the latter choice. ★

**9.3.2 Clebsch Euler–Poincaré principle**

Let us first define the concepts and notation that will arise in the course of the proof of Proposition 9.3.1.

**Definition 9.3.1 (The diamond operation  $\diamond$ )** *The diamond operation ( $\diamond$ ) in Equation (9.3.8) is defined as minus the dual of the Lie derivative with respect to the pairing induced by the variational derivative in  $q$ , namely,*

$$\langle p \diamond q, \xi \rangle = \langle\langle p, -\mathcal{L}_\xi q \rangle\rangle. \quad (9.3.5)$$

**Definition 9.3.2 (Transpose of the Lie derivative)** *The transpose of the Lie derivative  $\mathcal{L}_\xi^T p$  is defined via the pairing  $\langle\langle \cdot, \cdot \rangle\rangle$  between  $(q, p) \in T^*M$  and  $(q, \dot{q}) \in TM$  as*

$$\langle\langle \mathcal{L}_\xi^T p, q \rangle\rangle = \langle\langle p, \mathcal{L}_\xi q \rangle\rangle. \quad (9.3.6)$$

**Proof.** The variations of the action integral

$$S(\xi, q, \dot{q}, p) = \int_a^b \left[ l(\xi) + \langle\langle p, \dot{q} + \mathcal{L}_\xi q \rangle\rangle \right] dt \quad (9.3.7)$$

from formula (9.3.3) are given by

$$\begin{aligned} \delta S &= \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle + \langle\langle \frac{\delta l}{\delta p}, \delta p \rangle\rangle + \langle\langle \frac{\delta l}{\delta q}, \delta q \rangle\rangle + \langle\langle p, \mathcal{L}_{\delta \xi} q \rangle\rangle dt \\ &= \int_a^b \left\langle \frac{\delta l}{\delta \xi} - p \diamond q, \delta \xi \right\rangle + \langle\langle \delta p, \dot{q} + \mathcal{L}_\xi q \rangle\rangle - \langle\langle \dot{p} - \mathcal{L}_\xi^T p, \delta q \rangle\rangle dt. \end{aligned}$$

Thus, stationarity of this implicit variational principle implies the following set of equations:

$$\frac{\delta l}{\delta \xi} = p \diamond q, \quad \dot{q} = -\mathcal{L}_\xi q, \quad \dot{p} = \mathcal{L}_\xi^T p. \quad (9.3.8)$$

In these formulas, the notation distinguishes between the two types of pairings,

$$\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R} \quad \text{and} \quad \langle\langle \cdot, \cdot \rangle\rangle : T^*M \times TM \mapsto \mathbb{R}. \quad (9.3.9)$$

(The third pairing in the formula for  $\delta S$  is not distinguished because it is equivalent to the second one under integration by parts in time.)

The Euler–Poincaré equation emerges from elimination of  $(q, p)$  using these formulas and the properties of the diamond operation

## 204 9 : THE EULER–POINCARÉ THEOREM

that arise from its definition, as follows, for any vector  $\eta \in \mathfrak{g}$ :

$$\begin{aligned} \left\langle \frac{d}{dt} \frac{\delta l}{\delta \xi}, \eta \right\rangle &= \frac{d}{dt} \left\langle \frac{\delta l}{\delta \xi}, \eta \right\rangle, \\ \text{[Definition of } \diamond] &= \frac{d}{dt} \left\langle p \diamond q, \eta \right\rangle = \frac{d}{dt} \left\langle \left\langle p, -\mathcal{L}_\eta q \right\rangle \right\rangle, \\ \text{[Equations (9.3.8)]} &= \left\langle \left\langle \mathcal{L}_\xi^T p, -\mathcal{L}_\eta q \right\rangle \right\rangle + \left\langle \left\langle p, \mathcal{L}_\eta \mathcal{L}_\xi q \right\rangle \right\rangle, \\ \text{[Transpose, } \diamond \text{ and ad]} &= \left\langle \left\langle p, -\mathcal{L}_{[\xi, \eta]} q \right\rangle \right\rangle = \left\langle p \diamond q, \text{ad}_\xi \eta \right\rangle, \\ \text{[Definition of ad}^*] &= \left\langle \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \eta \right\rangle. \end{aligned}$$

This is the Euler–Poincaré Equation (9.3.1). ■

**Exercise.** Show that the diamond operation defined in Equation (9.3.5) is antisymmetric,

$$\left\langle p \diamond q, \xi \right\rangle = - \left\langle q \diamond p, \xi \right\rangle. \quad (9.3.10)$$

★

**Exercise. (Euler–Poincaré equation for right action)** Compute the Euler–Poincaré equation for the Lie group action  $G \times M \mapsto M : q(t) = q(0)g(t)$  in which the group acts from the right on a point  $q(0)$  in the configuration manifold  $M$  along a time-dependent curve  $g(t) \in G$ . Explain why the result differs in sign from the case of left  $G$ -action on manifold  $M$ . ★

**Exercise. (Clebsch approach for motion on  $T^*(G \times V)$ )** Often the Lagrangian will contain a parameter taking values



in a vector space  $V$  that represents a feature of the potential energy of the motion. We have encountered this situation already with the heavy top, in which the parameter is the vector in the body pointing from the contact point to the centre of mass. Since the potential energy will affect the motion we assume an action  $G \times V \rightarrow V$  of the Lie group  $G$  on the vector space  $V$ . The Lagrangian then takes the form  $L : TG \times V \rightarrow \mathbb{R}$ .

Compute the variations of the action integral

$$S(\xi, q, \dot{q}, p) = \int_a^b \left[ \tilde{l}(\xi, q) + \left\langle \left\langle p, \dot{q} + \mathcal{L}_\xi q \right\rangle \right\rangle \right] dt$$

and determine the effects in the Euler–Poincaré equation of having  $q \in V$  appear in the Lagrangian  $\tilde{l}(\xi, q)$ .

Show first that stationarity of  $S$  implies the following set of equations:

$$\frac{\delta \tilde{l}}{\delta \xi} = p \diamond q, \quad \dot{q} = -\mathcal{L}_\xi q, \quad \dot{p} = \mathcal{L}_\xi^T p + \frac{\delta \tilde{l}}{\delta q}.$$

Then transform to the variable  $\delta l / \delta \xi$  to find the associated Euler–Poincaré equations on the space  $\mathfrak{g}^* \times V$ ,

$$\begin{aligned} \frac{d}{dt} \frac{\delta \tilde{l}}{\delta \xi} &= \text{ad}_\xi^* \frac{\delta \tilde{l}}{\delta \xi} + \frac{\delta \tilde{l}}{\delta q} \diamond q, \\ \frac{dq}{dt} &= -\mathcal{L}_\xi q. \end{aligned}$$

Perform the Legendre transformation to derive the Lie–Poisson Hamiltonian formulation corresponding to  $\tilde{l}(\xi, q)$ .



## 9.4 Lie–Poisson Hamiltonian formulation

The Clebsch variational principle for the Euler–Poincaré equation provides a natural path to its canonical and Lie–Poisson Hamiltonian formulations. The Legendre transform takes the Lagrangian

$$l(p, q, \dot{q}, \xi) = l(\xi) + \left\langle\left\langle p, \dot{q} + \mathcal{L}_\xi q \right\rangle\right\rangle$$

in the action (9.3.7) to the Hamiltonian,

$$H(p, q) = \left\langle\left\langle p, \dot{q} \right\rangle\right\rangle - l(p, q, \dot{q}, \xi) = \left\langle\left\langle p, -\mathcal{L}_\xi q \right\rangle\right\rangle - l(\xi),$$

whose variations are given by

$$\begin{aligned} \delta H(p, q) &= \left\langle\left\langle \delta p, -\mathcal{L}_\xi q \right\rangle\right\rangle + \left\langle\left\langle p, -\mathcal{L}_\xi \delta q \right\rangle\right\rangle \\ &\quad + \left\langle\left\langle p, -\mathcal{L}_{\delta\xi} q \right\rangle\right\rangle - \left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle \\ &= \left\langle\left\langle \delta p, -\mathcal{L}_\xi q \right\rangle\right\rangle + \left\langle\left\langle -\mathcal{L}_\xi^T p, \delta q \right\rangle\right\rangle + \left\langle p \diamond q - \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle. \end{aligned}$$

These variational derivatives recover Equations (9.3.8) in canonical Hamiltonian form,

$$\dot{q} = \delta H / \delta p = -\mathcal{L}_\xi q \quad \text{and} \quad \dot{p} = -\delta H / \delta q = \mathcal{L}_\xi^T p.$$

Moreover, independence of  $H$  from  $\xi$  yields the momentum relation,

$$\frac{\delta l}{\delta \xi} = p \diamond q. \quad (9.4.1)$$

The Legendre transformation of the Euler–Poincaré equations using the Clebsch canonical variables leads to the *Lie–Poisson Hamiltonian form* of these equations,

$$\frac{d\mu}{dt} = \{\mu, h\} = \text{ad}_{\delta h / \delta \mu}^* \mu, \quad (9.4.2)$$

with

$$\mu = p \diamond q = \frac{\delta l}{\delta \xi}, \quad h(\mu) = \langle \mu, \xi \rangle - l(\xi), \quad \xi = \frac{\delta h}{\delta \mu}. \quad (9.4.3)$$

By Equation (9.4.3), the evolution of a smooth real function  $f : \mathfrak{g}^* \rightarrow \mathbb{R}$  is governed by

$$\begin{aligned} \frac{df}{dt} &= \left\langle \frac{\delta f}{\delta \mu}, \frac{d\mu}{dt} \right\rangle \\ &= \left\langle \frac{\delta f}{\delta \mu}, \text{ad}_{\delta h / \delta \mu}^* \mu \right\rangle \\ &= \left\langle \text{ad}_{\delta h / \delta \mu} \frac{\delta f}{\delta \mu}, \mu \right\rangle \\ &= - \left\langle \mu, \left[ \frac{\delta f}{\delta \mu}, \frac{\delta h}{\delta \mu} \right] \right\rangle \\ &=: \{f, h\}. \end{aligned} \quad (9.4.4)$$

The last equality defines the *Lie–Poisson bracket*  $\{f, h\}$  for smooth real functions  $f$  and  $h$  on the dual Lie algebra  $\mathfrak{g}^*$ . One may check directly that this bracket operation is a bilinear, skew-symmetric derivation that satisfies the Jacobi identity. Thus, it defines a proper Poisson bracket on  $\mathfrak{g}^*$ .

### 9.4.1 Cotangent-lift momentum maps

Although it is more elaborate than the Hamilton–Pontryagin principle and it requires input about the action of a Lie algebra on the configuration space, the Clebsch variational principle for the Euler–Poincaré equation reveals useful information.

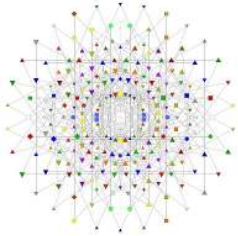
As we shall see, the Clebsch approach provides a direct means of computing the *momentum map* for the specified Lie algebra action on a given configuration manifold  $M$ . In fact, the first equation in (9.4.3) is the standard example of the momentum map obtained by the *cotangent lift* of a Lie algebra action on a configuration manifold.

**208** 9 : THE EULER–POINCARÉ THEOREM

Momentum maps will be discussed later, in Chapter 11. For now, the reader may wish to notice that the formulas (9.4.3) and (11.2.1) involving the diamond operation have remarkable similarities. In particular, the term  $q \diamond p$  in these formulas has the same meaning. Consequently, we may state the following proposition.

**Proposition 9.4.1 (Momentum maps)** *The Lie–Poisson form (9.4.2) of the Euler–Poincaré Equation (9.3.1) governs the evolution of the momentum map derived from the cotangent lift of the Lie algebra action on the configuration manifold.*

The remainder of the present text should provide the means to fully understand this statement.



# 10

## LIE–POISSON HAMILTONIAN FORM OF A CONTINUUM SPIN CHAIN

---

### Contents

10.1	Formulating continuum spin chain equations	210
10.2	Euler–Poincaré equations	212
10.3	Hamiltonian formulation	213

---

## 210 10 : LIE-POISSON HAMILTONIAN FORM OF A CONTINUUM SPIN CHAIN

In this chapter we will begin thinking in terms of Hamiltonian partial differential equations in the specific example of *G-strands*, which are evolutionary maps into a Lie group  $g(t, x) : \mathbb{R} \times \mathbb{R} \rightarrow G$  that follow from Hamilton's principle for a certain class of  $G$ -invariant Lagrangians. The case when  $G = SO(3)$  may be regarded physically as a smooth distribution of  $so(3)$ -valued spins attached to a one-dimensional straight strand lying along the  $x$ -axis. We will investigate its three-dimensional orientation dynamics at each point along the strand. For no additional cost, we may begin with the Euler–Poincaré theorem for a left-invariant Lagrangian defined on the tangent space of an *arbitrary* Lie group  $G$  and later specialise to the case where  $G$  is the rotation group  $SO(3)$ .

The Lie–Poisson Hamiltonian formulation of the Euler–Poincaré Equation (9.3.1) for this problem will be derived via the Legendre Transformation by following calculations similar to those done previously for the rigid body in Section 2.5. To emphasise the systematic nature of the Legendre transformation from the Euler–Poincaré picture to the Lie–Poisson picture, we will lay out the procedure in well-defined steps.

### 10.1 Formulating continuum spin chain equations

We shall consider Hamilton's principle  $\delta S = 0$  for a left-invariant Lagrangian,

$$S = \int_a^b \int_{-\infty}^{\infty} \ell(\Omega, \Xi) dx dt, \quad (10.1.1)$$

with the following definitions of the tangent vectors  $\Omega$  and  $\Xi$ ,

$$\Omega(t, x) = g^{-1} \partial_t g(t, x) \quad \text{and} \quad \Xi(t, x) = g^{-1} \partial_x g(t, x), \quad (10.1.2)$$

where  $g(t, x) \in G$  is a real-valued map  $g : \mathbb{R} \times \mathbb{R} \rightarrow G$  for a Lie group  $G$ . Later, we shall specialise to the case where  $G$  is the rotation group  $SO(3)$ . We shall apply the by now standard Euler–Poincaré procedure, modulo the partial spatial derivative in the definition of  $\Xi(t, x) = g^{-1} \partial_x g(t, x) \in \mathfrak{g}$ . This procedure takes the following steps:

(i) Write the auxiliary equation for the evolution of  $\Xi : \mathbb{R} \times \mathbb{R} \rightarrow \mathfrak{g}$ , obtained by differentiating its definition with respect to time and invoking equality of cross derivatives.

(ii) Use the Euler–Poincaré theorem for left-invariant Lagrangians to obtain the equation of motion for the momentum variable  $\partial\ell/\partial\Omega : \mathbb{R} \times \mathbb{R} \rightarrow \mathfrak{g}^*$ , where  $\mathfrak{g}^*$  is the dual Lie algebra. Use the  $L^2$  pairing defined by the spatial integration.

(These will be partial differential equations. Assume homogeneous boundary conditions on  $\Omega(t, x)$ ,  $\Xi(t, x)$  and vanishing endpoint conditions on the variation  $\eta = g^{-1}\delta g(t, x) \in \mathfrak{g}$  when integrating by parts.)

(iii) Legendre-transform this Lagrangian to obtain the corresponding Hamiltonian. Differentiate the Hamiltonian and determine its partial derivatives. Write the Euler–Poincaré equation in terms of the new momentum variable  $\Pi = \delta\ell/\delta\Omega \in \mathfrak{g}^*$ .

(iv) Determine the Lie–Poisson bracket implied by the Euler–Poincaré equation in terms of the Legendre-transformed quantities  $\Pi = \delta\ell/\delta\Omega$ , by rearranging the time derivative of a smooth function  $f(\Pi, \Xi) : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$ .

(v) Specialise to  $G = SO(3)$  and write the Lie–Poisson Hamiltonian form in terms of vector operations in  $\mathbb{R}^3$ .

(vi) For  $G = SO(3)$  choose the Lagrangian

$$\begin{aligned} \ell &= \frac{1}{2} \int_{-\infty}^{\infty} \text{Tr} \left( [g^{-1}\partial_t g, g^{-1}\partial_x g]^2 \right) dx \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \text{Tr} \left( [\Omega, \Xi]^2 \right) dx, \end{aligned} \quad (10.1.3)$$

## 212 10 : LIE-POISSON HAMILTONIAN FORM OF A CONTINUUM SPIN CHAIN

where  $[\Omega, \Xi] = \Omega\Xi - \Xi\Omega$  is the commutator in the Lie algebra  $\mathfrak{g}$ . Use the hat map to write the Euler–Poincaré equation and its Lie–Poisson Hamiltonian form in terms of vector operations in  $\mathbb{R}^3$ .

## 10.2 Euler–Poincaré equations

The Euler–Poincaré procedure systematically produces the following results.

**Auxiliary equations** By definition,  $\Omega(t, x) = g^{-1}\partial_t g(t, x)$  and  $\Xi(t, x) = g^{-1}\partial_x g(t, x)$  are Lie-algebra-valued functions over  $\mathbb{R} \times \mathbb{R}$ . The evolution of  $\Xi$  is obtained from these definitions by taking the difference of the two equations for the partial derivatives

$$\begin{aligned}\partial_t \Xi(t, x) &= -(g^{-1}\partial_t g)(g^{-1}\partial_x g) + g^{-1}\partial_t \partial_x g(t, x), \\ \partial_x \Omega(t, x) &= -(g^{-1}\partial_x g)(g^{-1}\partial_t g) + g^{-1}\partial_x \partial_t g(t, x),\end{aligned}$$

and invoking equality of cross derivatives. Hence,  $\Xi$  evolves by the adjoint operation, much like in the derivation of the variational derivative of  $\Omega$ ,

$$\partial_t \Xi(t, x) - \partial_x \Omega(t, x) = \Xi\Omega - \Omega\Xi = [\Xi, \Omega] =: -\text{ad}_\Omega \Xi. \quad (10.2.1)$$

This is the auxiliary equation for  $\Xi(t, x)$ . In differential geometry, this relation is called a **zero curvature relation**, because it implies that the curvature vanishes for the Lie-algebra-valued connection one-form  $A = \Omega dt + \Xi dx$  [doCa1976].

**Hamilton’s principle** For  $\eta = g^{-1}\delta g(t, x) \in \mathfrak{g}$ , Hamilton’s principle  $\delta S = 0$  for  $S = \int_a^b \ell(\Omega, \Xi) dt$  leads to

$$\delta S = \int_a^b \left\langle \frac{\delta \ell}{\delta \Omega}, \delta \Omega \right\rangle + \left\langle \frac{\delta \ell}{\delta \Xi}, \delta \Xi \right\rangle dt$$



$$\begin{aligned}
&= \int_a^b \left\langle \frac{\delta \ell}{\delta \Omega}, \partial_t \eta + \text{ad}_\Omega \eta \right\rangle + \left\langle \frac{\delta \ell}{\delta \Xi}, \partial_x \eta + \text{ad}_\Xi \eta \right\rangle dt \\
&= \int_a^b \left\langle -\partial_t \frac{\delta \ell}{\delta \Omega} + \text{ad}_\Omega^* \frac{\delta \ell}{\delta \Omega}, \eta \right\rangle + \left\langle -\partial_x \frac{\delta \ell}{\delta \Xi} + \text{ad}_\Xi^* \frac{\delta \ell}{\delta \Xi}, \eta \right\rangle dt \\
&= \int_a^b \left\langle -\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \Omega} + \text{ad}_\Omega^* \frac{\delta \ell}{\delta \Omega} - \frac{\partial}{\partial x} \frac{\delta \ell}{\delta \Xi} + \text{ad}_\Xi^* \frac{\delta \ell}{\delta \Xi}, \eta \right\rangle dt,
\end{aligned}$$

where the formulas for the variations  $\delta \Omega$  and  $\delta \Xi$  are obtained by essentially the same calculation as in part (i). Hence,  $\delta S = 0$  yields

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \Omega} = \text{ad}_\Omega^* \frac{\delta \ell}{\delta \Omega} - \frac{\partial}{\partial x} \frac{\delta \ell}{\delta \Xi} + \text{ad}_\Xi^* \frac{\delta \ell}{\delta \Xi}. \quad (10.2.2)$$

This is the Euler–Poincaré equation for  $\delta \ell / \delta \Omega \in \mathfrak{g}^*$ .

**Exercise.** Use Equation (4.2.9) in Proposition 4.2.2 to show that the Euler–Poincaré Equation (10.2.2) is a *conservation law* for spin angular momentum  $\Pi = \delta \ell / \delta \Omega$ ,

$$\frac{\partial}{\partial t} \left( \text{Ad}_{g(t,x)}^* \frac{\delta \ell}{\delta \Omega} \right) = - \frac{\partial}{\partial x} \left( \text{Ad}_{g(t,x)}^* \frac{\delta \ell}{\delta \Xi} \right). \quad (10.2.3)$$

★

### 10.3 Hamiltonian formulation

**Legendre transform** Legendre-transforming the Lagrangian  $\ell(\Omega, \Xi) : \mathfrak{g} \times V \rightarrow \mathbb{R}$  yields the Hamiltonian  $h(\Pi, \Xi) : \mathfrak{g}^* \times V \rightarrow \mathbb{R}$ ,

$$h(\Pi, \Xi) = \langle \Pi, \Omega \rangle - \ell(\Omega, \Xi). \quad (10.3.1)$$

**214 10 : LIE-POISSON HAMILTONIAN FORM OF A CONTINUUM SPIN CHAIN**

Differentiating the Hamiltonian determines its partial derivatives:

$$\begin{aligned}
 \delta h &= \left\langle \delta \Pi, \frac{\delta h}{\delta \Pi} \right\rangle + \left\langle \frac{\delta h}{\delta \Xi}, \delta \Xi \right\rangle \\
 &= \left\langle \delta \Pi, \Omega \right\rangle + \left\langle \Pi - \frac{\delta l}{\delta \Omega}, \delta \Omega \right\rangle - \left\langle \frac{\delta \ell}{\delta \Xi}, \delta \Xi \right\rangle \\
 \Rightarrow \frac{\delta l}{\delta \Omega} &= \Pi, \quad \frac{\delta h}{\delta \Pi} = \Omega \quad \text{and} \quad \frac{\delta h}{\delta \Xi} = -\frac{\delta \ell}{\delta \Xi}.
 \end{aligned}$$

The middle term vanishes because  $\Pi - \delta l/\delta \Omega = 0$  defines  $\Pi$ . These derivatives allow one to rewrite the Euler–Poincaré equation solely in terms of momentum  $\Pi$  as

$$\begin{aligned}
 \partial_t \Pi &= \text{ad}_{\delta h/\delta \Pi}^* \Pi + \partial_x \frac{\delta h}{\delta \Xi} - \text{ad}_{\Xi}^* \frac{\delta h}{\delta \Xi}, \\
 \partial_t \Xi &= \partial_x \frac{\delta h}{\delta \Pi} - \text{ad}_{\delta h/\delta \Pi} \Xi.
 \end{aligned} \tag{10.3.2}$$

**Hamiltonian equations** The corresponding Hamiltonian equation for any functional of  $f(\Pi, \Xi)$  is then

$$\begin{aligned}
 \frac{\partial}{\partial t} f(\Pi, \Xi) &= \left\langle \partial_t \Pi, \frac{\delta f}{\delta \Pi} \right\rangle + \left\langle \partial_t \Xi, \frac{\delta f}{\delta \Xi} \right\rangle \\
 &= \left\langle \text{ad}_{\delta h/\delta \Pi}^* \Pi + \partial_x \frac{\delta h}{\delta \Xi} - \text{ad}_{\Xi}^* \frac{\delta h}{\delta \Xi}, \frac{\delta f}{\delta \Pi} \right\rangle \\
 &\quad + \left\langle \partial_x \frac{\delta h}{\delta \Pi} - \text{ad}_{\delta h/\delta \Pi} \Xi, \frac{\delta f}{\delta \Xi} \right\rangle \\
 &= -\left\langle \Pi, \left[ \frac{\delta f}{\delta \Pi}, \frac{\delta h}{\delta \Pi} \right] \right\rangle \\
 &\quad + \left\langle \partial_x \frac{\delta h}{\delta \Xi}, \frac{\delta f}{\delta \Pi} \right\rangle - \left\langle \partial_x \frac{\delta f}{\delta \Xi}, \frac{\delta h}{\delta \Pi} \right\rangle \\
 &\quad + \left\langle \Xi, \text{ad}_{\delta f/\delta \Pi}^* \frac{\delta h}{\delta \Xi} - \text{ad}_{\delta h/\delta \Pi}^* \frac{\delta f}{\delta \Xi} \right\rangle \\
 &=: \{f, h\}(\Pi, \Xi).
 \end{aligned}$$

Assembling these equations into Hamiltonian form gives, symbolically,

$$\frac{\partial}{\partial t} \begin{bmatrix} \Pi \\ \Xi \end{bmatrix} = \begin{bmatrix} \text{ad}_{\square}^* \Pi & (\text{div} - \text{ad}_{\Xi}^*) \square \\ (\text{grad} - \text{ad}_{\square}) \Xi & 0 \end{bmatrix} \begin{bmatrix} \delta h/\delta \Pi \\ \delta h/\delta \Xi \end{bmatrix} \tag{10.3.3}$$

The boxes  $\square$  in Equation (10.3.3) indicate how the  $\text{ad}$  and  $\text{ad}^*$  operations are applied in the matrix multiplication. For example,

$$\text{ad}_{\square}^* \Pi(\delta h / \delta \Pi) = \text{ad}_{\delta h / \delta \Pi}^* \Pi,$$

so each matrix entry acts on its corresponding vector component.<sup>1</sup>

**Higher dimensions** Although it is beyond the scope of the present text, we shall make a few short comments about the meaning of the terms appearing in the Hamiltonian matrix (10.3.3). First, the notation indicates that the natural jump to higher dimensions has been made. This is done by using the spatial gradient to define the left-invariant auxiliary variable  $\Xi \equiv g^{-1} \nabla g$  in higher dimensions. The lower left entry of the matrix (10.3.3) defines the *covariant spatial gradient*, and its upper right entry defines the adjoint operator, the *covariant spatial divergence*. More explicitly, in terms of indices and partial differential operators, this Hamiltonian matrix becomes,

$$\frac{\partial}{\partial t} \begin{bmatrix} \Pi_\alpha \\ \Xi_i^\alpha \end{bmatrix} = B_{\alpha\beta} \begin{bmatrix} \delta h / \delta \Pi_\beta \\ \delta h / \delta \Xi_j^\beta \end{bmatrix}, \quad (10.3.4)$$

where the Hamiltonian structure matrix  $B_{\alpha\beta}$  is given explicitly as

$$B_{\alpha\beta} = \begin{bmatrix} -\Pi_\kappa t_{\alpha\beta}^\kappa & \delta_\alpha^\beta \partial_j + t_{\alpha\kappa}^\beta \Xi_j^\kappa \\ \delta_\beta^\alpha \partial_i - t_{\beta\kappa}^\alpha \Xi_i^\kappa & 0 \end{bmatrix}. \quad (10.3.5)$$

Here, the summation convention is enforced on repeated indices. Superscript Greek indices refer to the Lie algebraic basis set, subscript Greek indices refer to the dual basis and Latin indices refer to the spatial reference frame. The partial derivative  $\partial_j = \partial / \partial x_j$ , say, acts to the right on all terms in a product by the chain rule.

<sup>1</sup>This is the lower right corner of the Hamiltonian matrix for a perfect complex fluid [Ho2002, GBRa2008]. It also appears in the Lie–Poisson brackets for Yang–Mills fluids [GiHoKu1982] and for spin glasses [HoKu1988].

216 10 : LIE-POISSON HAMILTONIAN FORM OF A CONTINUUM SPIN CHAIN

**Lie–Poisson bracket** For the case that  $t_{\beta\kappa}^\alpha$  are structure constants for the Lie algebra  $so(3)$ , then  $t_{\beta\kappa}^\alpha = \epsilon_{\alpha\beta\kappa}$  with  $\epsilon_{123} = +1$ . By using the hat map (2.1.11), the Lie–Poisson Hamiltonian matrix in (10.3.5) may be rewritten for the  $so(3)$  case in  $\mathbb{R}^3$  vector form as

$$\frac{\partial}{\partial t} \begin{bmatrix} \mathbf{\Pi} \\ \mathbf{\Xi}_i \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi} \times & \partial_j + \mathbf{\Xi}_j \times \\ \partial_i + \mathbf{\Xi}_i \times & 0 \end{bmatrix} \begin{bmatrix} \delta h / \delta \mathbf{\Pi} \\ \delta h / \delta \mathbf{\Xi}_j \end{bmatrix}. \quad (10.3.6)$$

Returning to one dimension, stationary solutions  $\partial_t \rightarrow 0$  and spatially independent solutions  $\partial_x \rightarrow 0$  both satisfy equations of the same  $se(3)$  form as the heavy top. For example, the time-independent solutions satisfy, with  $\mathbf{\Omega} = \delta h / \delta \mathbf{\Pi}$  and  $\mathbf{\Lambda} = \delta h / \delta \mathbf{\Xi}$ ,

$$\frac{d}{dx} \mathbf{\Lambda} = -\mathbf{\Xi} \times \mathbf{\Lambda} - \mathbf{\Pi} \times \mathbf{\Omega} \quad \text{and} \quad \frac{d}{dx} \mathbf{\Omega} = -\mathbf{\Xi} \times \mathbf{\Omega}.$$

That the equations have the same form is to be expected because of the exchange symmetry under  $t \leftrightarrow x$  and  $\mathbf{\Omega} \leftrightarrow \mathbf{\Xi}$ . Perhaps less expected is that the heavy-top form reappears.

For  $G = SO(3)$  and the Lagrangian  $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  in one spatial dimension  $\ell(\mathbf{\Omega}, \mathbf{\Xi})$  the Euler–Poincaré equation and its Hamiltonian form are given in terms of vector operations in  $\mathbb{R}^3$ , as follows. First, the Euler–Poincaré Equation (10.2.2) becomes

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{\Omega}} = -\mathbf{\Omega} \times \frac{\delta \ell}{\delta \mathbf{\Omega}} - \frac{\partial}{\partial x} \frac{\delta \ell}{\delta \mathbf{\Xi}} - \mathbf{\Xi} \times \frac{\delta \ell}{\delta \mathbf{\Xi}}. \quad (10.3.7)$$

**Choices for the Lagrangian**

- Interesting choices for the Lagrangian include those symmetric under exchange of  $\mathbf{\Omega}$  and  $\mathbf{\Xi}$ , such as

$$\ell_{\perp} = |\mathbf{\Omega} \times \mathbf{\Xi}|^2 / 2 \quad \text{and} \quad \ell_{\parallel} = (\mathbf{\Omega} \cdot \mathbf{\Xi})^2 / 2,$$

for which the variational derivatives are, respectively,

$$\begin{aligned} \frac{\delta \ell_{\perp}}{\delta \mathbf{\Omega}} &= \mathbf{\Xi} \times (\mathbf{\Omega} \times \mathbf{\Xi}) =: |\mathbf{\Xi}|^2 \mathbf{\Omega}_{\perp}, \\ \frac{\delta \ell_{\perp}}{\delta \mathbf{\Xi}} &= \mathbf{\Omega} \times (\mathbf{\Xi} \times \mathbf{\Omega}) =: |\mathbf{\Omega}|^2 \mathbf{\Xi}_{\perp}, \end{aligned}$$

for  $\ell_{\perp}$  and the complementary quantities,

$$\begin{aligned}\frac{\delta \ell_{\parallel}}{\delta \Omega} &= (\Omega \cdot \Xi) \Xi =: |\Xi|^2 \Omega_{\parallel}, \\ \frac{\delta \ell_{\parallel}}{\delta \Xi} &= (\Omega \cdot \Xi) \Omega =: |\Omega|^2 \Xi_{\parallel},\end{aligned}$$

for  $\ell_{\parallel}$ . With either of these choices,  $\ell_{\perp}$  or  $\ell_{\parallel}$ , Equation (10.3.7) becomes a local conservation law for spin angular momentum

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \Omega} = - \frac{\partial}{\partial x} \frac{\delta \ell}{\delta \Xi}.$$

The case  $\ell_{\perp}$  is reminiscent of the *Skyrme model* [Sk1961], a nonlinear topological model of pions in nuclear physics.

- Another interesting choice for  $G = SO(3)$  and the Lagrangian  $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  in one spatial dimension is

$$\ell(\Omega, \Xi) = \frac{1}{2} \int_{-\infty}^{\infty} \Omega \cdot \mathbb{A} \Omega + \Xi \cdot \mathbb{B} \Xi \, dx,$$

for symmetric matrices  $\mathbb{A}$  and  $\mathbb{B}$ , which may also be  $L^2$ -symmetric differential operators. In this case the variational derivatives are given by

$$\delta \ell(\Omega, \Xi) = \int_{-\infty}^{\infty} \delta \Omega \cdot \mathbb{A} \Omega + \delta \Xi \cdot \mathbb{B} \Xi \, dx,$$

and the Euler–Poincaré Equation (10.2.2) becomes

$$\frac{\partial}{\partial t} \mathbb{A} \Omega + \Omega \times \mathbb{A} \Omega + \frac{\partial}{\partial x} \mathbb{B} \Xi + \Xi \times \mathbb{B} \Xi = 0. \quad (10.3.8)$$

This is the sum of two coupled rotors, one in space and one in time, again suggesting the one-dimensional spin glass, or spin chain. When  $\mathbb{A}$  and  $\mathbb{B}$  are taken to be the identity, Equation (10.3.8) recovers the *chiral model*, or *sigma model*, which is completely integrable, cf. [Wi1984, ZaMi1980].

218 10 : LIE-POISSON HAMILTONIAN FORM OF A CONTINUUM SPIN CHAIN

**Hamiltonian structures** The Hamiltonian structures of these equations on  $so(3)^*$  are obtained from the Legendre-transform relations

$$\frac{\delta \ell}{\delta \Omega} = \mathbf{\Pi}, \quad \frac{\delta h}{\delta \mathbf{\Pi}} = \Omega \quad \text{and} \quad \frac{\delta h}{\delta \Xi} = -\frac{\delta \ell}{\delta \Xi}.$$

Hence, the Euler–Poincaré Equation (10.2.2) becomes

$$\frac{\partial}{\partial t} \mathbf{\Pi} = \mathbf{\Pi} \times \frac{\delta h}{\delta \mathbf{\Pi}} + \frac{\partial}{\partial x} \frac{\delta h}{\delta \Xi} + \Xi \times \frac{\delta h}{\delta \Xi}, \quad (10.3.9)$$

and the auxiliary Equation (10.3.10) becomes

$$\frac{\partial}{\partial t} \Xi = \frac{\partial}{\partial x} \frac{\delta h}{\delta \mathbf{\Pi}} + \Xi \times \frac{\delta h}{\delta \mathbf{\Pi}}, \quad (10.3.10)$$

which recovers the Lie–Poisson structure in Equation (10.3.6).

Finally, the reconstruction equations may be expressed using the hat map as

$$\begin{aligned} \partial_t O(t, x) &= O(t, x) \widehat{\Omega}(t, x) \quad \text{and} \\ \partial_x O(t, x) &= O(t, x) \widehat{\Xi}(t, x). \end{aligned} \quad (10.3.11)$$

**Remark 10.3.1** The Euler–Poincaré equations for the continuum spin chain discussed here and their Lie–Poisson Hamiltonian formulation provide a framework for systematically investigating three-dimensional orientation dynamics along a one-dimensional strand. These partial differential equations are interesting in their own right and they have many possible applications. For an idea of where the applications of these equations could lead, consult [SiMaKr1988,EGHPR2010].

□

**Exercise.** Write the Euler–Poincaré equations of the continuum spin chain for  $SE(3)$ , in which each point is both rotating and translating. Recall that

$$\left( \frac{d}{dt} \frac{\delta l}{\delta \xi}, \frac{d}{dt} \frac{\delta l}{\delta \alpha} \right) = \text{ad}_{(\xi, \alpha)}^* \left( \frac{\delta l}{\delta \xi}, \frac{\delta l}{\delta \alpha} \right), \quad (10.3.12)$$

where the  $\text{ad}^*$  operation of the Lie algebra  $\mathfrak{se}(3)$  on its dual  $\mathfrak{se}(3)^*$  is given in Equations (6.3.2) and (6.3.3).

Apply formula (10.3.12) to express the space-time Euler–Poincaré Equation (10.2.2) for  $SE(3)$  in vector form.

Complete the computation of the Lie–Poisson Hamiltonian form for the continuum spin chain on  $SE(3)$ . ★

**Exercise.** Let the set of  $2 \times 2$  matrices  $M_i$  with  $i = 1, 2, 3$  satisfy the defining relation for the symplectic Lie group  $Sp(2)$ ,

$$M_i J M_i^T = J \quad \text{with} \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (10.3.13)$$

The corresponding elements of its Lie algebra  $m_i = \dot{M}_i M_i^{-1} \in \mathfrak{sp}(2)$  satisfy  $(J m_i)^T = J m_i$  for each  $i = 1, 2, 3$ . Thus,  $X_i = J m_i$  satisfying  $X_i^T = X_i$  is a set of three symmetric  $2 \times 2$  matrices. Define  $X = J \dot{M} M^{-1}$  with time derivative  $\dot{M} = \partial M(t, x) / \partial t$  and  $Y = J M' M^{-1}$  with space derivative  $M' = \partial M(t, x) / \partial x$ . Then show that

$$X' = \dot{Y} + [X, Y]_J, \quad (10.3.14)$$

for the J-bracket defined by

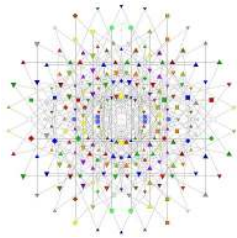
$$[X, Y]_J := X J Y - Y J X =: 2 \text{sym}(X J Y) =: \text{ad}_X^J Y.$$

In terms of the J-bracket, compute the continuum Euler–Poincaré equations for a Lagrangian  $\ell(X, Y)$  defined on the symplectic Lie algebra  $\mathfrak{sp}(2)$ .

Compute the Lie–Poisson Hamiltonian form of the system comprising the continuum Euler–Poincaré equations on  $\mathfrak{sp}(2)^*$  and the compatibility equation (10.3.14) on  $\mathfrak{sp}(2)$ . ★







# 11

## MOMENTUM MAPS

---

### Contents

11.1	The momentum map	222
11.2	Cotangent lift	224
11.3	Examples of momentum maps	226
11.3.1	The Poincaré sphere $S^2 \in S^3$	237
11.3.2	Overview	242

---

## 11.1 The momentum map

A momentum map  $J : M \mapsto \mathfrak{g}^*$  arises when the smooth Lie group action of  $G$  on a manifold  $M$  preserves either the symplectic structure, or the Poisson structure on  $M$ . Here  $\mathfrak{g}$  is the Lie algebra of  $G$  and  $\mathfrak{g}^*$  is its dual. We concentrate on the situation in which  $M = T^*Q$  is the cotangent bundle of a configuration manifold  $Q$  on which the Lie group  $G$  acts smoothly.

**Example 11.1.1** *An example of a momentum map  $J : M \mapsto \mathfrak{g}^*$  is the quantity  $J(p, q)$  defined by*

$$\begin{aligned} J^\xi(p, q) &:= \left\langle J(p, q), \xi \right\rangle_{\mathfrak{g}^* \times \mathfrak{g}} \\ &= \left\langle\left\langle (p, q), \xi_Q(q) \right\rangle\right\rangle_{T^*Q \times TQ} \\ &= \left\langle\left\langle p, \mathcal{L}_\xi q \right\rangle\right\rangle_{T^*Q \times TQ}. \end{aligned} \quad (11.1.1)$$

*In this formula, the infinitesimal action of a Lie group  $G$  by an element  $\xi \in \mathfrak{g}$  of its Lie algebra is expressed as a Lie derivative  $\mathcal{L}_\xi q = \xi_Q(q)$  on the configuration space  $Q$ . The momentum map lifts this expression into phase space  $M = T^*Q$  by expressing it as a Hamiltonian vector field  $X_{J^\xi(p, q)} = \{ \cdot, J^\xi(p, q) \}$  on  $T^*Q$ .*

Denote the action of the Lie group  $G$  on the configuration manifold  $Q$  as  $q(s) = g(s)q(0)$  for  $g(s) \in G$ ,  $s \in \mathbb{R}$  and  $q \in Q$ . As usual, a vector field  $\xi_Q(q) \in TQ$  at a point  $q \in Q$  is obtained by differentiating  $q(s) = g(s)q(0)$  with respect to  $s$  in the direction  $\xi$  at the identity  $s = 0$ , where  $g(0) = e$ . That is,

$$\xi_Q(q) = q'(s)|_{s=0} = g'(s)q(0)|_{s=0} = (g'g^{-1})|_{s=0}q(0) =: \xi q,$$

for  $q = q(0)$ . In other notation, the vector field  $\xi_Q(q) \in TQ$  may be expressed as a Lie derivative,

$$\xi_Q(q) = \frac{d}{ds} [\exp(s\xi)q] \Big|_{s=0} =: \mathcal{L}_\xi q = \xi q \in TQ. \quad (11.1.2)$$

**Remark 11.1.1** The formula for  $\xi_Q(q) \in TQ$  is the *tangent lift* of the action of  $G$  on  $Q$  at  $q \in Q$ . The tangent lift action of  $G$  on  $TQ$  induces an action of  $G$  on  $T^*Q$  by the *cotangent lift* (the inverse transpose of the tangent lift). The cotangent lift action of  $G$  on  $T^*Q$  is always symplectic; so it may be written using canonical coordinates  $(p, q) \in T^*Q$  as a Hamiltonian vector field  $X_{J^\xi} = \{ \cdot, J^\xi(p, q) \}$ .  $\square$

For the case when the symplectic manifold  $M$  is the cotangent bundle  $T^*Q$  of a configuration manifold  $Q$ , the quantity

$$J^\xi(p, q) := \left\langle J(p, q), \xi \right\rangle_{\mathfrak{g}^* \times \mathfrak{g}}$$

is the Hamiltonian on  $T^*Q$ . The canonical Poisson bracket

$$X_{J^\xi} = \{ \cdot, J^\xi(p, q) \}$$

is the Hamiltonian vector field  $X_{J^\xi}$  for the infinitesimal action of the Lie group  $G$  on  $Q$  (configuration space), lifted to the cotangent bundle  $T^*Q$  (phase space) with symplectic form  $\omega = dq \wedge dp$ . Equivalently,

$$dJ^\xi(p, q) = X_{J^\xi} \lrcorner \omega = \omega(X_{J^\xi}, \cdot).$$

This property defines the standard momentum map.

**Definition 11.1.1 (Standard momentum map)** *The standard momentum map  $J : T^*Q \mapsto \mathfrak{g}^*$  is defined by requiring*

$$X_{J^\xi(p, q)} = \xi_{(p, q)} \quad \text{for each } \xi \in \mathfrak{g}. \quad (11.1.3)$$

*That is, the infinitesimal generator  $\xi_{(p, q)}$  of the cotangent lift action of  $G$  for each element  $\xi \in \mathfrak{g}$  is equal to the Hamiltonian vector field*

$$X_{J^\xi(p, q)} = \{ \cdot, J^\xi(p, q) \}, \quad (11.1.4)$$

*of the function  $J^\xi : T^*Q \mapsto \mathbb{R}$ . The momentum map  $J : T^*Q \mapsto \mathfrak{g}^*$  is defined as  $J(p, q)$ , satisfying*

- $J^\xi(p, q) := \langle J(p, q), \xi \rangle$  for each  $\xi \in \mathfrak{g}$  and any  $(p, q) \in T^*Q$ ;
- $\langle \cdot, \cdot \rangle$  is the natural pairing  $\mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$ ; and
- $\{ \cdot, \cdot \} : \mathcal{F}(p, q) \times \mathcal{F}(p, q) \mapsto \mathcal{F}(p, q)$  is the canonical Poisson bracket on  $T^*Q$ .

## 11.2 Cotangent lift

**Theorem 11.2.1 (Cotangent-lift momentum map)** *The cotangent-lift momentum map  $J : T^*Q \mapsto \mathfrak{g}^*$  satisfies*

$$\begin{aligned}
 J^\xi(p, q) &:= \langle J(p, q), \xi \rangle \\
 &= \langle\langle (p, q), \xi_Q(q) \rangle\rangle \\
 &= \langle\langle p, \mathcal{L}_\xi q \rangle\rangle \\
 &=: \langle q \diamond p, \xi \rangle. \tag{11.2.1}
 \end{aligned}$$

**Proof.** The first equality repeats the definition of  $J^\xi(p, q)$ . The second equality inserts the definition of the infinitesimal action  $\xi_Q(q) \in TQ$  of the Lie group  $G$  on  $Q$  at the point  $q \in Q$ . The pairing  $\langle\langle \cdot, \cdot \rangle\rangle : T^*Q \times TQ \mapsto \mathbb{R}$  in this equality is between the tangent and cotangent spaces of the configuration  $Q$ . The third equality inserts the definition of the infinitesimal action  $\xi_Q(q) = \mathcal{L}_\xi q$  in terms of the Lie derivative. The last equality provides the required Hamiltonian  $J^\xi(p, q) = \langle q \diamond p, \xi \rangle$  for the Hamiltonian vector field  $X_{J^\xi} = \{ \cdot, J^\xi(p, q) \}$  in the cotangent lift action of  $G$  on  $T^*Q$  by defining the diamond operation ( $\diamond$ ) in terms of the two pairings and the Lie derivative. ■

**Remark 11.2.1** The diamond operation was introduced in (6.2.9) and in the Clebsch Equation (9.3.8), where it was defined using the dual of the Lie derivative with respect to the  $T^*Q \times TQ$  pairing induced by the variational derivative in  $q$ , namely,

$$\langle q \diamond p, \xi \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = \langle\langle p, \mathcal{L}_\xi q \rangle\rangle_{T^*Q \times TQ}. \tag{11.2.2}$$

Thus, as discussed in Section 9.3.2, the variational relation in the Clebsch procedure associated with the dynamical Clebsch

constraint, or reconstruction Equation (9.3.4), defines the corresponding cotangent-lift momentum map in Equation (9.4.3). (The sign is correct because  $\diamond$  is antisymmetric.)  $\square$

**Theorem 11.2.2 (Hamiltonian Noether's theorem)** *If the Hamiltonian  $H(p, q)$  on  $T^*Q$  is invariant under the action of the Lie group  $G$ , then  $J^\xi(p, q)$  is conserved on trajectories of the corresponding Hamiltonian vector field,*

$$X_H = \{ \cdot, H(p, q) \}.$$

**Proof.** Differentiating the invariance condition  $H(gp, gq) = H(p, q)$  with respect to  $g$  for fixed  $(p, q) \in T^*Q$  yields

$$\begin{aligned} \mathcal{L}_\xi H(p, q) &= dH(p, q) \cdot \xi_{(p, q)} = 0 = X_{J^\xi(p, q)} H(p, q) \\ &= -\{J^\xi, H\}(p, q) = -X_{H(p, q)} J^\xi(p, q). \end{aligned}$$

Consequently, the momentum map  $J^\xi(p, q)$  is conserved on trajectories of the Hamiltonian vector field  $X_H = \{ \cdot, H(p, q) \}$  for a  $G$ -invariant Hamiltonian.  $\blacksquare$

**Proposition 11.2.1 (Equivariant group actions)**

- A group action  $\Phi_g : G \times T^*Q \mapsto T^*Q$  is said to be **equivariant** if it satisfies

$$J \circ \Phi_g = \text{Ad}_{g^{-1}}^* \circ J.$$

This means the following diagram commutes:

$$\begin{array}{ccc} P & \xrightarrow{\Phi_{g(t)}} & P \\ J \downarrow & & \downarrow J \\ \mathfrak{g}^* & \xrightarrow{\text{Ad}_{g(t)^{-1}}^*} & \mathfrak{g}^* \end{array}$$

## 226 11 : MOMENTUM MAPS

- *Equivariance implies infinitesimal equivariance. Namely,*

$$\left. \frac{d}{dt} \right|_{t=0} J(\Phi_{g(t)}(z)) = \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{g^{-1}}^* \circ J(z)$$

*implies*

$$dJ(z) \cdot \xi_P(z) = -\text{ad}_\xi^* J(z),$$

*with  $z = (p, q)$ . Setting  $dJ(z) \cdot \xi_P(z) = X_{J\xi}J$  and pairing with a fixed Lie algebra element  $\eta$  yields the  $\eta$ -component:*

$$\begin{aligned} \langle dJ(z) \cdot \xi_P(z), \eta \rangle &= \langle -\text{ad}_\xi^* J(z), \eta \rangle, \\ X_{J\xi}J^\eta &= \langle J(z), -\text{ad}_\xi \eta \rangle, \\ \{J^\eta(z), J^\xi(z)\} &= \langle J(z), [\eta, \xi] \rangle. \end{aligned} \quad (11.2.3)$$

- *Consequently, infinitesimal equivariance implies*

$$\left\{ \langle J(p, q), \eta \rangle, \langle J(p, q), \xi \rangle \right\} = \langle J(p, q), [\eta, \xi] \rangle. \quad (11.2.4)$$

*This means that the map  $(\mathfrak{g}, [\cdot, \cdot]) \rightarrow (C^\infty(T^*Q), \{\cdot, \cdot\})$  defined by  $\xi \mapsto J^\xi$ ,  $\xi \in \mathfrak{g}$  is a **Lie algebra homomorphism** (i.e., it preserves bracket relations).*

- *Infinitesimal equivariance implies that the momentum map*

$$J : T^*Q \mapsto \mathfrak{g}^* \quad \text{is a Poisson map.}$$

*That is,  $J$  corresponding to left (resp., right) group action produces a + (resp., -) Lie–Poisson bracket on  $\mathfrak{g}^*$ .*

### 11.3 Examples of momentum maps

**Example 11.3.1 (Momentum map for  $SO(3)$  acting on  $\mathbb{R}^3$ )** For  $Q = \mathbb{R}^3$  and  $\mathfrak{g} = \mathfrak{so}(3)$  one finds  $\xi_Q(q) = \mathcal{L}_\xi q = \xi \times q$  by the hat map and

$$\langle\langle p, \mathcal{L}_\xi q \rangle\rangle = p \cdot \xi \times q = (q \times p) \cdot \xi = \langle J(p, q), \xi \rangle = J^\xi(p, q),$$

which is the Hamiltonian for an infinitesimal rotation around  $\xi$  in  $\mathbb{R}^3$ . In the case that  $\mathfrak{g} = \mathfrak{so}(3)$ , the pairings  $\langle \cdot, \cdot \rangle$  and  $\langle\langle \cdot, \cdot \rangle\rangle$  may both be taken as dot products of vectors in  $\mathbb{R}^3$ , the momentum map  $J(p, q) = q \diamond p = q \times p \in \mathbb{R}^3$  is the phase-space expression for angular momentum and the  $\diamond$  operation is  $\times$ , the cross product of vectors in  $\mathbb{R}^3$ . This is an example of a **cotangent-lift** momentum map.

**Example 11.3.2 (Momentum map for  $SU(2)$  acting on  $\mathbb{C}^2$ )** The Lie group  $SU(2)$  of complex  $2 \times 2$  unitary matrices  $U(s)$  with unit determinant acts on  $\mathbf{a} \in \mathbb{C}^2$  by matrix multiplication as

$$\mathbf{a}(s) = U(s)\mathbf{a}(0) = \exp(is\xi)\mathbf{a}(0),$$

in which  $i\xi = U'U^{-1}|_{s=0}$  is a  $2 \times 2$  traceless skew-Hermitian matrix, as seen from the following:

$$UU^\dagger = \text{Id} \quad \text{implies} \quad U'U^\dagger + UU'^\dagger = 0 = U'U^\dagger + (U'U^\dagger)^\dagger.$$

Likewise,  $\xi$  alone (that is, not multiplied by  $i$ ) is a  $2 \times 2$  traceless Hermitian matrix.

The infinitesimal generator  $\xi(\mathbf{a}) \in \mathbb{C}^2$  may be expressed as a linear transformation,

$$\xi(\mathbf{a}) = \frac{d}{ds} [\exp(is\xi)\mathbf{a}] \Big|_{s=0} = i\xi\mathbf{a},$$

in which the product  $(\xi\mathbf{a})$  of the Hermitian matrix  $(\xi)$  and the two-component complex vector  $(\mathbf{a})$  has components  $\xi_{kl}a_l$ , with  $k, l = 1, 2$ .

To be a momentum map,  $J : \mathbb{C}^2 \mapsto \mathfrak{su}(2)^*$  must satisfy the defining relation (11.2.1),

$$\begin{aligned} J^\xi(\mathbf{a}) &:= \left\langle J(\mathbf{a}), \xi \right\rangle_{\mathfrak{su}(2)^* \times \mathfrak{su}(2)} = \left\langle\left\langle \mathbf{a}, \xi(\mathbf{a}) \right\rangle\right\rangle_{\mathbb{C}^2} = \left\langle\left\langle \mathbf{a}, i\xi\mathbf{a} \right\rangle\right\rangle_{\mathbb{C}^2} \\ &= \text{Im}(a_k^*(i\xi)_{kl}a_l) = a_k^*\xi_{kl}a_l = \text{tr}((\mathbf{a} \otimes \mathbf{a}^*)\xi) = \text{tr}(Q^\dagger\xi). \end{aligned}$$

Being traceless,  $\xi$  has zero pairing with any multiple of the identity; so one may subtract the trace of  $Q = \mathbf{a} \otimes \mathbf{a}^*$ . Thus, the traceless Hermitian quantity

$$J(\mathbf{a}) = Q - \frac{1}{2}(\text{tr } Q)\text{Id} = \mathbf{a} \otimes \mathbf{a}^* - \frac{1}{2}\text{Id}|\mathbf{a}|^2 \in \mathfrak{su}(2)^* \quad (11.3.1)$$

## 228 11 : MOMENTUM MAPS

defines a momentum map  $J : \mathbb{C}^2 \mapsto su(2)^*$ . That is,  $J$  maps  $\mathbf{a} \in \mathbb{C}^2$  to the traceless Hermitian matrix  $J(\mathbf{a})$ , which is an element of  $su(2)^*$ , the dual space to  $su(2)$  under the pairing  $\langle \cdot, \cdot \rangle : su(2)^* \times su(2) \mapsto \mathbb{R}$  given by the trace of the matrix product,

$$\langle J, \xi \rangle_{su(2)^* \times su(2)} = \text{tr}(J(\mathbf{a})^\dagger \xi), \quad (11.3.2)$$

$$\text{for } J(\mathbf{a}) = J(\mathbf{a})^\dagger \in su(2)^* \text{ and } i\xi \in su(2). \quad (11.3.3)$$

**Proposition 11.3.1 (Momentum map equivariance)** Let  $U \in SU(2)$  and  $\mathbf{a} \in \mathbb{C}^2$ . The momentum map for  $SU(2)$  acting on  $\mathbb{C}^2$  defined by

$$\langle J(\mathbf{a}), \xi \rangle_{su(2)^* \times su(2)} = \langle\langle \mathbf{a}, i\xi \mathbf{a} \rangle\rangle_{\mathbb{C}^2} \quad (11.3.4)$$

is equivariant. That is,

$$J(U\mathbf{a}) = \text{Ad}_{U^{-1}}^* J(\mathbf{a}).$$

**Proof.** Substitute  $\text{Ad}_{U^{-1}} \xi$  into the momentum map definition,

$$\begin{aligned} \langle \text{Ad}_{U^{-1}}^* J(\mathbf{a}), \xi \rangle_{su(2)^* \times su(2)} &= \langle J(\mathbf{a}), \text{Ad}_{U^{-1}} \xi \rangle_{su(2)^* \times su(2)} \\ &= \langle\langle \mathbf{a}, U^\dagger i\xi U \mathbf{a} \rangle\rangle_{\mathbb{C}^2} \\ &= \langle\langle (U\mathbf{a}), i\xi (U\mathbf{a}) \rangle\rangle_{\mathbb{C}^2} \\ &= \langle J(U\mathbf{a}), \xi \rangle_{su(2)^* \times su(2)}. \end{aligned}$$

Therefore,  $J(U\mathbf{a}) = \text{Ad}_{U^{-1}}^* J(\mathbf{a})$ , as claimed.  $\blacksquare$

**Remark 11.3.1 (Poincaré sphere momentum map)** Looking at Equation (3.2.16) reveals that the momentum map  $\mathbb{C}^2 \mapsto su(2)^*$  for the action of  $SU(2)$  acting on  $\mathbb{C}^2$  in Equation (11.3.1) is a component of the map  $\mathbb{C}^2 \mapsto S^2$  to the Poincaré sphere, which defines the Hopf fibration  $S^3 \simeq S^2 \times S^1$ . To see this, one simply replaces  $\xi \in su(2)$



with the vector of Pauli matrices  $\sigma$  in Equation (3.2.18) to find

$$J(\mathbf{a}) = \frac{1}{2} \mathbf{n} \cdot \sigma \quad \text{in which} \quad \mathbf{n} = \text{tr } Q \sigma = a_k^* \sigma_{kl} a_l. \quad (11.3.5)$$

Thus, the quadratic  $S^1$ -invariant quantities comprising the components of the unit vector  $\mathbf{n} = (n_1, n_2, n_3)$  given by

$$n_1 + i n_2 = 2a_1 a_2^*, \quad n_3 = |a_1|^2 - |a_2|^2 \quad (11.3.6)$$

on the Poincaré sphere  $|\mathbf{n}|^2 = n_0^2$  in (3.2.20) are precisely the components of the momentum map in Equation (11.3.1).  $\square$

**Exercise.** Compute the Lie–Poisson brackets among the components of the unit vector  $\mathbf{n} = (n_1, n_2, n_3)$  by using their definitions in terms of  $\mathbf{a} \in \mathbb{C}^2$  and applying the canonical Poisson brackets  $\{a_k^*, a_l\} = 2i\delta_{kl}$ .  $\star$

**Answer.**

$$\{n_i, n_j\} = 4\epsilon_{ijk} n_k.$$



**Remark 11.3.2 (Cotangent-lift momentum maps)** The formula determining the momentum map for the cotangent lift action of a Lie group  $G$  on a smooth manifold  $Q$  may be expressed in terms of the pairings

$$\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R} \quad \text{and} \quad \langle \langle \cdot, \cdot \rangle \rangle : T^*M \times TM \mapsto \mathbb{R},$$

as

$$\langle J(p, q), \xi \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = \langle q \diamond p, \xi \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = \langle \langle p, \mathcal{L}_\xi q \rangle \rangle_{T^*M \times TM}, \quad (11.3.7)$$

where  $(q, p) \in T_q^*M$  and  $\mathcal{L}_\xi q \in T_qM$  is the infinitesimal generator of the action of the Lie algebra element  $\xi$  on the coordinate  $q$ .  $\square$

**Proposition 11.3.2 (Equivariance of cotangent lifts)** *Cotangent-lift momentum maps (11.3.7) are equivariant. That is,*

$$J(g \cdot p, g \cdot q) = \text{Ad}_{g^{-1}}^* J(p, q), \quad (11.3.8)$$

where  $(g \cdot p, g \cdot q)$  denotes the cotangent lift to  $T^*M$  of the action of  $G$  on manifold  $M$ .

**Proof.** The proof follows from Remark 4.2.2, that  $\text{Ad}_{g^{-1}}^*$  is a representation of the coAdjoint action  $\Phi_g^*$  of the group  $G$  on its dual Lie algebra  $\mathfrak{g}^*$ . This means that  $\text{Ad}_{g^{-1}}^*(q \diamond p) = (g \cdot q \diamond g \cdot p)$ , and we have

$$\begin{aligned} \left\langle \text{Ad}_{g^{-1}}^* J(p, q), \xi \right\rangle_{\mathfrak{g}^* \times \mathfrak{g}} &= \left\langle \text{Ad}_{g^{-1}}^*(q \diamond p), \xi \right\rangle_{\mathfrak{g}^* \times \mathfrak{g}} \\ &= \left\langle g \cdot q \diamond g \cdot p, \xi \right\rangle_{\mathfrak{g}^* \times \mathfrak{g}} \\ &= \left\langle J(g \cdot p, g \cdot q), \xi \right\rangle_{\mathfrak{g}^* \times \mathfrak{g}}. \end{aligned} \quad (11.3.9)$$

Thus, Equation (11.3.8) holds and cotangent-lift momentum maps are equivariant. ■

### **Importance of equivariance**

Equivariance of a momentum map is important, because Poisson brackets among the components of an equivariant momentum map close among themselves and satisfy the Jacobi identity. That is, the following theorem holds.

**Theorem 11.3.1** *Equivariant momentum maps are Poisson.*

**Proof.** As we know, a momentum map  $J : P \rightarrow \mathfrak{g}^*$  is equivariant, if

$$J \circ \Phi_{g(t)} = \text{Ad}_{g(t)^{-1}}^* \circ J,$$

for any curve  $g(t) \in G$ . As discussed earlier, the time derivative of the equivariance relation leads to the infinitesimal equivariance relation,

$$\{\langle J, \xi \rangle, \langle J, \eta \rangle\} = \langle J, [\xi, \eta] \rangle, \quad (11.3.10)$$

where  $\xi, \eta \in \mathfrak{g}$  and  $\{\cdot, \cdot\}$  denotes the Poisson bracket on the manifold  $P$ . This in turn implies that the momentum map preserves Poisson brackets in the sense that

$$\{F_1 \circ J, F_2 \circ J\} = \{F_1, F_2\}_{LP} \circ J, \quad (11.3.11)$$

for all  $F_1, F_2 \in \mathcal{F}(\mathfrak{g}^*)$ , where  $\{F_1, F_2\}_{LP}$  denotes the Lie–Poisson bracket for the appropriate left or right action of  $\mathfrak{g}$  on  $P$ . That is, equivariance implies infinitesimal equivariance, which is sufficient for the momentum map to be Poisson. ■

**Exercise. (Compute  $N$ -dimensional momentum maps)**

Define appropriate pairings and determine the momentum maps explicitly for the following actions:

- (i)  $\mathcal{L}_\xi q = \xi \times q$  for  $\mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$ .
- (ii)  $\mathcal{L}_\xi q = \text{ad}_\xi q$  for adjoint action  $\text{ad} : \mathfrak{g} \times \mathfrak{g} \mapsto \mathfrak{g}$  in a Lie algebra  $\mathfrak{g}$ .
- (iii)  $AqA^{-1}$  for  $A \in GL(3, \mathbb{R})$  acting on  $q \in GL(3, \mathbb{R})$  by matrix conjugation.
- (iv)  $Aq$  for left action of  $A \in SO(3)$  on  $q \in SO(3)$ .
- (v)  $AqA^T$  for  $A \in GL(3, \mathbb{R})$  acting on  $q \in \text{Sym}(3)$ , that is  $q = q^T$ .
- (vi) Adjoint action of the Lie algebra of the semidirect-product group  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$  on itself. See Section 6.5 for notation and coadjoint actions. ★

## 232 11 : MOMENTUM MAPS

**Answer.**

- (i) For the pairing by scalar product of vectors, one writes

$$\left\langle\left\langle p, \mathcal{L}_\xi q \right\rangle\right\rangle_{T^*M \times TM} = p \cdot \xi \times q = q \times p \cdot \xi,$$

so that the momentum map for the spatial rotation (2.1.13) is

$$J = q \times p. \quad (11.3.12)$$

- (ii) Similarly, for the pairing  $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$ ,

$$\langle p, \text{ad}_\xi q \rangle = -\langle \text{ad}_q^* p, \xi \rangle \Rightarrow J = -\text{ad}_q^* p.$$

- (iii) Compute the ad action for  $GL(3, R)$  conjugation as

$$T_e(AqA^{-1}) = \xi q - q\xi = [\xi, q],$$

for  $\xi = A'(0) \in \mathfrak{gl}(3, R)$  acting on  $q \in GL(3, R)$  by matrix Lie bracket  $[\cdot, \cdot]$ . For the matrix pairing

$$\langle A, B \rangle = \text{tr}(A^T B),$$

one finds the momentum map,

$$\text{tr}(p^T [\xi, q]) = \text{tr}\left((pq^T - q^T p)^T \xi\right) \Rightarrow J = pq^T - q^T p.$$

- (iv) Compute  $T_e(Aq) = \xi q$  for  $\xi = A'(0) \in \mathfrak{so}(3)$  acting on  $q \in SO(3)$  by left matrix multiplication. For the matrix pairing  $\langle A, B \rangle = \text{trace}(A^T B)$ , one finds the following expression for the momentum map,

$$\text{trace}(p^T \xi q) = \text{trace}((pq^T)^T \xi) \Rightarrow J = \frac{1}{2}(pq^T - qp^T),$$

upon using antisymmetry of the matrix  $\xi \in \mathfrak{so}(3)$ .

(v) Compute

$$T_e(AqA^T) = \xi q + q\xi^T$$

for  $\xi = A'(0) \in gl(3, R)$  acting on  $q \in Sym(3)$ . For the matrix pairing

$$\begin{aligned} \langle A, B \rangle &= \text{tr}(A^T B), \text{ one finds} \\ \text{tr}(p^T(\xi q + q\xi^T)) &= \text{tr}(q(p^T + p)\xi) \\ &= \text{tr}(2qp)^T \xi \Rightarrow J = 2qp \end{aligned}$$

upon using symmetry of the matrix  $\xi q + q\xi^T$  to choose  $p = p^T$ . (The momentum canonical to the symmetric matrix  $q = q^T$  should be symmetric, in order to have the correct number of components.)

(vi) For the pairing  $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$ ,

$$\langle p, \text{ad}_\xi q \rangle = -\langle \text{ad}_q^* p, \xi \rangle \Rightarrow J = -\text{ad}_q^* p.$$

From Equation (6.5.22), this is

$$J = -\text{ad}_{(A,h)}^* \begin{pmatrix} D \\ k \end{pmatrix} = - \begin{pmatrix} [D, A] - hk + \frac{1}{2}(\mathbf{h} \cdot \mathbf{k})\mathbb{1} \\ kA \end{pmatrix}$$

for  $(A, h) \in \mathfrak{g}$  and  $(D, k) \in \mathfrak{g}^*$ . ▲

**Exercise. (Unitary transformations of Hermitian matrices)**

Consider the manifold  $\mathcal{Q}$  of  $n \times n$  Hermitian matrices, so that  $Q^\dagger = Q$  for  $Q \in \mathcal{Q}$ . The Poisson (symplectic) manifold is  $T^*\mathcal{Q}$ , whose elements are pairs  $(Q, P)$  of Hermitian matrices. The corresponding Poisson bracket is

$$\{F, H\} = \text{tr} \left( \frac{\partial F}{\partial Q} \frac{\partial H}{\partial P} - \frac{\partial H}{\partial Q} \frac{\partial F}{\partial P} \right).$$

Let  $G$  be the group  $U(n)$  of  $n \times n$  unitary matrices. The group  $G$  acts on  $T^*\mathcal{Q}$  through

$$(Q, P) \mapsto (UQU^\dagger, UPU^\dagger), \quad UU^\dagger = Id.$$

## 234 11 : MOMENTUM MAPS

- (i) What is the linearisation of this group action?
- (ii) What is its momentum map?
- (iii) Is this momentum map equivariant?
- (iv) Is this momentum map conserved by the Hamiltonian  $H = \frac{1}{2}\text{tr } P^2$ ? Prove it.
- (v) What changes occur in the solution for *orthogonal* transformations of the manifold of  $n \times n$  *symmetric* matrices, instead? ★

**Answer. (Unitary transformations of Hermitian matrices)**

- (i) The linearisation of this group action with  $U = \exp(t\xi)$ , with skew-Hermitian  $\xi^\dagger = -\xi$ , yields the vector field with  $(Q, P)$  components,

$$X_\xi = \left( [\xi, Q], [\xi, P] \right).$$

- (ii) This is the Hamiltonian vector field for

$$H_\xi = \text{tr}([Q, P]\xi),$$

thus yielding the momentum map  $J(Q, P) = [Q, P]$ .

That is, the momentum map for the lifted action of the unitary transformations on the phase space  $(Q, P) \in T^*Q$  of the Hermitian matrices is the matrix commutator  $[Q, P]$ .

This is entirely natural from the viewpoint of quantum mechanics, in which the commutator  $[Q, P]$  is responsible for the *uncertainty principle*.

- (iii) Being defined by a cotangent lift, this momentum map is equivariant.

(iv) For  $H = \frac{1}{2}\text{tr } P^2$ ,

$$\{[Q, P], H\} = \text{tr} \left( \frac{\partial [Q, P]}{\partial Q} \frac{\partial H}{\partial P} \right) = \text{tr} (P^2 - P^2) = 0,$$

so the momentum map  $J(Q, P) = [Q, P]$  is conserved by this Hamiltonian.

Alternatively, one may simply observe that the map

$$(Q, P) \mapsto (UQU^\dagger, UPU^\dagger), \quad UU^\dagger = Id,$$

preserves  $\text{tr}(P^2)$ , since it takes

$$\text{tr}(P^2) \mapsto \text{tr}(UPU^\dagger UPU^\dagger) = \text{tr}(P^2).$$

(v) The computation for orthogonal transformations of the manifold of  $n \times n$  symmetric matrices is entirely analogous, except for minor changes in interpretation. ▲

**Example 11.3.3 (The 1:1 resonance [Ku1978])** *This example extends the  $\mathbb{C}^2 \mapsto su(2)^*$  momentum map (11.3.1) to  $\mathbb{C}^2 \mapsto u(2)^*$  and thereby completes the relation of the momentum map to the Poincaré sphere and Hopf fibration.*

*A unitary  $2 \times 2$  matrix  $U(s)$  acts on a complex two-vector  $\mathbf{a} \in \mathbb{C}^2$  by matrix multiplication as*

$$\mathbf{a}(s) = U(s)\mathbf{a}(0) = \exp(is\xi)\mathbf{a}(0),$$

*in which  $i\xi = U'U^{-1}|_{s=0}$  is a  $2 \times 2$  skew-Hermitian matrix. Therefore, the infinitesimal generator  $\xi(\mathbf{a}) \in \mathbb{C}^2$  may be expressed as a linear transformation,*

$$\xi(\mathbf{a}) = \frac{d}{ds} [\exp(is\xi)\mathbf{a}] \Big|_{s=0} = i\xi\mathbf{a},$$

*in which the matrix  $\xi^\dagger = \xi$  is Hermitian.*

**Definition 11.3.1 (Momentum map  $J : \mathbb{C}^2 \mapsto u(2)^*$ )** *The momentum map  $J(\mathbf{a}) : \mathbb{C}^2 \mapsto u(2)^*$  for the matrix action of  $U(2)$  on  $\mathbb{C}^2$  is defined by*

$$\begin{aligned} J^\xi(\mathbf{a}) &:= \langle J(\mathbf{a}), \boldsymbol{\xi} \rangle_{u(2)} = \frac{i}{2} \langle\langle \mathbf{a}, \xi(\mathbf{a}) \rangle\rangle_{\mathbb{C}^2} \\ &= \frac{1}{2} \omega(\mathbf{a}, \xi(\mathbf{a})) \quad \text{with} \quad \xi(\mathbf{a}) = i\xi\mathbf{a}, \end{aligned} \quad (11.3.13)$$

and  $\xi^\dagger = \xi$ . The  $\mathbb{C}^2$  pairing  $\langle\langle \cdot, \cdot \rangle\rangle_{\mathbb{C}^2}$  in this map is the Hermitian pairing, which for skew-Hermitian  $\xi(\mathbf{a})^\dagger = -\xi(\mathbf{a})$  is also the canonical symplectic form,  $\omega(\mathbf{a}, \mathbf{b}) = \text{Im}(\mathbf{a}^* \cdot \mathbf{b})$  on  $\mathbb{C}^2$ , as discussed in [MaRa1994]. Thus,

$$\begin{aligned} 2J^\xi(\mathbf{a}) &:= \omega(\mathbf{a}, \xi(\mathbf{a})) = \omega(\mathbf{a}, i\xi\mathbf{a}) \\ &= \text{Im}(a_k^* (i\xi)_{kl} a_l) \\ &= a_k^* \xi_{kl} a_l \\ &= \text{tr}((\mathbf{a} \otimes \mathbf{a}^*) \xi) \\ &= \text{tr}(J^\dagger(\mathbf{a}^*, \mathbf{a}) \xi). \end{aligned} \quad (11.3.14)$$

Consequently, the momentum map  $J : \mathbb{C}^2 \mapsto u(2)^*$  is given by the Hermitian expression

$$J(\mathbf{a}) = \frac{1}{2} \mathbf{a} \otimes \mathbf{a}^*. \quad (11.3.15)$$

This conclusion may be checked by computing the differential of the Hamiltonian  $dJ^\xi(\mathbf{a})$  for the momentum map, which should be canonically related to its Hamiltonian vector field  $X_{J^\xi(\mathbf{a})} = \{\cdot, J^\xi(\mathbf{a})\}$ . As the infinitesimal generator  $\xi(\mathbf{a}) = i\xi\mathbf{a}$  is linear, we have

$$\begin{aligned} dJ^\xi(\mathbf{a}) &= d\langle J(\mathbf{a}), \boldsymbol{\xi} \rangle_{u(2)} = \frac{i}{2} \langle\langle \mathbf{a}, \xi(d\mathbf{a}) \rangle\rangle_{\mathbb{C}^2} + \frac{i}{2} \langle\langle d\mathbf{a}, \xi(\mathbf{a}) \rangle\rangle_{\mathbb{C}^2} \\ &= \mathfrak{S} \langle\langle \xi(\mathbf{a}), d\mathbf{a} \rangle\rangle_{\mathbb{C}^2} = \omega(\xi(\mathbf{a}), \cdot) = X_{J^\xi(\mathbf{a})} \lrcorner \omega, \end{aligned}$$

which is the desired canonical relation.



### 11.3.1 The Poincaré sphere $S^2 \in S^3$

We expand the Hermitian matrix  $J = \frac{1}{2}\mathbf{a} \otimes \mathbf{a}^*$  in (11.3.15) in a basis of four  $2 \times 2$  unit Hermitian matrices  $(\sigma_0, \boldsymbol{\sigma})$ , with  $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$  given by

$$\begin{aligned}\sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \sigma_1 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \sigma_2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, & \sigma_3 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.\end{aligned}\quad (11.3.16)$$

The result is the decomposition

$$J = \frac{1}{4}(R\sigma_0 + \mathbf{Y} \cdot \boldsymbol{\sigma}). \quad (11.3.17)$$

Here we denote  $R := \text{tr}(J\sigma_0) = |a_1|^2 + |a_2|^2$  and

$$\mathbf{Y} = \text{tr}(J\boldsymbol{\sigma}) = a_k^* \boldsymbol{\sigma}_{kl} a_l, \quad (11.3.18)$$

with vector notation  $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ . In components, one finds

$$J = \frac{1}{2} \begin{bmatrix} a_1^* a_1 & a_1^* a_2 \\ a_2^* a_1 & a_2^* a_2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} R + Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & R - Y_3 \end{bmatrix}, \quad (11.3.19)$$

with trace  $\text{tr} J = R$ . Thus, the decomposition (11.3.17) splits the momentum map into its trace part  $R \in \mathbb{R}$  and its traceless part  $\mathbf{Y} \in \mathbb{R}^3$ , given by

$$\mathbf{Y} = J - \frac{1}{2}(\text{tr} J) \text{Id} \in \mathfrak{su}(2)^* \cong \mathbb{R}^3. \quad (11.3.20)$$

This formula recovers the  $SU(2)$  momentum map in Equation (11.3.5) found earlier for the Poincaré sphere.

**Definition 11.3.2 (Poincaré sphere)** *The coefficients  $R \in \mathbb{R}$  and  $\mathbf{Y} \in \mathbb{R}^3$  in the expansion of the matrix  $J$  in (11.3.17) comprise the four real quadratic quantities,*

$$\begin{aligned}R &= \frac{1}{2}(|a_1|^2 + |a_2|^2), \\ Y_3 &= \frac{1}{2}(|a_1|^2 - |a_2|^2) \quad \text{and} \\ Y_1 - iY_2 &= a_1^* a_2.\end{aligned}\quad (11.3.21)$$

## 238 11 : MOMENTUM MAPS

These quantities are all invariant under the action  $\mathbf{a} \rightarrow e^{i\phi}\mathbf{a}$  of  $\phi \in S^1$  on  $\mathbf{a} \in \mathbb{C}^2$ . The  $S^1$ -invariant coefficients in the expansion of the momentum map  $J = \mathbf{a} \otimes \mathbf{a}^*$  (11.3.15) in the basis of sigma matrices (11.3.16) satisfy the relation

$$4 \det J = R^2 - |\mathbf{Y}|^2 = 0, \quad \text{with} \quad |\mathbf{Y}|^2 \equiv Y_1^2 + Y_2^2 + Y_3^2. \quad (11.3.22)$$

This relation defines the **Poincaré sphere**  $S^2 \in S^3$  of radius  $R$  which, in turn, is related to the Hopf fibration  $\mathbb{C}^2/S^1 \simeq S^3$ . For more information about the Poincaré sphere and the Hopf fibration, consult, e.g., [Ho2008] and references therein.

**The  $U(2)$  Lie group structure**

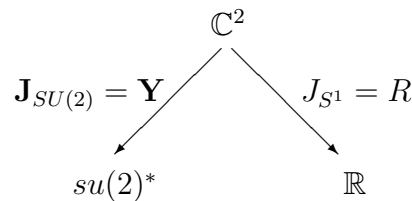
The Lie group  $U(2) = S^1 \times SU(2)$  is the direct product of its centre,

$$Z(U(2)) = \{zI \text{ with } |z| = 1\} \equiv S^1,$$

and the special unitary group in two dimensions,

$$SU(2) = \left\{ \begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix} \text{ with } |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

As a consequence, the momentum map  $J(\mathbf{a}) = \frac{1}{2}\mathbf{a} \otimes \mathbf{a}^*$  in (11.3.15) for the action  $U(2) \times \mathbb{C}^2 \rightarrow \mathbb{C}^2$  decomposes into two momentum maps obtained by separating  $J$  into its trace part  $J_{S^1} = R \in \mathbb{R}$  and its traceless part  $\mathbf{J}_{SU(2)} = \mathbf{Y} \in \mathbb{R}^3$ . This decomposition may be sketched, as follows.



The target spaces  $su(2)^*$  and  $\mathbb{R}$  of the left and right legs of this pair of momentum maps are each Poisson manifolds, with coordi-

nates  $\mathbf{Y} \in su(2)^*$  and  $R \in \mathbb{R}$ , respectively. The corresponding Poisson brackets are given in tabular form as

$\{\cdot, \cdot\}$	$Y_1$	$Y_2$	$Y_3$	$R$
$Y_1$	0	$Y_3$	$-Y_2$	0
$Y_2$	$-Y_3$	0	$Y_1$	0
$Y_3$	$Y_2$	$-Y_1$	0	0
$R$	0	0	0	0

(11.3.23)

In index notation, these Poisson brackets are given as

$$\{Y_k, Y_l\} = \epsilon_{klm} Y_m \quad \text{and} \quad \{Y_k, R\} = 0. \quad (11.3.24)$$

The last Poisson bracket relation means that the spaces with coordinates  $\mathbf{Y} \in su(2)^*$  and  $R \in \mathbb{R}$  are *symplectically orthogonal* in  $u(2)^* = su(2)^* \times \mathbb{R}$ .

Equations (11.3.24) prove the following.

**Theorem 11.3.2 (Momentum map (11.3.15) is Poisson)** *The direct-product structure of  $U(2) = S^1 \times SU(2)$  decomposes the momentum map  $J$  in Equation (11.3.15) into two other momentum maps,  $J_{SU(2)} : \mathbb{C}^2 \mapsto su(2)^*$  and  $J_{S^1} : \mathbb{C}^2 \mapsto \mathbb{R}$ . These other momentum maps are also Poisson maps. That is, they each satisfy the Poisson property for smooth functions  $F$  and  $H$ ,*

$$\{F \circ J, H \circ J\} = \{F, H\} \circ J. \quad (11.3.25)$$

*This relation defines a Lie–Poisson bracket on  $su(2)^*$  that inherits the defining properties of a Poisson bracket from the canonical relations*

$$\{a_k, a_l^*\} = -2i\delta_{kl},$$

*for the canonical symplectic form,  $\omega = \Im(da_j \wedge da_j^*)$ .*

**Remark 11.3.3** The Poisson bracket table in (11.3.23) is the  $so(3)^*$  Lie–Poisson bracket table for angular momentum in the *spatial* frame. It differs by an overall sign from the  $so(3)^*$  Lie–Poisson bracket table for angular momentum in the *body* frame, see (2.5.13). □

## 240 11 : MOMENTUM MAPS

**Definition 11.3.3 (Dual pairs)** Let  $(M, \omega)$  be a symplectic manifold and let  $P_1, P_2$  be two Poisson manifolds. A pair of Poisson mappings

$$P_1 \xleftarrow{J_1} (M, \omega) \xrightarrow{J_2} P_2$$

is called a **dual pair** [We1983b] if  $\ker TJ_1$  and  $\ker TJ_2$  are symplectic orthogonal complements of one another, that is,

$$(\ker TJ_1)^\omega = \ker TJ_2. \quad (11.3.26)$$

A systematic treatment of dual pairs can be found in Chapter 11 of [OrRa2004]. The infinite-dimensional case is treated in [GaVi2010].

**Remark 11.3.4 (Summary)** In the pair of momentum maps

$$su(2)^* \equiv \mathbb{R}^3 \xleftarrow{\mathbf{Y}} (\mathbb{C}^2, \omega) \xrightarrow{R} \mathbb{R}, \quad (11.3.27)$$

$\mathbf{Y}$  maps the fibres of  $R$ , which are three-spheres, into two-spheres, that are coadjoint orbits of  $SU(2)$ . The restriction of  $\mathbf{Y}$  to these three-spheres is the Hopf fibration. Further pursuit of the theory of dual pairs is beyond our present scope. See [HoVi2010] for a recent discussion of dual pairs for resonant oscillators from the present viewpoint.  $\square$

**Example 11.3.4 (An infinite-dimensional momentum map)** Let  $\mathcal{F} : T^*\mathcal{M} \mapsto \mathbb{R}$  be the space of real-valued functions on phase space defined by the cotangent bundle  $T^*\mathcal{M}$  with coordinates  $(q, p)$  of a manifold  $\mathcal{M}$  with coordinates  $q$ . The dual of the phase-space functions comprises the phase-space densities, denoted  $\mathcal{F}^*$ . The pairing  $\langle \cdot, \cdot \rangle : \mathcal{F}^* \times \mathcal{F} \mapsto \mathbb{R}$  between a phase-space function  $g \in \mathcal{F}$  and a phase-space density  $f dp \wedge dq \in \mathcal{F}^*$  is defined by the integral over phase space,

$$\langle f, g \rangle := \int_{T^*\mathcal{M}} f g dq \wedge dp.$$

We identify the space  $\mathcal{F}^* \times \mathcal{F} \simeq T^*\mathcal{F}$  as the cotangent bundle of the space of phase-space functions  $\mathcal{F}$ .

Let  $X_h = \{ \cdot, h \}$  be the Hamiltonian vector field defined by applying the canonical Poisson bracket on  $T^*\mathcal{M}$  using the phase-space function

$h \in \mathcal{F}$ . The momentum map for the action of canonical (symplectic) transformations on phase-space functions is given by

$$J_h(f, g) = \langle f, X_h g \rangle = \langle f, \{g, h\} \rangle = \langle h, \{f, g\} \rangle,$$

where we have integrated by parts and invoked homogeneous boundary conditions in the third step. This discussion reveals the following.

**Lemma 11.3.1** *The Poisson bracket  $\{f, g\}(q, p)$  is the momentum map for the action of canonical transformations on phase-space functions.*

**Remark 11.3.5** The preservation of this momentum map (i.e., the Poisson bracket) under canonical transformations on phase-space functions is no surprise; this preservation is part of the definition of a canonical transformation.  $\square$

Equivalently, one may write this momentum map as a pairing between the Lie algebra  $\mathfrak{g}$  of Hamiltonian vector fields and its dual,  $\mathfrak{g}^*$ , the one-form densities on phase space. The following lemma provides this representation.

**Lemma 11.3.2** *The Poisson one-form density  $f dg \otimes dq \wedge dp$  is the momentum map dual to the action of Hamiltonian vector fields on phase-space functions  $f$  and  $g$ .*

**Proof.** The momentum map for this action is defined in terms of the pairing,

$$\begin{aligned} J_{X_h}(f, g) &= \langle J(f, g), X_h \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = - \langle f \diamond g, X_h \rangle_{\mathfrak{g}^* \times \mathfrak{g}} \\ &= \langle f, X_h g \rangle \\ &= \int_{T^* \mathcal{M}} f(g_q h_p - h_q g_p) dq \wedge dp \\ &= \int_{T^* \mathcal{M}} (X_h \lrcorner f dg) dq \wedge dp \\ &= \langle f dg, X_h \rangle_{\mathfrak{g}^* \times \mathfrak{g}}, \end{aligned} \tag{11.3.28}$$

where  $X_h \lrcorner f dg$  denotes substitution of the Hamiltonian vector field  $X_h = \{\cdot, h\}$  into the differential one-form on phase space

$$f dg(q, p) = f(g_q dq + g_p dp). \quad \blacksquare$$

**Remark 11.3.6** The computation in (11.3.28) shows that the exterior differential ( $d$ ) on phase space is minus the diamond operation for the action of Hamiltonian vector fields on phase-space functions.  $\square$

### 11.3.2 Overview

Looking back on Section 2.5.5 which discussed the Clebsch variational principle for the rigid body, we notice that the first of Equations (2.5.22) is exactly the momentum map for body angular momentum. Likewise, the momentum map ( $q \diamond p$ ) emerged in Equation (9.3.8) in the Clebsch variational approach to the Euler–Poincaré equation in Chapter 9.

This observation affirms the main message of the book: Lie symmetry reduction on the Lagrangian side produces the Euler–Poincaré equation, as discussed in Chapter 9. Its formulation on the Hamiltonian side as a Lie–Poisson equation in Section 9.4 governs the dynamics of the momentum map defined in Equation (11.2.1), which derives from the cotangent lift of the Lie algebra action of the original Lie symmetry on the configuration manifold defined in (11.1.2). This is the relation between the results of reduction by Lie symmetry on the Lagrangian and Hamiltonian sides.

The primary purpose of this book has been to explain that statement, so that it is understood by undergraduate students in mathematics, physics and engineering.

**Remark 11.3.7** Looking forward, the reader may wish to pursue these ideas further, and go beyond the cotangent-lift momentum maps discussed here. At that point, the reader will need to consult the books [OrRa2004, MaMiOrPeRa2007] for discussions of momentum maps obtained by singular reduction techniques and reduction by stages. For discussions of the use of momentum maps derived from Hamilton–Pontryagin and Clebsch variational principles, respectively, in designing geometric integrators for discrete dynamical systems, see [BoMa2009] and [CoHo2007].  $\square$

**Exercise.** The canonical Poisson bracket  $\{g, h\}$  between two phase-space functions  $g$  and  $h$  induces a Lie–Poisson bracket  $\{\cdot, \cdot\}_{LP}$  between linear functionals of phase-space functions defined by

$$J_{X_h}(f, g) = \langle f, \{g, h\} \rangle = \{ \langle f, g \rangle, \langle f, h \rangle \}_{LP}.$$

Show that this Lie–Poisson bracket  $\{\cdot, \cdot\}_{LP}$  satisfies the properties that define a Poisson bracket, including the Jacobi identity. ★

**Exercise.** Let the components of the angular momentum vector  $\mathbf{L} \in \mathbb{R}^3$  be defined by

$$L_i := \int (p \times q)_i f(q, p) dq \wedge dp = \langle f(p \times q)_i \rangle,$$

where  $i = 1, 2, 3$  and  $(\times)$  denotes vector product in  $\mathbb{R}^3$ . Compute the Lie–Poisson bracket  $\{L_i, L_j\}_{LP}$ . ★

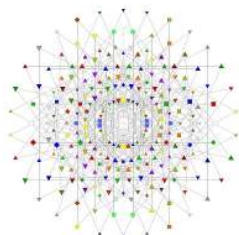
**Exercise.** Compute the Lie–Poisson bracket relation

$$\{J_{X_h}(f, g), J_{X_H}(f, G)\}_{LP} = J_{X_{\{G, H\}}}(f, \{g, h\})$$

for phase-space functions  $g, h, G, H$ . ★







# 12

## ROUND, ROLLING RIGID BODIES

---

### Contents

12.1	Introduction	246
12.1.1	Holonomic versus nonholonomic	246
12.1.2	The Chaplygin ball	248
12.2	Nonholonomic Hamilton–Pontryagin variational principle	252
12.2.1	HP principle for the Chaplygin ball	256
12.2.2	Circular disk rocking in a vertical plane	265
12.2.3	Euler’s rolling and spinning disk	268
12.3	Nonholonomic Euler–Poincaré reduction	275
12.3.1	Semidirect-product structure	276
12.3.2	Euler–Poincaré theorem	278
12.3.3	Constrained reduced Lagrangian	282

---

**246** 12 : ROUND, ROLLING RIGID BODIES

This chapter deals with round rigid bodies rolling on a flat surface, by applying the Hamilton–Pontryagin and Euler–Poincaré approaches to Lagrangian reduction for variational principles with symmetries. These approaches were introduced for the heavy top in Chapter 8. To the potential energy and rotational energy of the heavy top, one now adds translational energy for its centre of mass motion and imposes the nonholonomic constraint that the rigid body is rolling. We focus on the constrained rocking, rolling, but not sliding motion of a rigid body on a perfectly rough horizontal plane in two examples. These examples are the Chaplygin ball (a rolling spherical ball with an off-centre mass distribution, as sketched in Figure 12.1) and the Euler disk (modelling a gyrating coin, as sketched in Figure 12.2). The chapter closes by showing that the Hamilton–Pontryagin and Euler–Poincaré approaches to Lagrangian reduction yield the same equations as those obtained from the standard Lagrange–d’Alembert variational approach [Bl2004].

We hope the examples of geometric mechanics problems treated in this chapter will inspire confidence in interested readers. Having mastered the material so far, they should find themselves ready and completely equipped with the tools needed to model nonholonomic dynamics of rolling without sliding.

---

## 12.1 Introduction

### 12.1.1 Holonomic versus nonholonomic

Constraints which restrict the possible configurations for a mechanical system are called *holonomic*. Examples are restriction of a pendulum to have constant length, or restriction of a particle in three dimensions to move on the surface of a sphere. Constraints on the velocities which cannot be reduced to holonomic constraints are termed *nonholonomic*.

Nonholonomic systems typically arise when constraints on velocity are imposed, such as the constraint that the bodies roll without slipping on a surface. Cars, bicycles, unicycles – anything with rolling wheels – are all examples of nonholonomic systems.

Nonholonomic mechanics has a rich history that dates back to the time of Euler, Lagrange and d'Alembert. The classical work in this subject is summarised in Routh [Ro1860], Jellett [Je1872] and has been discussed more recently in the context of control theory in Neimark and Fufaev [NeFu1972]. The geometry of nonholonomic systems shares its mathematical foundations with geometric control theory. An introduction to nonholonomic constraints in geometric control applications (such as feedback laws that stabilise or generate locomotion) is given in [Bl2004]. A branch of mathematics known as sub-Riemannian geometry has developed for dealing with non-holonomic geometric control systems [Mo2002]. Control problems and sub-Riemannian geometry are beyond the scope of the present text.

**Definition 12.1.1** *A constrained dynamical system consists of*

- *a smooth manifold  $Q$ , which is the configuration space;*
- *a smooth function  $L : TQ \rightarrow \mathbb{R}$ , which is the Lagrangian (typically taken to be the difference between the kinetic energy and potential energy); and*
- *a smooth distribution  $\mathcal{D} \subset TQ$ , which determines the constraints.*

**Definition 12.1.2** *A distribution is a collection of linear subspaces of the tangent spaces of  $Q$  such that  $\mathcal{D}_q \subset T_qQ$  for  $q \in Q$ .*

**Remark 12.1.1** The  $\mathcal{D}_q$  tangent spaces of  $Q$  parameterise the allowable directions for the system at a given point  $q$  of the configuration space. A curve  $q(t) \in Q$  satisfies the constraints, provided  $\dot{q}(t) \in \mathcal{D}_{q(t)}$  for all  $t$ . (This implies that the constraints are linear in the velocities.) The distribution  $\mathcal{D} \subset TQ$  will in general not be the differential of a function on  $Q$ . That is, in general, the constraints will be *nonholonomic*.  $\square$

This chapter applies the same ideas of symmetry reduction underlying the Hamilton–Pontryagin and Euler–Poincaré approaches that were explained earlier in studying Hamilton's principle for the

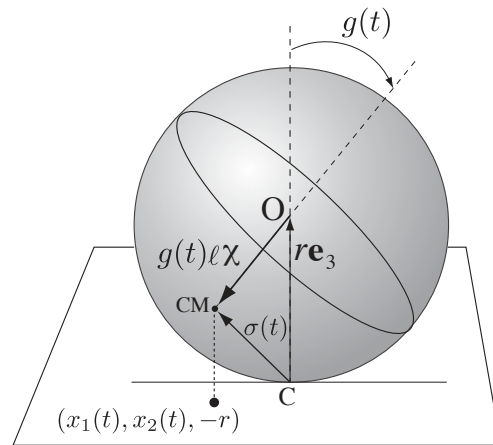
## 248 12 : ROUND, ROLLING RIGID BODIES

heavy top in Chapter 8. In addition, it introduces nonholonomic constraints into the variational principle by using *Lagrange multipliers*. The approach is illustrated by deriving the dynamical equations for the rocking, rolling, but not sliding motion of two classical nonholonomic problems. The first problem involves a spherical ball whose mass distribution is off-centre, and is rolling without slipping on a horizontal plane in the presence of gravity. This problem is called the *Chaplygin ball*. The second problem treats a disk that is simultaneously rolling, spinning and falling. This is the problem of the *Euler disk*. Both of these problems have stimulated many interesting investigations, and they are likely to keep doing so.

### 12.1.2 The Chaplygin ball

Consider a spherical ball of mass  $m$ , radius  $r$  and moment of inertia  $I$ , whose mass distribution is inhomogeneous, so that its centre of mass lies anywhere in the ball as it rolls without slipping on a horizontal plane in the presence of gravity. This problem was first solved by Chaplygin [Ch1903]. Extensive references for the history of this problem and the associated problem of the tippe top involving both rolling and sliding are given in [GrNi2000]. Modern geometric perspectives of this problem appear in [Cu1998, He1995, Ze1995, Sc2002, BoMaRo2004, GIPeWo2007, CiLa2007, Ki2011].

**Definition 12.1.3 (Geometry for the rolling ball)** *Let  $(E_1, E_2, E_3)$  denote the reference system of body coordinates, chosen to coincide with the ball's principal axes, in which the inertia tensor is diagonal,  $I = \text{diag}(I_1, I_2, I_3)$ . The coordinate directions in space are denoted  $(e_1, e_2, e_3)$ . These spatial unit vectors are chosen so that  $(e_1, e_2)$  are horizontal and  $e_3$  is vertical. The origin of the body coordinate system coincides with the centre of mass of the ball, which has coordinates  $x(t) = (x_1(t), x_2(t), x_3(t))$  in space. In the spatial coordinate system, the body frame is the moving frame  $(g(t)E_1, g(t)E_2, g(t)E_3)$ , where  $g(t) \in SO(3)$  defines the attitude of the body relative to its reference configuration.*



**Figure 12.1.** The Chaplygin ball. The position  $(x_1(t), x_2(t), x_3(t))$  is the position of the centre of mass, not the centre of the sphere. The spatial vector  $\sigma(t)$  points from the contact point  $\mathbf{C}$  to the centre of mass. The projection of the centre of mass location onto the plane is the point  $(x_1(t), x_2(t), -r)$ .

Thus, the motion of the ball is given by a curve in the group of special Euclidean transformations

$$(g(t), x(t)) \in SE(3) \simeq SO(3) \times \mathbb{R}^3,$$

where  $g(t) \in SO(3)$  and  $x(t) \in \mathbb{R}^3$ . An element of  $SE(3)$  maps a generic point  $P$  in the ball's reference coordinates to a point in space,

$$Q(t) = g(t)P + x(t),$$

at time  $t$ .

**Exercise.** Show that the left-invariant and right-invariant tangent vectors for this motion on  $SE(3)$  are, respectively,

$$(\Omega, Y) := (g^{-1}\dot{g}, g^{-1}\dot{x}),$$

and

## 250 12 : ROUND, ROLLING RIGID BODIES

$$(\omega, \zeta) := (\dot{g}g^{-1}, -\dot{g}g^{-1}x + \dot{x}).$$

See Equations (6.3.10) and (6.3.11).



**Definition 12.1.4 (Spatial moment vector)** Let  $\ell\chi$  denote the vector in the body frame that points from the geometric centre of the ball to its centre of mass, where

- $\ell$  is the fixed distance between these two points; and
- $\chi$  is a unit vector in the body, not necessarily aligned with a principal axis.

At any instant, the vector in space pointing from the contact point  $C$  to the centre of mass at  $x(t)$  is the **spatial moment vector**,

$$\sigma(t) = re_3 + \ell g(t)\chi.$$

**Definition 12.1.5 (Rolling constraint)** Physically, rolling of the body means that the velocity of its centre of mass  $\dot{x}(t)$  in the spatial frame at any instant  $t$  must be obtained by a rotation with angular frequency

$$\hat{\omega} = \dot{g}g^{-1} \in so(3),$$

around the fixed contact point. Consequently, the **rolling constraint** is imposed by requiring

$$\dot{x}(t) = \dot{g}g^{-1}\sigma(t) = \hat{\omega}\sigma(t) = \omega \times \sigma(t), \quad (12.1.1)$$

where we have used the hat map in the last step for the (right-invariant) spatial angular velocity. Equivalently, one may rewrite this rolling constraint as

$$\dot{x}(t) = \dot{g}(t)(rg^{-1}(t)e_3 + \ell\chi). \quad (12.1.2)$$

This formula defines the **spatial nonholonomic constraint distribution**  $\mathcal{D}$  for the Chaplygin ball.

**Exercise.** What is the velocity  $\dot{x}(t)$  in the body coordinate system? ★

### ***The Hamilton–Pontryagin approach for the Chaplygin ball***

The Hamilton–Pontryagin method proceeds after assembling the following ingredients:

- **Potential energy:** The potential energy due to gravity for the Chaplygin ball is  $m\gamma x_3(t)$ , where

$$x_3(t) = \langle e_3, g(t)\ell\chi \rangle = \langle g^{-1}(t)e_3, \ell\chi \rangle \quad (12.1.3)$$

is the vertical displacement of the centre of mass relative to the centre of the ball and  $\gamma$  is the constant acceleration of gravity.

- **Kinetic energy:** The kinetic energy of the Chaplygin ball is the sum of the energy of translation of its centre of mass and the energy of rotation about the centre of mass.
- **Lagrangian:** The Lagrangian is defined to be the difference between the total kinetic energy and the gravitational potential energy,

$$L = \frac{1}{2} \langle \Omega, I\Omega \rangle + \frac{m}{2} |\dot{x}|^2 - m\gamma \ell \langle g^{-1}e_3, \chi \rangle,$$

where  $\Omega = g^{-1}\dot{g}$  is the (left-invariant) body angular velocity.

- **Rolling constraint in the body:** The spatial distribution (12.1.2) for the rolling constraint transforms into body coordinates upon setting

$$Y = g^{-1}\dot{x}, \quad \Gamma = g^{-1}e_3, \quad s = g^{-1}\sigma = r\Gamma + \ell\chi, \quad (12.1.4)$$

with vectors  $Y, \Gamma, s$  each in  $\mathbb{R}^3$ . The rolling constraint in body coordinates then simplifies to

$$Y = \Omega s. \quad (12.1.5)$$

## 252 12 : ROUND, ROLLING RIGID BODIES

The hat map expresses this rolling constraint equivalently in vector form as  $Y = \Omega \times s$ .

- **Reduced Lagrangian in body coordinates:** The reduced Lagrangian is defined in body coordinates by using the definitions in (12.1.5), as

$$l(\Omega, Y, \Gamma) = \frac{1}{2} \langle \Omega, I\Omega \rangle + \frac{m}{2} |Y|^2 - m\gamma \ell \langle \Gamma, \chi \rangle. \quad (12.1.6)$$

These considerations have expressed a spatial Lagrangian  $L : TG \times TV \times V \rightarrow \mathbb{R}$  in its body form as  $l : \mathfrak{g} \times TV \times V \rightarrow \mathbb{R}$ , plus the various constraint relations.

## 12.2 Nonholonomic Hamilton–Pontryagin variational principle

We shall write Hamilton’s principle  $\delta S = 0$  for a class of nonholonomically constrained action principles that includes the Chaplygin ball, by simply using Lagrange multipliers to impose the constraints on the reduced Lagrangian, as in the Hamilton–Pontryagin principle for the matrix Euler equations in Section 2.4.1. In particular, we apply the Hamilton–Pontryagin approach to the following class of constrained action integrals,

$$\begin{aligned} S &= \int L(g, \dot{g}, \dot{x}, e_3) dt \\ &= \int \left\{ l(\Omega, Y, \Gamma) + \langle \Pi, g^{-1}\dot{g} - \Omega \rangle \right. \\ &\quad \left. + \langle \kappa, g^{-1}e_3 - \Gamma \rangle + \langle \lambda, g^{-1}\dot{x} - Y \rangle \right\} dt, \end{aligned} \quad (12.2.1)$$

where

$$l(\Omega, Y, \Gamma) = L(e, g^{-1}\dot{g}, g^{-1}\dot{x}, g^{-1}e_3). \quad (12.2.2)$$

This class of action integrals produces constrained equations of motion determined from the Hamilton–Pontryagin principle.



**Theorem 12.2.1 (Hamilton–Pontryagin principle)** *The stationarity condition for the nonholonomically constrained Hamilton–Pontryagin principle defined in Equations (12.2.1) and (12.2.2) under the rolling constraint in body coordinates (12.1.5) implies the following equation of motion,*

$$\left(\frac{d}{dt} - \text{ad}_\Omega^*\right)(\Pi - \lambda \diamond s) = \kappa \diamond \Gamma - \lambda \diamond \dot{s}, \quad (12.2.3)$$

where  $\Pi = \delta l / \delta \Omega$ ,  $\kappa = \delta l / \delta \Gamma$ ,  $\lambda = \delta l / \delta Y$  and  $s = s(\Gamma)$  defines the vector in the body directed from the point of rolling contact to the centre of mass.

**Remark 12.2.1 (The  $\text{ad}^*$  and  $\diamond$  operations)** The ubiquitous coadjoint action  $\text{ad}^* : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathfrak{g}^*$  and the diamond operation  $\diamond : V^* \times V \rightarrow \mathfrak{g}^*$  in Equation (12.2.3) are defined as the duals of the Lie algebra adjoint action  $\text{ad} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  and of its (left) action  $\mathfrak{g} \times V \rightarrow V$  on a vector representation space  $V$ , respectively, with respect to the corresponding pairings as

$$\langle \text{ad}_\Omega^* \Pi, \eta \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = \langle \Pi, \text{ad}_\Omega \eta \rangle_{\mathfrak{g}^* \times \mathfrak{g}}, \quad (12.2.4)$$

with  $\Omega, \eta \in \mathfrak{g}$  and  $\Pi \in \mathfrak{g}^*$  for  $\text{ad}^*$  and

$$\langle \kappa \diamond \Gamma, \eta \rangle_{\mathfrak{g}^* \times \mathfrak{g}} = \langle \kappa, -\eta \Gamma \rangle_{V^* \times V}, \quad (12.2.5)$$

with  $\eta \in \mathfrak{g}$ ,  $\kappa \in V^*$  and  $\Gamma \in V$  for  $\diamond$ . The  $\text{ad}^*$  and  $\diamond$  operations first appeared in the dual actions of  $SE(3)$  in (6.2.9). They arose again in the Clebsch Equation (9.3.8) and in the proof of the Clebsch variational principle for the Euler–Poincaré equation in Proposition 9.3.1. They featured in the definition of the cotangent-lift momentum map (11.2.1) and now they have emerged again in the nonholonomic Hamilton–Pontryagin Equation (12.2.3)  $\square$

## 254 12 : ROUND, ROLLING RIGID BODIES

**Proof.** One evaluates the variational derivatives in this constrained Hamilton's principle from the definitions of the variables as

$$\begin{aligned}\delta\Omega &= \delta(g^{-1}\dot{g}) = \frac{d\eta}{dt} + \text{ad}_\Omega\eta, \\ \delta\Gamma &= \delta(g^{-1}e_3) = -\eta(g^{-1}e_3) = -\eta\Gamma, \\ \delta Y &= \delta(g^{-1}\dot{x}) = \frac{d}{dt}(\eta s) + (\text{ad}_\Omega\eta)s.\end{aligned}\quad (12.2.6)$$

Here  $\eta := g^{-1}\delta g$  and the last equation is computed from

$$\begin{aligned}\delta(g^{-1}\dot{x}) &= -\eta g^{-1}\dot{x} + g^{-1}\delta\dot{x} \\ &= -\eta\Omega s + \frac{d}{dt}(g^{-1}\delta x) + \Omega\eta s.\end{aligned}\quad (12.2.7)$$

**Remark 12.2.2** In computing formulas (12.2.6), one must first take variations of the definitions, and only then evaluate the result on the constraint distribution defined by  $Y = g^{-1}\dot{x} = \Omega s$  and

$$g^{-1}\delta x = (g^{-1}\delta g)(rg^{-1}e_3 + \ell\chi) = \eta s,$$

which is the analogue of Equation (12.1.2). This type of subtlety in taking variations of Hamilton's principle in deriving Euler-Lagrange equations for nonholonomic systems is also discussed in [Ru2000].  $\square$

Expanding the variations of the Hamilton-Pontryagin action integral (12.2.1) using relations (12.2.6) and then integrating by parts yields

$$\begin{aligned}\delta S &= \int \left\{ \left\langle \frac{\delta l}{\delta\Omega} - \Pi, \delta\Omega \right\rangle + \left\langle \frac{\delta l}{\delta\Gamma} - \kappa, \delta\Gamma \right\rangle + \left\langle \frac{\delta l}{\delta Y} - \lambda, \delta Y \right\rangle \right. \\ &\quad \left. - \left\langle \left( \frac{d}{dt} - \text{ad}_\Omega^* \right) (\Pi - \lambda \diamond s) - \kappa \diamond \Gamma + \lambda \diamond \dot{s}, \eta \right\rangle \right\} dt \\ &\quad + \left\langle \Pi + s \diamond \lambda, \eta \right\rangle \Big|_a^b.\end{aligned}\quad (12.2.8)$$

The last entry in the integrand arises from varying in the group element  $g$  using formulas (12.2.6) obtained from relation (12.2.7). Stationarity ( $\delta S = 0$ ) for the class of action integrals  $S$  in Equations (12.2.1) and (12.2.2) for variations  $\eta$  that vanish at the endpoints now defines the Lagrange multipliers,  $\Pi = \delta l / \delta \Omega$ ,  $\kappa = \delta l / \delta \Gamma$ ,  $\lambda = \delta l / \delta Y$ , in terms of variational derivatives of the Lagrangian and thereby proves the formula for the constrained equation of motion (12.2.3) in the statement of the theorem. ■

**Remark 12.2.3 (Preparation for Noether’s theorem)** The endpoint term that arises from integration by parts in (12.2.8) will be discussed later in Remark 12.2.13 about the Jellett and Routh integrals for the cylindrically symmetric Chaplygin ball, and in Remark 12.3.5, about Noether’s theorem for the nonholonomic Euler–Poincaré equations. As usual, this endpoint term is the source of Noether’s theorem for the system. However, in nonholonomic systems the variations must respect the constraints, so Noether symmetries have more than just their geometric meaning. They also have a dynamical meaning. □

**Remark 12.2.4 (Vector notation)** For  $g \in SO(3)$  the equation of constrained motion (12.2.3) arising from stationarity ( $\delta S = 0$ ) of the action in (12.2.8) may be expressed in  $\mathbb{R}^3$  vector notation via the hat map as

$$\left(\frac{d}{dt} + \Omega \times\right)(\Pi - \lambda \times s) = \kappa \times \Gamma - \lambda \times \dot{s}. \quad (12.2.9)$$

These equations are completed by the formulas

$$\kappa = \delta l / \delta \Gamma, \quad \lambda = \delta l / \delta Y, \quad \dot{\Gamma} = -\Omega \times \Gamma,$$

and  $\dot{s}$  with  $s = s(\Gamma)$ . This vector version of Theorem 12.2.1 will yield equations of motion for two classic nonholonomic problems, the Chaplygin ball and the Euler disk. □

**Exercise.** Set  $M = \Pi - \lambda \times s$  and compute  $d(M \cdot \Omega)/dt$ . How does this simplify when  $\ell = 0$ , so that the centre of mass coincides with the centre of the sphere? ★

### 12.2.1 HP principle for the Chaplygin ball

In vector notation, the reduced Lagrangian in Equation (12.1.6) for the Chaplygin ball in body coordinates is

$$l(\Omega, \mathbf{Y}, \Gamma) = \frac{1}{2} \Omega \cdot I \Omega + \frac{m}{2} |\mathbf{Y}|^2 - m\gamma \ell \Gamma \cdot \chi. \quad (12.2.10)$$

This is the sum of the kinetic energies due to rotation and translation, minus the potential energy of gravity. One evaluates its vector-valued variational relations as

$$\begin{aligned} \Pi &= \frac{\delta l}{\delta \Omega} = I \Omega, & \kappa &= \frac{\delta l}{\delta \Gamma} = -m\gamma \ell \chi, \\ \lambda &= \frac{\delta l}{\delta \mathbf{Y}} = m \mathbf{Y} = m \Omega \times \mathbf{s}, \end{aligned} \quad (12.2.11)$$

in which the *last step* is to substitute the rolling constraint.

From their definitions for the Chaplygin ball

$$\Gamma(t) = g^{-1}(t) \mathbf{e}_3 \quad \text{and} \quad \mathbf{s} = r \Gamma(t) + \ell \chi, \quad (12.2.12)$$

one also finds the auxiliary equations,

$$\dot{\Gamma} = -\Omega \times \Gamma \quad \text{and} \quad \dot{\mathbf{s}} = -\Omega \times (\mathbf{s} - \ell \chi). \quad (12.2.13)$$

As always, the auxiliary equation for the unit vector  $\Gamma$  preserves  $|\Gamma|^2 = 1$ .

Substituting these two auxiliary equations into the stationarity condition (12.2.9) and rearranging yields the vector motion equation for the Chaplygin ball,

$$\begin{aligned} \left(\frac{d}{dt} + \boldsymbol{\Omega} \times\right) \mathbf{M} &= \underbrace{m\gamma\ell \boldsymbol{\Gamma} \times \boldsymbol{\chi}}_{\text{gravity}} - \underbrace{mr (\boldsymbol{\Omega} \times \boldsymbol{\Gamma}) \times (\boldsymbol{\Omega} \times \ell \boldsymbol{\chi})}_{\text{rolling constraint torque}} \\ &= m\ell \left[ \gamma \boldsymbol{\Gamma} \times \boldsymbol{\chi} + r \boldsymbol{\Omega} (\boldsymbol{\Omega} \cdot \boldsymbol{\chi} \times \boldsymbol{\Gamma}) \right] \end{aligned} \quad (12.2.14)$$

$$\text{with } \mathbf{M} = I\boldsymbol{\Omega} + m\mathbf{s} \times (\boldsymbol{\Omega} \times \mathbf{s}). \quad (12.2.15)$$

**Exercise. (Energy conservation)** Show that the Chaplygin ball in rolling motion without sliding conserves the following energy in body coordinates,

$$\begin{aligned} E(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) &= \frac{1}{2} \boldsymbol{\Omega} \cdot I\boldsymbol{\Omega} + \frac{m}{2} |\boldsymbol{\Omega} \times \mathbf{s}|^2 \\ &\quad + m\gamma\ell \boldsymbol{\Gamma} \cdot \boldsymbol{\chi}. \end{aligned} \quad (12.2.16)$$

This is the sum of the kinetic energies due to rotation and translation, plus the potential energy of gravity. (The rolling constraint does no work.) ★

**Exercise. (Constrained reduced Lagrangian)** Determine whether the motion Equation (12.2.14) for the Chaplygin ball also results from the constrained reduced Lagrangian,

$$l_c(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) = \frac{1}{2} \boldsymbol{\Omega} \cdot I\boldsymbol{\Omega} + \frac{m}{2} |\boldsymbol{\Omega} \times \mathbf{s}|^2 - m\gamma\ell \boldsymbol{\Gamma} \cdot \boldsymbol{\chi}, \quad (12.2.17)$$

in which the rolling constraint is applied *before* varying. ★

## 258 12 : ROUND, ROLLING RIGID BODIES

**Remark 12.2.5** The distinction between the motion equations that result from the reduced Lagrangian (12.2.10) when the rolling constraint is applied after the variations and those that result from the *constrained* reduced Lagrangian (12.2.17) on applying the rolling constraint before the variations will be discussed in Section 12.3.3.

□

**Remark 12.2.6 (The issue of sliding)** The rolling constraint exerts the torque on the body that keeps it rolling. This torque appears in the motion Equation (12.2.14) directed along the angular velocity vector  $\Omega$  and proportional to the square of its magnitude. For sufficiently rapid rotation, one could expect that friction might no longer be able to sustain the constraint torque. In this situation, additional modelling steps to include sliding would be required. However, considerations of sliding seriously affect the solution procedure, because they change the nature of the problem.

For pure rolling, one integrates the motion Equation (12.2.14) for angular momentum and solves for the body angular velocity and orientation of the vertical in the body. One then reconstructs the path on the rotation group  $g(t)$ . Finally, one applies  $g(t)$  in Equation (12.1.2) for the rolling constraint to obtain the position of the centre of mass  $\mathbf{x}(t)$ .

The ball is sliding, to the extent that its *slip velocity* is nonzero. The slip velocity of the contact point is the velocity of the contact point on the rigid body relative to the centre of mass. This is

$$\mathbf{v}_S(t) = \dot{\mathbf{x}}(t) - \dot{g}(t)(rg^{-1}(t)\mathbf{e}_3 + \ell\boldsymbol{\chi}). \quad (12.2.18)$$

Sliding introduces frictional effects that differ considerably from pure rolling. In particular, sliding introduces an additional friction force in Newton's law for the acceleration of the centre of mass,

$$m\ddot{\mathbf{x}}(t) = \mathbf{F}_S.$$

Sometimes, the sliding force  $\mathbf{F}_S$  is taken as being proportional to the slip velocity  $\mathbf{v}_S(t)$  and of opposite sign. The torque  $\mathbf{F}_S \times \mathbf{s}$  associated with the sliding force also enters the angular momentum equation. Moreover, a normal reaction force and its torque must

be determined to ensure that the sphere stays in contact with the plane. In some applications, such as the tippe top, sliding is essential. However, we shall forgo the opportunity to discuss sliding in the present text. See [GrNi2000, CiLa2007] for discussions of frictional effects for the Chaplygin ball, in the context of the tippe top phenomenon.  $\square$

**Remark 12.2.7 (The Chaplygin ball vs the heavy top)** The motion Equation (12.2.14) for the Chaplygin ball is reminiscent of the heavy-top motion Equation (8.1.2) and it has the same auxiliary Equation (8.1.3) for  $\Gamma$ . However, it also has two important differences. First, the translational kinetic energy associated with rolling of the Chaplygin ball enters effectively as an angular momentum. Second, an additional torque due to the rolling constraint appears on the right-hand side of the motion equation.

### ***Hamilton–Pontryagin principle for the heavy top***

The heavy-top equations also arise from the Hamilton–Pontryagin principle (12.2.1) with the following reduced Lagrangian in vector body coordinates, obtained by simply ignoring the kinetic energy of translation in (12.2.10) for the Chaplygin ball:

$$l(\Omega, \Gamma) = \frac{1}{2} \Omega \cdot I \Omega - m\gamma\ell \Gamma \cdot \chi. \quad (12.2.19)$$

This is just the difference between the kinetic energy of rotation and the potential energy of gravity. Substituting the vector-valued variational relations in Equation (12.2.11) and the auxiliary equation,

$$\dot{\Gamma} = -\Omega \times \Gamma, \quad (12.2.20)$$

into the stationarity condition (12.2.9) and rearranging recovers the vector motion Equation (8.1.2) for the heavy top,

$$\left( \frac{d}{dt} + \Omega \times \right) I \Omega = m\gamma\ell \Gamma \times \chi,$$

which we see now is rather simpler than Equation (12.2.14) for the motion of the Chaplygin ball. For example, one easily sees that the heavy-top equations conserve  $|\Gamma|^2$  and  $\Gamma \cdot I \Omega$ .  $\square$

**Exercise.** Show that the vector Equation (12.2.9) for non-holonomic motion in body coordinates and the auxiliary Equations (12.2.11)–(12.2.13) imply an evolution equation,

$$\frac{d}{dt}(M \cdot \Gamma) = mr\ell(\Omega \cdot \Gamma)(\Omega \cdot \chi \times \Gamma), \quad (12.2.21)$$

for the projection  $M \cdot \Gamma$  of the total body angular momentum  $M$  onto the vertical direction as seen from the body  $\Gamma = g^{-1}(t)\mathbf{e}_3$ , with

$$M = \Pi - \lambda \times s \quad \text{with} \quad \lambda = \delta l / \delta Y = mY = m\Omega \times s$$

and the vector  $\Gamma = g^{-1}(t)\mathbf{e}_3$ , which represents the vertical direction as seen from the body. ★

**Answer.** The required evolution Equation (12.2.21) follows from a direct calculation,

$$\begin{aligned} \frac{d}{dt}(M \cdot \Gamma) &= -mr(\Omega \times \Gamma) \cdot (Y \times \Gamma) \\ &= mr(\Omega \cdot \Gamma)(Y \cdot \Gamma) \\ &= mr(\Omega \cdot \Gamma)((\Omega \times s) \cdot \Gamma) \\ &= mr\ell(\Omega \cdot \Gamma)(\Omega \cdot \chi \times \Gamma), \end{aligned}$$

as obtained by using Equations (12.2.11)–(12.2.13). ▲

**Remark 12.2.8** The corresponding *spatial* quantity is the vertical component of the spatial angular momentum about the point of contact, given by

$$M \cdot \Gamma = m \cdot \mathbf{e}_3 = m_3. \quad (12.2.22)$$

Here the spatial vector  $m$  is defined by  $m = g(t)M$  for motion along the curve  $g(t) \in SO(3)$ . Relation (12.2.21) implies that the vertical component of spatial angular momentum  $m_3$  in (12.2.22) is *not* conserved, even for a cylindrical body. □



**Remark 12.2.9** The curve  $g(t)$  may be reconstructed from the solution of the motion Equation (12.2.14) for the body angular momentum  $\mathbf{\Pi}(t) := I\mathbf{\Omega}(t)$  by inverting the Legendre transformation and then solving the linear differential equation  $\dot{g}(t) = g(t)\mathbf{\Omega}(t)$  obtained from the definition of body angular velocity.  $\square$

**Exercise. (Jellett’s relation for the Chaplygin ball)** Show that the vector Equation (12.2.9) for nonholonomic motion in body coordinates implies the following evolution equation for the projection of the angular momentum  $\mathbf{\Pi}$  onto the vector  $\mathbf{s}$  directed from the point of contact to the centre of mass:

$$\frac{d}{dt}(\mathbf{\Pi} \cdot \mathbf{s}) = \left( -\mathbf{\Omega} \times (\mathbf{\Pi} - \boldsymbol{\lambda} \times \mathbf{s}) + \boldsymbol{\kappa} \times \boldsymbol{\Gamma} \right) \cdot \mathbf{s} + \mathbf{\Pi} \cdot \dot{\mathbf{s}}. \quad (12.2.23)$$

For the Chaplygin ball, Equations (12.2.11)–(12.2.13) imply *Jellett’s relation*,

$$\frac{d}{dt}(\mathbf{\Pi} \cdot \mathbf{s}) = \ell \boldsymbol{\chi} \cdot \mathbf{\Pi} \times \mathbf{\Omega}. \quad (12.2.24)$$

★

**Remark 12.2.10 (Jellett’s integral for cylindrical symmetry)** The mass distribution of Chaplygin’s ball is usually assumed to be cylindrically symmetric about  $\boldsymbol{\chi}$ . For this case, treated for example in [LyBu2009], the unit vector  $\boldsymbol{\chi}$  points along one of the principal axes and the right-hand side of (12.2.24) vanishes. Thus, rotational symmetry about  $\boldsymbol{\chi}$  produces an additional constant of motion,  $\mathbf{\Pi} \cdot \mathbf{s}$ , called *Jellett’s integral* [Je1872]. See, e.g., [Cu1998, GrNi2000, BoMaRo2004, LyBu2009] for references to the original literature and further discussions of the relation of cylindrical symmetry to the conservation of Jellett’s integral.

## 262 12 : ROUND, ROLLING RIGID BODIES

As emphasised in [LyBu2009], the concept of cylindrical symmetry has two aspects:

- (i) a principal axis passes through the geometric centre;
- (ii) the moments of inertia about all axes perpendicular to this axis are equal.

Either of these conditions may be satisfied without the other one holding. The rock'n'roller treated in [LyBu2009] satisfies condition (i) but not condition (ii). Chaplygin's ball satisfies both. An example of a body for which (ii) holds but not (i) is a light polystyrene sphere into which an arbitrarily oriented ellipsoid of heavy metal has been embedded off-centre.

□

**Exercise.** Prove that the right-hand side of (12.2.24) vanishes when both conditions (i) and (ii) of cylindrical symmetry are satisfied for an axis along  $\chi$ . ★

**Answer.** Let  $\chi = -\mathbf{E}_3$ , so that the centre of mass lies along the  $\mathbf{E}_3$  principal axis. Then

$$\ell \chi \cdot \mathbf{\Pi} \times \boldsymbol{\Omega} = -\ell(I_1 \Omega_1 \Omega_2 - I_2 \Omega_2 \Omega_1) = -\ell(I_1 - I_2) \Omega_1 \Omega_2.$$

For cylindrical symmetry about  $\mathbf{E}_3$ , we have  $I_1 = I_2$ . Consequently, the right-hand side vanishes and Jellett's integral  $\mathbf{\Pi} \cdot \mathbf{s}$  is conserved. ▲

**Remark 12.2.11 (Equivalent form of Jellett's integral)** Jellett's integral  $J = \mathbf{\Pi} \cdot \mathbf{s}$  may also be written as a linear expression in the total angular momentum,  $M$ , defined as

$$M = \mathbf{\Pi} + m \mathbf{s} \times (\boldsymbol{\Omega} \times \mathbf{s}).$$

The equivalent form of Jellett’s integral is

$$J = \mathbf{M} \cdot \mathbf{s} = r\mathbf{M} \cdot \boldsymbol{\Gamma} + \ell\mathbf{M} \cdot \boldsymbol{\chi} = rm_3 - \ell M_3 \quad (12.2.25)$$

for  $\boldsymbol{\chi} = -\mathbf{E}_3$ . Thus, Jellett’s integral is a weighted difference between the geometric spatial and body angular momentum components in the corresponding three directions. Separately,  $m_3$  and  $M_3$  are not conserved.  $\square$

**Remark 12.2.12 (Routh’s integral)** An additional integral of motion for the cylindrically symmetric Chaplygin ball was found by Routh [Ro1860],

$$R = \Omega_3 \sqrt{I_1 I_3 + m(\mathbf{s} \cdot I\mathbf{s})}.$$

The Routh integral  $R$  may also be written equivalently as a linear expression in the total angular momentum [KiPuHo2011]. Namely,

$$R = \frac{\mathbf{M} \cdot (I_1 \boldsymbol{\chi} + m(\mathbf{s} \cdot \boldsymbol{\chi})\mathbf{s})}{\sqrt{I_1 I_3 + m(\mathbf{s} \cdot I\mathbf{s})}}. \quad (12.2.26)$$

$\square$

**Exercise. (Endpoint term in vector notation)** Express the endpoint term in (12.2.8) for the Hamilton–Pontryagin principle for the Chaplygin ball in vector notation. Discuss its relation to Jellett’s integral  $J$ .  $\star$

**Answer.** In vector notation, this endpoint term becomes

$$\langle \boldsymbol{\Pi} + \mathbf{s} \diamond \lambda, \boldsymbol{\eta} \rangle = (\boldsymbol{\Pi} + m\mathbf{s} \times (\boldsymbol{\Omega} \times \mathbf{s})) \cdot \boldsymbol{\eta} = \mathbf{M} \cdot \boldsymbol{\eta}.$$

Thus, the endpoint term in (12.2.8) recovers Jellett’s integral  $J$  when  $\boldsymbol{\eta} = \mathbf{s}$ .  $\blacktriangle$

**Remark 12.2.13 (Relation of  $J$ ,  $R$  to Noether’s theorem)** Both Jellett’s integral  $J$  in (12.2.25) and Routh’s integral  $R$  in (12.2.26) arise

## 264 12 : ROUND, ROLLING RIGID BODIES

via Noether's theorem from symmetries of the Lagrangian (12.2.10) for the Chaplygin ball.

In vector notation, the reduced Lagrangian in Equation (12.2.10) for the Chaplygin ball in body coordinates has variations

$$\begin{aligned}\delta l(\boldsymbol{\Omega}, \mathbf{Y}, \boldsymbol{\Gamma}) &= \delta \boldsymbol{\Omega} \cdot \frac{\delta l}{\delta \boldsymbol{\Omega}} + \delta \mathbf{Y} \cdot \frac{\delta l}{\delta \mathbf{Y}} + \delta \boldsymbol{\Gamma} \cdot \frac{\delta l}{\delta \boldsymbol{\Gamma}} \\ &= \delta \boldsymbol{\Omega} \cdot I \boldsymbol{\Omega} + \delta \mathbf{Y} \cdot m \mathbf{Y} + \delta \boldsymbol{\Gamma} \cdot (-m\gamma\ell \boldsymbol{\chi}).\end{aligned}\quad (12.2.27)$$

The vector variations  $(\delta \boldsymbol{\Omega}, \delta \boldsymbol{\Gamma}, \delta \mathbf{Y})$  here may be expressed in terms of a single vector  $\boldsymbol{\eta}$  by rewriting (12.2.6) equivalently as

$$\begin{aligned}\delta \boldsymbol{\Omega} &= \dot{\boldsymbol{\eta}} + \boldsymbol{\Omega} \times \boldsymbol{\eta}, \\ \delta \boldsymbol{\Gamma} &= -\boldsymbol{\eta} \times \boldsymbol{\Gamma}, \\ \delta \mathbf{Y} &= \frac{d}{dt}(\boldsymbol{\eta} \times \mathbf{s}) + (\boldsymbol{\Omega} \times \boldsymbol{\eta}) \times \mathbf{s}.\end{aligned}\quad (12.2.28)$$

When  $\boldsymbol{\eta} = \mathbf{s}$ , these expressions become

$$\begin{aligned}\delta \boldsymbol{\Omega} &= \dot{\mathbf{s}} + \boldsymbol{\Omega} \times \mathbf{s} = \ell \boldsymbol{\Omega} \times \boldsymbol{\chi}, \\ \delta \boldsymbol{\Gamma} &= -\mathbf{s} \times \boldsymbol{\Gamma} = -\ell \boldsymbol{\chi} \times \boldsymbol{\Gamma}, \\ \delta \mathbf{Y} &= (\boldsymbol{\Omega} \times \mathbf{s}) \times \mathbf{s} = \mathbf{Y} \times \mathbf{s},\end{aligned}\quad (12.2.29)$$

where we have used  $\dot{\mathbf{s}} = -\boldsymbol{\Omega} \times (\mathbf{s} - \ell \boldsymbol{\chi})$  and  $\mathbf{Y} = \boldsymbol{\Omega} \times \mathbf{s}$ .

Substituting Equations (12.2.29) for the variations when  $\boldsymbol{\eta} = \mathbf{s}$  into the variation of the Lagrangian in (12.2.27) yields

$$\delta l(\boldsymbol{\Omega}, \mathbf{Y}, \boldsymbol{\Gamma}) = \ell \boldsymbol{\Omega} \times \boldsymbol{\chi} \cdot \boldsymbol{\Pi}.\quad (12.2.30)$$

Perhaps not unexpectedly, this variation vanishes and, thus, the reduced Lagrangian is invariant, for  $\boldsymbol{\eta} = \mathbf{s}$  when cylindrical symmetry ( $I_1 = I_2$ ) is imposed and the centre of mass lies along the axis of symmetry. Hence, Jellett's integral  $J$  in (12.2.25) arises via Noether's theorem from this cylindrical symmetry of the Lagrangian (12.2.10) for the Chaplygin ball.

Routh's integral  $R$  in (12.2.26) also arises via Noether's theorem from a symmetry of the Lagrangian (12.2.10) for the Chaplygin ball, but the calculation is more involved [Ki2011].  $\square$

**Exercise. (Routh’s integral and symmetry)** Show that Routh’s integral for the Chaplygin ball  $R$  in (12.2.26) arises via Noether’s theorem from a symmetry of the Lagrangian (12.2.10). ★

**Remark 12.2.14 (Chaplygin’s concentric sphere, for  $\ell \rightarrow 0$ )** When the centre of mass coincides with the centre of symmetry, then  $\ell = 0$ , the right-hand side of Equation (12.2.14) vanishes and one recovers the equations of motion for *Chaplygin’s concentric sphere*:

$$\frac{d\mathbf{\Pi}_{tot}}{dt} + \mathbf{\Omega} \times \mathbf{\Pi}_{tot} = 0, \quad \frac{d\mathbf{\Gamma}}{dt} + \mathbf{\Omega} \times \mathbf{\Gamma} = 0, \quad (12.2.31)$$

$$\mathbf{\Pi}_{tot} = I\mathbf{\Omega} + mr^2 \mathbf{\Gamma} \times (\mathbf{\Omega} \times \mathbf{\Gamma}). \quad (12.2.32)$$

The equations for Chaplygin’s concentric sphere preserve  $|\mathbf{\Pi}_{tot}|^2$ ,  $|\mathbf{\Gamma}|^2$ ,  $I\mathbf{\Omega} \cdot \mathbf{\Gamma}$  and the corresponding sum of kinetic energies in (12.2.16) when  $\ell \rightarrow 0$ . For the solution behaviour of Chaplygin’s concentric sphere, see, e.g., [Ki2001]. □

### 12.2.2 Circular disk rocking in a vertical plane

Consider an inhomogeneous circular disk that is rocking and rolling in a vertical plane, as its contact point moves backward and forward along a *straight* line. At the initial moment, the centre of mass is assumed to lie a distance  $\ell$  directly beneath the geometric centre of the disk. That is, in the initial configuration,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad (12.2.33)$$

where  $\mathbf{e}_1$  and  $\mathbf{e}_2$  are spatial unit vectors in the vertical plane. The moment of inertia is  $I = I_3$  for rotations about the third axis,  $\mathbf{e}_3$ , which is *horizontal*. The orientation of the disk is given by

## 266 12 : ROUND, ROLLING RIGID BODIES

$A(\theta) \in SO(2)$  for rotations about  $\mathbf{e}_3$  by an angle variable  $\theta(t)$  and its corresponding body angular velocity  $\widehat{\Omega} = A^{-1}\dot{A}$  is given by

$$A(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad \widehat{\Omega} = A^{-1}\dot{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \dot{\theta},$$

so that in vector notation the body angular velocity is

$$\mathbf{\Omega} = \dot{\theta} \mathbf{e}_3.$$

The remaining body vector variables for this problem are

- the vertical unit vector,  $\mathbf{e}_2$ , as seen from the body,

$$\mathbf{\Gamma}(\theta) = A^{-1}(\theta)\mathbf{e}_2 = \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix};$$

- the vector  $\mathbf{s}(\theta)$  pointing from the contact point  $C$  to the centre of mass location in the body,

$$\mathbf{s}(\theta) = r\mathbf{\Gamma}(\theta) + \ell\boldsymbol{\chi} = \begin{pmatrix} r \sin \theta \\ r \cos \theta - \ell \end{pmatrix};$$

- the rolling constraint in body coordinates (12.1.5),

$$\mathbf{Y} = \mathbf{\Omega} \times \mathbf{s} = \dot{\theta} \mathbf{e}_3 \times \mathbf{s} = \dot{\theta} \begin{pmatrix} -r \cos \theta + \ell \\ r \sin \theta \end{pmatrix};$$

- the angular momentum in the body,

$$\mathbf{\Pi} = I\mathbf{\Omega} = I_3\dot{\theta} \mathbf{e}_3;$$

- the total angular momentum,

$$\begin{aligned} \mathbf{M} &= \mathbf{\Pi} - m\mathbf{Y} \times \mathbf{s} \\ &= \left( I_3\dot{\theta} + m|\mathbf{s}(\theta)|^2\dot{\theta} \right) \mathbf{e}_3 \end{aligned}$$

so that  $\mathbf{\Omega} \times \mathbf{M} = 0$ .

## 12.2 HAMILTON–PONTRYAGIN PRINCIPLE 267

The constrained motion equation in vector form (12.2.9) with variational derivatives (12.2.11) for  $so(2)^*$  in this problem becomes

$$\left(\frac{d}{dt} + \boldsymbol{\Omega} \times\right) \mathbf{M} = -m\gamma\ell\boldsymbol{\chi} \times \boldsymbol{\Gamma} - m\mathbf{Y} \times \dot{\mathbf{s}}. \quad (12.2.34)$$

The problem of the motion of a rolling disk in a vertical plane is considerably simpler than the fully three-dimensional problem, because each term has only a horizontal ( $\mathbf{e}_3$ ) component, which implies that the vector cross product  $\boldsymbol{\Omega} \times \mathbf{M}$  vanishes. The horizontal  $\mathbf{e}_3$ -component of each of the remaining terms on the right-hand side of the Euler–Poincaré Equation (12.2.34) may be calculated as

$$\begin{aligned} \boldsymbol{\Gamma} \times m\gamma\ell\boldsymbol{\chi} &= -m\gamma\ell \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= -m\gamma\ell \sin \theta \mathbf{e}_3, \end{aligned}$$

$$\begin{aligned} \dot{\mathbf{s}} \times m\mathbf{Y} &= -(\boldsymbol{\Omega} \times r \boldsymbol{\Gamma}) \times m\mathbf{Y} \\ &= mr\dot{\theta}^2 \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix} \times \begin{pmatrix} -r \cos \theta + \ell \\ r \sin \theta \end{pmatrix} \\ &= mrl\dot{\theta}^2 \sin \theta \mathbf{e}_3. \end{aligned}$$

The right-hand side of the motion Equation (12.2.14), or equivalently Equation (12.2.34), for the Chaplygin disk then becomes

$$m\ell \left[ \gamma \boldsymbol{\Gamma} \times \boldsymbol{\chi} + r \boldsymbol{\Omega} (\boldsymbol{\Omega} \cdot \boldsymbol{\chi} \times \boldsymbol{\Gamma}) \right] = \left[ m\ell(-\gamma + r\dot{\theta}^2) \sin \theta \right] \mathbf{e}_3.$$

**Remark 12.2.15** The Jellett and Routh integrals both vanish for the Chaplygin disk.  $\square$

Assembling the terms in the motion Equation (12.2.34) for the Chaplygin disk yields

$$\frac{d}{dt} \left[ \left( I_3 + m|\mathbf{s}(\theta)|^2 \right) \dot{\theta} \right] = m\ell(-\gamma + r\dot{\theta}^2) \sin \theta, \quad (12.2.35)$$

with

$$|\mathbf{s}(\theta)|^2 = r^2 + \ell^2 - 2r\ell \cos \theta. \quad (12.2.36)$$

**Remark 12.2.16 (Chaplygin disk vs simple pendulum)** Compared to the simple pendulum equation,

$$m\ell^2\ddot{\theta} = -m\ell\gamma \sin \theta ,$$

the rocking and rolling motion Equation (12.2.35) for the Chaplygin disk has two additional terms. These terms represent the rate of change of angular momentum due to translation and the torque arising from the rolling constraint.  $\square$

**Remark 12.2.17 (Conservation of Chaplygin disk energy)** Although the constraint torque is time-dependent, it does no work. Therefore, it preserves the total energy. Energy conservation is seen by expanding the motion Equation (12.2.35), multiplying it by  $\dot{\theta}$  and rearranging to find

$$\frac{d}{dt} \left[ \frac{1}{2} I_3 \dot{\theta}^2 + \frac{1}{2} m |s(\theta)|^2 \dot{\theta}^2 + m\gamma\ell(1 - \cos \theta) \right] = 0 , \quad (12.2.37)$$

which expresses the conserved energy for the constrained motion of the Chaplygin disk rocking and rolling, backward and forward, in a vertical plane. The middle term is the translational kinetic energy.  $\square$

**Remark 12.2.18 (Euler–Lagrange equation)** The constrained reduced Lagrangian in Equation (12.2.10) for the Chaplygin ball in body coordinates may be expressed as

$$l_c(\theta, \dot{\theta}) = \frac{1}{2} I_3 \dot{\theta}^2 + \frac{1}{2} m |s(\theta)|^2 \dot{\theta}^2 + m\gamma\ell \cos \theta . \quad (12.2.38)$$

Remarkably, in this case, the equation of motion for the two-dimensional rocking Chaplygin disk is also the Euler–Lagrange equation for the constrained reduced Lagrangian (12.2.38).  $\square$

### 12.2.3 Euler’s rolling and spinning disk

One of the best-known examples of a nonholonomically constrained rigid body is the rolling circular disk. Ignoring dissipation, the



equations of motion for the rolling disk were formulated by Slesser [Sl1861] and shown to be integrable by Chaplygin [Ch1897], Appel [Ap1900] and Korteweg [Ko1900], by a transformation of variables that takes the dynamics miraculously to Legendre's equation. Nearly a century later these results were used by Cushman *et al.* [CuHeKe1995] and O'Reilly [Or1996] to examine, among other matters, the stability and bifurcations of the steady motions of flat (infinitely thin) circular disks. For additional historical accounts, see [BoMa2002, BoMaKi2002].

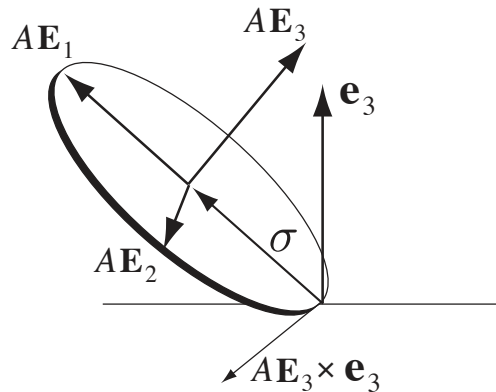
Recently, a toy called the *Euler disk* appeared, whose inventors [BeShWy1999] optimised the choice of disk weight, finite thickness, shape, material and surface so that the disk will spin and gyrate for well over a minute before coming to rest. In the last few tens of seconds before it comes to rest, it also shows a fascinating increase in its oscillation frequency which has caught the imagination of many people. See, e.g., [McDMcD2000, PeHuGr2002, LeLeGl2005]. Batista [Ba2006] found the transformation which solves the Euler disk in terms of hypergeometric functions.

Here we will use the Hamilton–Pontryagin variational principle to formulate the equations for the dynamics of the rolling and spinning flat circular disk that were originally solved by Chaplygin [Ch1897], Appel [Ap1900] and Korteweg [Ko1900] using the transformation to Legendre's equation.

Consider a flat circular disk with homogeneous mass distribution which rolls without slipping on a horizontal plane, and whose orientation is allowed to tilt away from the vertical plane. Denote its mass and radius by  $m$  and  $r$ , respectively. Let  $I = I_3$  be its moment of inertia about its axis of circular symmetry and  $I_1 = J = I_2$  be its moment of inertia about any diameter. Because its mass distribution is homogeneous, the disk's centre of mass coincides with its centre of circular symmetry.

**Remark 12.2.19** The case in which the mass distribution of the disk is unbalanced; that is, when its centre of mass does not coincide with the geometric centre of the disk, is investigated in [Ja2011]. This non-integrable, nonholonomic system combines the motion of a symmetric disk with the rolling of an unbalanced spherical ball.  $\square$

## 270 12 : ROUND, ROLLING RIGID BODIES



**Figure 12.2.** The geometry of the Euler disk is essentially the same as that of a rolling coin.

### Geometry of the Euler disk

Let  $(E_1, E_2, E_3)$  denote an orthonormal frame of unit vectors in the reference system whose origin is attached to the disk at its centre of symmetry, so that the disk lies in the  $(E_1, E_2)$  plane in body coordinates (Figure 12.2). Let  $(e_1, e_2, e_3)$  be an orthonormal coordinate frame in the inertial (spatial) system. Choose  $(e_1, e_2)$  to lie in a fixed horizontal plane and let  $e_3$  be a vertical unit vector. The motion is given by a curve  $(A(t), x(t)) \in SE(3)$  where  $A(t) \in SO(3)$  and  $x(t) \in \mathbb{R}^3$ . The origin of the reference system lies at the disk's centre of mass, so that  $x(t)$  gives the location of its centre of mass in space.

An element of  $SE(3)$  maps a generic point  $P$  in the body to a point  $A(t)P + x(t)$  in space at time  $t$ . The height of the centre of mass at a given time is  $z = \langle x(t), e_3 \rangle$ . The matrix  $A(t) \in SO(3)$  defines the attitude of the body relative to its reference configuration. The vector in the spatial frame,  $A(t)E_3$ , will be normal to the disk. Consequently, the vector  $e_3 \times A(t)E_3$  will be tangent to the edge of the disk at the point of contact with the plane. (One assumes  $A(t)E_3 \neq \pm e_3$ , so that the disk keeps rolling. The case  $\langle e_3, A(t)E_3 \rangle^2 = 1$  will turn out to be a singular situation.)

**Definition 12.2.1 (Vector to the centre of mass)** Let  $s(t) \in \mathbb{R}^3$  denote the vector in the body frame directed at time  $t$  from the point of contact  $C$  to the geometric centre of the disk. The corresponding vector in the spatial frame,

$$\sigma(t) = A(t)s(t),$$

is normal both to  $A(t)E_3$  and to the tangent direction  $e_3 \times A(t)E_3$ . Consequently,  $\sigma(t)$  lies along  $u(t) := A(t)E_3 \times (e_3 \times A(t)E_3)$  in the spatial frame and because it points from the edge to the centre, its magnitude is equal to the radius of the disk,  $r$ . Thus,

$$\sigma(t) = A(t)s(t) = r \frac{u(t)}{|u(t)|} \quad (12.2.39)$$

with

$$u(t) := A(t)E_3 \times (e_3 \times A(t)E_3). \quad (12.2.40)$$

The vector  $u$  has squared magnitude

$$|u|^2 = 1 - \langle e_3, A(t)E_3 \rangle^2 = 1 - \langle A^{-1}(t)e_3, E_3 \rangle^2 \neq 0.$$

The height of the centre of mass at a given time is

$$z(t) = \langle A(t)s(t), e_3 \rangle = r|u(t)|, \quad (12.2.41)$$

so the disk is lying flat ( $z = 0$ ) in the case where  $\langle e_3, A(t)E_3 \rangle^2 = 1$ .

**Definition 12.2.2 (Rolling constraint)** Physically, rolling of the disk means that the velocity of its centre of mass  $\dot{x}(t)$  in the spatial frame at any instant  $t$  must be obtained by a rotation of the spatial vector  $A(t)s$  at angular frequency  $\hat{\omega}(t) = \dot{A}A^{-1}(t) \in so(3)$  around the fixed contact point  $C$ . Consequently, the **rolling constraint** is imposed by requiring

$$\dot{x}(t) = \dot{A}A^{-1}\sigma(t) = \hat{\omega}\sigma(t) = \omega \times r \frac{u(t)}{|u(t)|}, \quad (12.2.42)$$

where the hat map  $\hat{\cdot} : \mathbb{R}^3 \rightarrow so(3)$  was used in the last step for the (right-invariant) spatial angular velocity. Equivalently, one may rewrite this rolling constraint using Equation (12.2.39) as

$$\dot{x}(t) = r\dot{A}A^{-1} \frac{u(t)}{|u(t)|}. \quad (12.2.43)$$

This formula defines the **spatial nonholonomic constraint distribution**  $\mathcal{D}$  for this problem.

**The Hamilton–Pontryagin approach for the Euler disk**

As before, the Hamilton–Pontryagin method proceeds by first assembling the following ingredients:

- Potential energy: The potential energy due to gravity for the disk is  $mgz(t)$ , where

$$z(t) = \langle \sigma(t), e_3 \rangle = \langle As, e_3 \rangle = \langle s(\Gamma), \Gamma(t) \rangle \quad (12.2.44)$$

is the height of the centre of mass above the horizontal plane and  $g$  is the constant acceleration of gravity.

- Kinetic energy: The kinetic energy of the disk is the sum of the energy of translation of its centre of mass and the energy of its rotation about the centre of mass.
- Lagrangian: The Lagrangian is defined as the difference between the total kinetic energy and the gravitational potential energy,

$$L = \frac{1}{2} \langle \Omega, I\Omega \rangle + \frac{m}{2} |\dot{x}|^2 - mg \langle s, A^{-1}e_3 \rangle,$$

where  $\Omega = A^{-1}\dot{A}$  is the (left-invariant) body angular velocity.

- Rolling constraint in the body: The spatial constraint distribution (12.2.43) transforms into body coordinates upon using Equation (12.2.39) and setting

$$Y = A^{-1}\dot{x}, \quad \Gamma = A^{-1}e_3, \quad (12.2.45)$$

$$s(\Gamma) = r \frac{A^{-1}(t)u(t)}{|u(t)|} = r \frac{E_3 \times (\Gamma(t) \times E_3)}{\sqrt{1 - \langle \Gamma(t), E_3 \rangle^2}}, \quad (12.2.46)$$

with vectors  $Y$ ,  $\Gamma$ ,  $s(\Gamma)$  each in  $\mathbb{R}^3$ . Note that

$$s(\Gamma) \cdot \Gamma(t) = r|u(t)| = z(t)$$

yields the height of the centre of mass and  $|s(\Gamma)|^2 = r^2$ .

In this notation, the rolling constraint in body coordinates (12.2.43) simplifies to an algebraic expression,

$$Y = \Omega \times s(\Gamma). \quad (12.2.47)$$

- Reduced Lagrangian in body coordinates: The reduced Lagrangian is defined in body coordinates by

$$l(\Omega, Y, \Gamma) = \frac{1}{2} \langle \Omega, I\Omega \rangle + \frac{m}{2} |Y|^2 - mg \langle s(\Gamma), \Gamma \rangle. \quad (12.2.48)$$

These considerations have expressed a spatial Lagrangian

$$L : TG \times TV \times V \rightarrow \mathbb{R}$$

in body form as  $l : \mathfrak{g} \times TV \times V \rightarrow \mathbb{R}$ , plus the various constraint relations.

### Vector equation for the Euler disk

One may now apply the results of the nonholonomic Hamilton-Pontryagin Theorem 12.2.1 to obtain the motion equation for the Euler disk, in the vector form (12.2.9),

$$\left( \frac{d}{dt} + \Omega \times \right) (\Pi - \lambda \times s) = \kappa \times \Gamma - \lambda \times \dot{s}. \quad (12.2.49)$$

This equation arises from the Lagrangian equivalent to (12.2.48) in terms of vector quantities,

$$L = \frac{1}{2} \Omega \cdot I\Omega + \frac{m}{2} |Y|^2 - mg s(\Gamma) \cdot \Gamma, \quad (12.2.50)$$

and its vector-valued variational relations,

$$\Pi = \frac{\delta l}{\delta \Omega} = I\Omega, \quad \kappa = \frac{\delta l}{\delta \Gamma} = -mg s(\Gamma), \quad (12.2.51)$$

$$\lambda = \frac{\delta l}{\delta Y} = mY = m\Omega \times s(\Gamma). \quad (12.2.52)$$

Recall the definitions,

$$\Gamma(t) = A^{-1}(t)e_3 \quad \text{and} \quad s(\Gamma) = r \frac{\mathbf{E}_3 \times (\Gamma \times \mathbf{E}_3)}{\sqrt{1 - (\Gamma(t) \cdot \mathbf{E}_3)^2}}, \quad (12.2.53)$$

## 274 12 : ROUND, ROLLING RIGID BODIES

so that

$$\boldsymbol{\Gamma} \cdot \mathbf{s} = \mathbf{e}_3 \cdot \boldsymbol{\sigma} = r \sqrt{1 - (\boldsymbol{\Gamma}(t) \cdot \mathbf{E}_3)^2}. \quad (12.2.54)$$

In vector notation, the two vectors  $\boldsymbol{\Gamma}$  and  $\mathbf{s}$  satisfy the auxiliary equations,

$$\dot{\boldsymbol{\Gamma}} = -\boldsymbol{\Omega} \times \boldsymbol{\Gamma} \quad (12.2.55)$$

and

$$\dot{\mathbf{s}} = \frac{r}{\sqrt{1 - (\boldsymbol{\Gamma} \cdot \mathbf{E}_3)^2}} \left( (\mathbf{E}_3 \times \boldsymbol{\Gamma}) \cdot (\boldsymbol{\Omega} \times \boldsymbol{\Gamma}) \right) \boldsymbol{\Gamma} \times \mathbf{E}_3, \quad (12.2.56)$$

so that

$$\boldsymbol{\Gamma} \cdot \dot{\mathbf{s}} = 0.$$

A similar formula holds for the variational derivative, namely

$$\boldsymbol{\Gamma} \cdot \delta \mathbf{s} = 0,$$

which immediately implies the variational derivative

$$\delta l / \delta \boldsymbol{\Gamma} = -mg \mathbf{s}(\boldsymbol{\Gamma})$$

in Equations (12.2.52). The auxiliary equations preserve  $|\boldsymbol{\Gamma}|^2 = 1$  and  $|\mathbf{s}|^2 = r^2$ .

On substituting these vector relations into (12.2.49) one obtains the motion equation for the Euler disk,

$$\begin{aligned} \left( \frac{d}{dt} + \boldsymbol{\Omega} \times \right) \left( I \boldsymbol{\Omega} + m \mathbf{s}(\boldsymbol{\Gamma}) \times (\boldsymbol{\Omega} \times \mathbf{s}(\boldsymbol{\Gamma})) \right) & \quad (12.2.57) \\ & = \underbrace{mg \boldsymbol{\Gamma} \times \mathbf{s}(\boldsymbol{\Gamma})}_{\text{gravity}} + \underbrace{m \dot{\mathbf{s}} \times (\boldsymbol{\Omega} \times \mathbf{s}(\boldsymbol{\Gamma}))}_{\text{rolling constraint torque}}. \end{aligned}$$

Thus, the motion equation for the Euler disk has the same form as Equation (12.2.14) for the Chaplygin ball. However, the rolling constraint torque is considerably more intricate for the Euler disk, because of the complexity of expression (12.2.56) for  $\dot{\mathbf{s}}(\boldsymbol{\Gamma}, \boldsymbol{\Omega})$ .

**Exercise. (Energy)** Show that the Euler disk in rolling motion without sliding conserves the following energy in body coordinates:

$$E(\Omega, \Gamma) = \frac{1}{2} \langle \Omega, I\Omega \rangle + \frac{m}{2} |\Omega \times \mathbf{s}|^2 + mg \langle \Gamma, \mathbf{s}(\Gamma) \rangle.$$

This is the sum of the kinetic energies of rotation and translation, plus the potential energy of gravity. (Recall that the rolling constraint does no work.) ★

**Exercise. (Jellett’s integral for the Euler disk)** Show that the Euler disk admits Jellett’s integral  $(\Pi \cdot \mathbf{s})$  as an additional constant of motion because of the cylindrical symmetry of its mass distribution.

Hint: Compute the analogue for the Euler disk of Equation (12.2.24) for the Chaplygin ball. ★

## 12.3 Nonholonomic Euler–Poincaré reduction

The treatments of the Chaplygin ball and the Euler disk in the previous section are consistent with an Euler–Poincaré version of the standard *Lagrange–d’Alembert principle* [Sc2002, Bl2004]. As in the cases of the Chaplygin ball and the Euler disk, the configuration space will be taken to be a semidirect-product space  $\mathfrak{G} = G \ltimes V$  such as the special Euclidean group  $SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$  of orientations and positions, whose various adjoint and coadjoint ( $AD, Ad, ad, Ad^*$  and  $ad^*$ ) actions needed for Euler–Poincaré theory are discussed in Chapter 7.

### 12.3.1 Semidirect-product structure

A semidirect-product structure involves a Lie group  $G$  and a representation vector space  $V$  for the action of  $G$ , which we choose to be a left action. Let  $\mathfrak{S}$  denote the semidirect product  $G \ltimes V$ . Topologically,  $\mathfrak{S}$  is the direct product of spaces  $G \times V$ . The left group action on  $\mathfrak{S}$  is given as in (6.1.1) for the special Euclidean group of rotations and translations,  $SE(3)$ ,

$$(g_1, y_1) \cdot (g_2, y_2) = (g_1 g_2, g_1 y_2 + y_1), \quad (12.3.1)$$

where for  $g \in G$  and  $y \in V$ , the left action of  $G$  on  $V$  is denoted as  $gy$ . The formulas for the right action of  $G$  on  $V$  may be derived by making the appropriate adjustments, as discussed in Section 6.3.3. The induced left action of  $\mathfrak{S}$  on its tangent space  $\Phi : \mathfrak{S} \times T\mathfrak{S} \rightarrow T\mathfrak{S}$  is given by

$$(h, y) \cdot (g, x) = (hg, h\dot{g}, hx + y, h\dot{x}), \quad (12.3.2)$$

where dot ( $\dot{\cdot}$ ) in  $(\dot{g}(t), \dot{x}(t))$  denotes the time derivative along a curve  $(g(t), x(t))$  in  $\mathfrak{S}$ . Let  $\mathfrak{g}$  denote the Lie algebra of the Lie group  $G$  and  $\mathfrak{s}$  the Lie algebra of  $\mathfrak{S}$ . (As a vector space, the Lie algebra  $\mathfrak{s}$  is  $\mathfrak{g} \times V$ .)

The semidirect-product Lie bracket on  $\mathfrak{s}$  is of the form in Equation (6.3.1),

$$[(\xi_1, Y_1), (\xi_2, Y_2)] = ([\xi_1, \xi_2], \xi_1 Y_2 - \xi_2 Y_1), \quad (12.3.3)$$

where the left induced action from  $\mathfrak{g} \times V \rightarrow V$  is denoted by concatenation from the left, as in  $\xi_1 Y_2$ . The dual action  $V^* \times V \rightarrow \mathfrak{g}^*$  for fixed  $b \in V^*$  and  $v \in V$  is defined for a symmetric real pairing  $\langle \cdot, \cdot \rangle : V^* \times V \rightarrow \mathbb{R}$  by the diamond operation:

$$\langle b, -\xi v \rangle = \langle b \diamond v, \xi \rangle. \quad (12.3.4)$$

**Definition 12.3.1** *The left- and right-invariant angular velocities  $\xi, \omega \in \mathfrak{g}$  are defined by*

$$\xi = g^{-1} \dot{g} \quad \text{and} \quad \omega = \dot{g} g^{-1}. \quad (12.3.5)$$



**Definition 12.3.2 (Linear constraints in space and body)** *Linear constraints in spatial form are expressed as*

$$\dot{x}(t) = \omega(k a_0 + g(t)\ell\chi), \quad (12.3.6)$$

and are defined by left Lie algebra and Lie group actions of  $\omega \in \mathfrak{g}$  and  $g \in G$  on fixed elements  $a_0$  and  $\chi$  of  $V$ , with constant  $k$ . In terms of the body variables defined by

$$\Gamma = g^{-1}a_0, \quad Y = g^{-1}\dot{x}, \quad (12.3.7)$$

the linear constraint in body form corresponding to (12.3.6) may be expressed in its body form as

$$\begin{aligned} Y &= g^{-1}\dot{x} = g^{-1}\dot{g}g^{-1}(k a_0 + g(t)\ell\chi) \\ &= (g^{-1}\dot{g})(k g^{-1}a_0 + \ell\chi) = \xi s(\Gamma), \end{aligned} \quad (12.3.8)$$

where  $s(\Gamma) = k\Gamma + \ell\chi$ .

**Remark 12.3.1 (Auxiliary equation for  $\Gamma$ )** By its definition, the quantity  $\Gamma(t) = g^{-1}(t)a_0$  satisfies the auxiliary equation,

$$\frac{d\Gamma}{dt} = -\xi(t)\Gamma(t). \quad (12.3.9)$$

Hence, the linear combination  $s(\Gamma) = k\Gamma + \ell\chi$  satisfies

$$\dot{s}(\Gamma) = -k\xi(t)\Gamma(t). \quad (12.3.10)$$

□

**Definition 12.3.3 (Lagrangians on semidirect products)** *Denote the Lagrangian by*

$$L_0 : T\mathfrak{S} \rightarrow \mathbb{R}.$$

Introduce parameter dependence into the Lagrangian  $L_0(g, \dot{g}, x, \dot{x})$  by setting it equal to the Lagrangian  $L(g, \dot{g}, x, \dot{x}, a_0)$ ,

$$L : T\mathfrak{S} \times V \rightarrow \mathbb{R},$$

with a fixed parameter  $a_0$  for the last factor of  $V$ .

**Remark 12.3.2** In what follows, we shall assume that Lagrangians  $L_0$  and  $L$  are independent of the position  $x$ . (This corresponds, for example, to rolling on a horizontal plane.)  $\square$

**Definition 12.3.4 (Reduced and constrained Lagrangians)** *The reduced Lagrangian is defined in terms of the  $x$ -independent Lagrangian  $L$  as*

$$l : \mathfrak{g} \times V, \quad l(\xi, Y, \Gamma) = L(g, g\xi, gY, g\Gamma) = L(g, \dot{g}, \dot{x}, a_0). \quad (12.3.11)$$

*The constrained reduced Lagrangian is defined by evaluating the reduced Lagrangian on the constraint distribution, as*

$$l_c : \mathfrak{g} \times V, \quad l_c(\xi, \xi s(\Gamma), \Gamma) = l(\xi, Y, \Gamma)|_{Y=\xi s(\Gamma)}. \quad (12.3.12)$$

*Thus, the reduced Lagrangian  $l$  is found by transforming from spatial to body variables in the Lagrangian  $L$ . Then the constrained reduced Lagrangian  $l_c$  is defined by evaluating the reduced Lagrangian  $l$  on the constraint distribution.*

### 12.3.2 Euler–Poincaré theorem

The definitions and assumptions for the constraints of the previous section form the semidirect-product framework that leads to the nonholonomic Euler–Poincaré theorem found in [Sc2002].

**Theorem 12.3.1 (Nonholonomic Euler–Poincaré theorem)** *The following three statements are equivalent.*

- *The curve  $(g(t), x(t)) \in \mathfrak{S}$  satisfies the Lagrange–d’Alembert principle for the Lagrangian  $L$ . That is,  $(\dot{g}(t), \dot{x}(t)) \in \mathfrak{S}$  satisfies the constraint (12.3.6) and*

$$\delta \int L(g(t), \dot{g}(t), \dot{x}(t), a_0) dt = 0,$$

*where  $\delta g$  is an independent variation vanishing at the endpoints, and the variation  $\delta x$  satisfies the constraint condition,*

$$\delta x = (\delta g g^{-1})(k a_0 + g \ell \chi). \quad (12.3.13)$$

## 12.3 NONHOLONOMIC EULER–POINCARÉ REDUCTION 279

- The reduced Lagrangian  $l$  defined on  $\mathfrak{s} \times V$  satisfies a constrained variational principle

$$\delta \int l(\xi(t), Y(t), \Gamma(t)) = 0, \quad (12.3.14)$$

where the variations take the form

$$\begin{aligned} \delta \xi &= \dot{\eta} + \text{ad}_\xi \eta, \\ \delta \Gamma &= -\eta \Gamma, \\ \delta Y &= \dot{\eta} s(\Gamma) + (\text{ad}_\xi \eta) s(\Gamma) + \eta \dot{s}(\Gamma), \end{aligned} \quad (12.3.15)$$

and  $\eta(t) = g^{-1} \delta g \in \mathfrak{g}$  is an independent variation vanishing at the endpoints.

- The following nonholonomic Euler–Poincaré equations hold on  $\mathfrak{g}^* \times V^*$ :

$$\frac{d}{dt} \frac{\partial l_c}{\partial \xi} - \text{ad}_\xi^* \frac{\partial l_c}{\partial \xi} = \frac{\partial l}{\partial \Gamma} \diamond \Gamma + \dot{s} \diamond \frac{\partial l}{\partial Y}, \quad (12.3.16)$$

$$\text{and } \frac{d\Gamma}{dt} + \xi \Gamma = 0, \quad (12.3.17)$$

where the constrained Lagrangian  $l_c$  is obtained by evaluating the reduced Lagrangian  $l$  on the constraint distribution.

**Remark 12.3.3** The theorem states that the Lagrange–d’Alembert principle on  $T\mathfrak{S}$  and equations of motion on  $T^*\mathfrak{S}$  is equivalent to a constrained variational principle on  $\mathfrak{s} \times V$  and Euler–Poincaré equations on  $\mathfrak{g}^* \times V^*$ .  $\square$

**Definition 12.3.5 (Lagrange–d’Alembert principle)** The Lagrange–d’Alembert principle on  $T\mathfrak{S}$  for  $L$  in Theorem 12.3.1 is the stationary principle,

$$\delta \int L(g(t), \dot{g}(t), \dot{x}(t)) dt = 0,$$

where  $\delta g(t)$  is an independent variation vanishing at the endpoints, and the variations  $\delta x$  of coordinates  $x(t) \in V$  must be consistent with the velocity constraint (12.3.6), rewritten now as a relation between one-forms,

$$dx = (dg g^{-1})(ka_0 + g\ell\chi) =: A(g, a_0) dg.$$

## 280 12 : ROUND, ROLLING RIGID BODIES

Correspondingly, the variations of  $L_0$  in  $x$  must satisfy

$$\delta x(t) = (\delta g g^{-1})(ka_0 + g\ell\chi) =: A(g, a_0) \delta g. \quad (12.3.18)$$

Stationarity under these variations of  $\int L_0 dt$  produces

$$\frac{d}{dt} \frac{\partial L_{0c}}{\partial \dot{g}} - \frac{\partial L_{0c}}{\partial g} = \frac{\partial L_0}{\partial \dot{x}} B \dot{g},$$

where the constrained Lagrangian  $L_{0c}$  is obtained by evaluating  $L_0$  on the velocity constraint  $\dot{x} = A(g, a_0) \dot{g}$  as

$$L_{0c}(g(t), \dot{g}(t)) = L_0(g(t), \dot{g}(t), \dot{x}(t)) \Big|_{\dot{x}=A(g,a_0)\dot{g}(t)}. \quad (12.3.19)$$

The quantity  $B$  appearing in this equation is viewed geometrically as the curvature associated with the connection  $A$  that defines the constraint distribution. The curvature two-form is defined by applying the exterior derivative to the connection one-form as

$$d(A(g, a_0)dg) = Bdg \wedge dg \quad \text{so that} \quad B = \text{skew} \left( \frac{\partial A}{\partial g} \right).$$

Of course, the variations of  $L_0$  were taken before rewriting the result using derivatives of  $L_{0c}$ . Details of the computations leading to these formulas and more discussions of their meaning are provided in [BlKrMaMu1996].

**Proof.** By comparing definitions, one finds that the integrand for the Lagrange–d’Alembert principle for  $L$  on  $\mathfrak{S}$  in the stationary principle is equal to the integrand  $l$  on  $\mathfrak{s} \times V$ . We must compute what the variations on the group  $\mathfrak{S}$  imply on the reduced space  $\mathfrak{s} \times V$ . Define  $\eta = g^{-1}\delta g$ . As for the pure Euler–Poincaré theory with left-invariant Lagrangians, the proof of the variational formula expressing  $\delta\xi$  in terms of  $\eta$  proceeds by direct computation.

$$\delta\xi = \delta(g^{-1}\dot{g}) = \dot{\eta} + \text{ad}_\xi \eta.$$

For the  $\delta\Gamma$  variation, one calculates

$$\delta\Gamma = \delta(g^{-1})a_0 = -g^{-1}\delta g g^{-1}a_0 = -\eta\Gamma.$$

The  $\delta Y$  variation is computed as follows:

$$\begin{aligned}\delta Y &= \delta(g^{-1})\dot{x} + g^{-1}\delta\dot{x} \\ &= -g^{-1}\delta g g^{-1}\dot{x} + \frac{d}{dt}(g^{-1}\delta x) - (g^{-1})'\delta x.\end{aligned}$$

Upon substituting the constraint  $\delta x = \delta g g^{-1}(ka_0 + g\ell\chi)$  and expanding terms, one finds  $g^{-1}\delta x = \eta s$  and  $g^{-1}\dot{x} = \xi s$  so that

$$\begin{aligned}\delta Y &= -\eta\xi s + \frac{d}{dt}(\eta s) + \xi\eta s \\ &= (\dot{\eta} + \text{ad}_\xi\eta)s + \eta\dot{s}.\end{aligned}$$

After this preparation, the nonholonomic EP equation finally emerges from a *direct* computation of the variation,  $\delta S$ , of the action  $S = \int l(\xi, Y, \Gamma) dt$ :

$$\begin{aligned}\delta S &= \int \left\langle \frac{\partial l}{\partial \xi}, \delta \xi \right\rangle + \left\langle \frac{\partial l}{\partial Y}, \delta Y \right\rangle + \left\langle \frac{\partial l}{\partial \Gamma}, \delta \Gamma \right\rangle dt \quad (12.3.20) \\ &= \int \left\langle \frac{\partial l}{\partial \xi}, \dot{\eta} + \text{ad}_\xi\eta \right\rangle + \left\langle \frac{\partial l}{\partial Y}, \dot{\eta}s + (\text{ad}_\xi\eta)s + \eta\dot{s} \right\rangle \\ &\quad + \left\langle \frac{\partial l}{\partial \Gamma}, -\eta\Gamma \right\rangle dt \\ &= \int \left\langle -\frac{d}{dt}\frac{\partial l}{\partial \xi} + \text{ad}_\xi^*\frac{\partial l}{\partial \xi} - \frac{d}{dt}\left(s \diamond \frac{\partial l}{\partial Y}\right) \right. \\ &\quad \left. + \text{ad}_\xi^*\left(s \diamond \frac{\partial l}{\partial Y}\right) + \dot{s} \diamond \frac{\partial l}{\partial Y} + \frac{\partial l}{\partial \Gamma} \diamond \Gamma, \eta \right\rangle dt \\ &\quad + \left\langle \left( \frac{\partial l}{\partial \xi} + s \diamond \frac{\partial l}{\partial Y} \right), \eta \right\rangle \Big|_a^b. \quad (12.3.21)\end{aligned}$$

Setting the variation  $\delta S$  equal to zero for any choice of  $\eta$  that vanishes at the endpoints in time,  $\eta(a) = 0 = \eta(b)$ , yields the nonholonomic Euler–Poincaré equation on  $\mathfrak{s} \times V$ ,

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \xi} + s \diamond \frac{\partial l}{\partial Y}\right) - \text{ad}_\xi^*\left(\frac{\partial l}{\partial \xi} + s \diamond \frac{\partial l}{\partial Y}\right) = \dot{s} \diamond \frac{\partial l}{\partial Y} + \frac{\partial l}{\partial \Gamma} \diamond \Gamma. \quad (12.3.22)$$

**Remark 12.3.4 (Equivalence theorem)** On comparison, one sees that the nonholonomic Euler–Poincaré Equation (12.3.22) coincides with the motion Equation (12.2.3) derived in Theorem 12.2.1 using the Hamilton–Pontryagin method.  $\square$

**Remark 12.3.5 (Noether’s theorem)** If the variation of the action  $\delta S$  vanishes because of a Lie symmetry of the Lagrangian, so that its first line in (12.3.20) vanishes,

$$\left\langle \frac{\partial l}{\partial \xi}, \delta \xi \right\rangle + \left\langle \frac{\partial l}{\partial \Gamma}, \delta \Gamma \right\rangle + \left\langle \frac{\partial l}{\partial Y}, \delta Y \right\rangle = 0, \quad (12.3.23)$$

with variations  $\delta \xi$ ,  $\delta \Gamma$  and the modified variation  $\delta Y$  given in (12.3.15) in terms of the symmetry generator  $\eta$ , then according to Noether’s theorem the endpoint term in (12.3.21)

$$\langle M, \eta \rangle \quad \text{with} \quad M := \frac{\partial l}{\partial \xi} + s \diamond \frac{\partial l}{\partial Y}$$

must be a constant of the motion governed by the constrained Euler–Poincaré Equation (12.3.22).

Notice that the Noether constant of motion  $\langle M, \eta \rangle$  is linear in the total angular momentum,  $M$ , just as for the Jellett and Routh integrals (12.2.25) and (12.2.26), respectively, for the cylindrically symmetric Chaplygin ball whose centre of mass lies on the axis of symmetry. The corresponding Lie algebra actions by  $\eta$  for those two constants of motion are symmetries of their Lagrangian (12.2.10). However, as shown in [Ki2011] and [KiPuHo2011], these symmetries should be understood in the sense of infinitesimal invariance of the Lagrangian in (12.3.23) under variations (12.3.15) that respect the nonholonomic rolling constraint (12.3.8) and the dynamical Equations (12.3.9) and (12.3.10) that govern the auxiliary quantities  $\Gamma$  and  $s(\Gamma)$ . See Remark 12.2.13 for the explicit calculations in the example of Jellett’s integral for the Chaplygin ball.  $\square$

### 12.3.3 Constrained reduced Lagrangian

The nonholonomic EP equation may also be written in terms of the constrained reduced Lagrangian. It is defined in (12.3.12) as

$$l_c(\xi, \Gamma) := l(\xi, Y, \Gamma) \Big|_{Y=\xi s(\Gamma)}.$$

Its differential is given by

$$\begin{aligned}
 dl_c &= \left\langle \frac{\partial l_c}{\partial \xi}, d\xi \right\rangle + \left\langle \frac{\partial l_c}{\partial \Gamma}, d\Gamma \right\rangle \\
 &= \left\langle \frac{\partial l}{\partial \xi}, d\xi \right\rangle + \left\langle \frac{\partial l}{\partial Y}, d\xi s(\Gamma) + \xi ds(\Gamma) \right\rangle + \left\langle \frac{\partial l}{\partial \Gamma}, d\Gamma \right\rangle \\
 &= \left\langle \frac{\partial l}{\partial \xi} - \frac{\partial l}{\partial Y} \diamond s(\Gamma), d\xi \right\rangle + \left\langle \frac{\partial l}{\partial \Gamma} + \left( \xi \frac{\partial s}{\partial \Gamma} \right)^T \frac{\partial}{\partial Y}, d\Gamma \right\rangle.
 \end{aligned}$$

This calculation implies the following relations among the derivatives of the reduced Lagrangian  $l$  evaluated on the constraint, and the derivatives of the constrained reduced Lagrangian  $l_c$ :

$$\frac{\partial_c}{\partial \xi} = \left( \frac{\partial}{\partial \xi} + s \diamond \frac{\partial}{\partial Y} \right) \text{ and } \frac{\partial_c}{\partial \Gamma} = \frac{\partial l}{\partial \Gamma} + \left( \xi \frac{\partial s}{\partial \Gamma} \right)^T \frac{\partial}{\partial Y}. \quad (12.3.24)$$

Substituting the first relation into (12.3.22) and recalling the advection equation for  $\Gamma$  yields the expression (12.3.17) in the statement of the theorem for the nonholonomic Euler–Poincaré equations, namely,

$$\frac{d}{dt} \frac{\partial l_c}{\partial \xi} - \text{ad}_\xi^* \frac{\partial l_c}{\partial \xi} = \frac{\partial l}{\partial \Gamma} \diamond \Gamma - \frac{\partial l}{\partial Y} \diamond \dot{s}, \quad (12.3.25)$$

$$\text{and } \frac{d\Gamma}{dt} + \xi \Gamma = 0. \quad (12.3.26)$$

This completes the proof of the nonholonomic Euler–Poincaré Theorem 12.3.1. ■

**Corollary 12.3.1 (Energy conservation)** *The nonholonomic Euler–Poincaré motion Equation (12.3.25) conserves the following energy in body coordinates:*

$$E(\xi, \Gamma) = \left\langle \frac{\partial l_c}{\partial \xi}, \xi \right\rangle - l_c(\xi, \Gamma). \quad (12.3.27)$$

*This is the Legendre transform of the constrained reduced Lagrangian.*

## 284 12 : ROUND, ROLLING RIGID BODIES

**Proof.** Conservation of the energy  $E(\xi, \Gamma)$  in (12.3.27) may be verified directly by differentiating it with respect to time and following a sequence of steps that use the auxiliary equation  $d\Gamma/dt = -\xi\Gamma$ , the equation of nonholonomic motion (12.3.25), the relations (12.3.24) and the antisymmetry of the adjoint (ad) and diamond ( $\diamond$ ) operations:

$$\begin{aligned} \frac{dE}{dt} &= \left\langle \frac{d}{dt} \frac{\partial l_c}{\partial \xi}, \xi \right\rangle - \left\langle \frac{\partial l_c}{\partial \Gamma}, \frac{d\Gamma}{dt} \right\rangle \\ \text{by (12.3.25)} &= \left\langle \text{ad}_\xi^* \frac{\partial l_c}{\partial \xi} + \frac{\partial l}{\partial \Gamma} \diamond \Gamma + \dot{s} \diamond \frac{\partial l}{\partial Y}, \xi \right\rangle + \left\langle \frac{\partial l_c}{\partial \Gamma}, \xi \Gamma \right\rangle \\ \text{by ad, ad}^* \text{ and } \diamond &= \left\langle \dot{s} \diamond \frac{\partial l}{\partial Y}, \xi \right\rangle + \left\langle \frac{\partial l_c}{\partial \Gamma} - \frac{\partial l}{\partial \Gamma}, \xi \Gamma \right\rangle. \end{aligned}$$

By the second relation in (12.3.24) and the auxiliary equation  $d\Gamma/dt = -\xi\Gamma$ , one finds

$$\begin{aligned} \left\langle \frac{\partial l_c}{\partial \Gamma} - \frac{\partial l}{\partial \Gamma}, \xi \Gamma \right\rangle &= - \left\langle \left( \xi \frac{\partial s}{\partial \Gamma} \right)^T \frac{\partial}{\partial Y}, \frac{d\Gamma}{dt} \right\rangle \\ &= - \left\langle \frac{\partial}{\partial Y}, \xi \frac{\partial s}{\partial \Gamma} \frac{d\Gamma}{dt} \right\rangle = - \left\langle \frac{\partial}{\partial Y}, \xi \dot{s}(\Gamma) \right\rangle. \end{aligned}$$

Thus,

$$\frac{dE}{dt} = \left\langle \dot{s} \diamond \frac{\partial l}{\partial Y} + \frac{\partial}{\partial Y} \diamond \dot{s}, \xi \right\rangle = 0,$$

in which the last step invokes the *antisymmetry* of the diamond operation when paired with a vector field,  $\eta$ , which holds because  $0 = \mathcal{L}_\eta \langle \lambda, s \rangle = \langle \lambda, \eta s \rangle + \langle \eta \lambda, s \rangle = -\langle \lambda \diamond s, \eta \rangle - \langle s \diamond \lambda, \eta \rangle$ . As a consequence,  $dE/dt = 0$  and the nonholonomic equations of motion (12.3.25) conserve the energy in Equation (12.3.27). ■

**Exercise.** Use Noether's theorem to derive conservation of energy (12.3.27) for the nonholonomic Euler–Poincaré Equation (12.3.22). ★



**Remark 12.3.6 (More general formulations)** It is possible to formulate this nonholonomic reduction procedure more generally in the Euler–Poincaré context. For example, one may allow the constraints and Lagrangian to depend on two or more different parameters, say,  $(a_0, b_0) \in V \times V$ , and take an arbitrary linear combination for the constraints, as explained, for example, in [Sc2002]. This is useful for nonspherical bodies, whose constraints and potential energy may depend on the other two (nonvertical) directions, as seen from the body.  $\square$

**Exercise.** Show that this generalisation adds other terms and their corresponding auxiliary equations of the same form as those involving  $\Gamma$  to the right-hand side of the nonholonomic Euler–Poincaré Equations (12.3.25). ★

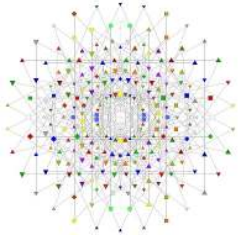
**Exercise.** What opportunities does this generalisation imply for the rocking and rolling of a nonspherical body? How about the case of a triaxial ellipsoid? ★

**Exercise.** Derive the Euler–Poincaré equations and discuss the properties of their solutions for the rolling motion of a circular disk whose centre of mass is displaced from its centre of symmetry.

This problem is a combination of the standard Euler disk, in which the centre of mass lies at the centre of symmetry, and the Chaplygin ball, in which it does not. ★

**286** 12 : ROUND, ROLLING RIGID BODIES

**Exercise.** Formulate the Clebsch method for dynamical systems with nonholonomic constraints. What role is played by the momentum map in this formulation? ★



# A

## GEOMETRICAL STRUCTURE OF CLASSICAL MECHANICS

---

### Contents

A.1	Manifolds	288
A.2	Motion: Tangent vectors and flows	296
A.2.1	Vector fields, integral curves and flows	297
A.2.2	Differentials of functions: The cotangent bundle	299
A.3	Tangent and cotangent lifts	300
A.3.1	Summary of derivatives on manifolds	301

---

## A.1 Manifolds

The main text refers to definitions and examples of smooth manifolds that are collected in this appendix for ready reference. We follow [AbMa1978, MaRa1994, Le2003, Ol2000, Wa1983], which should be consulted for further discussions as needed.

**Definition A.1.1** A smooth (i.e., differentiable) manifold  $M$  is a set of points together with a finite (or perhaps countable) set of subsets  $U_\alpha \subset M$  and one-to-one mappings  $\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^n$  such that

- $\bigcup_\alpha U_\alpha = M$ ;
- for every nonempty intersection  $U_\alpha \cap U_\beta$ , the set  $\phi_\alpha(U_\alpha \cap U_\beta)$  is an open subset of  $\mathbb{R}^n$  and the one-to-one mapping  $\phi_\beta \circ \phi_\alpha^{-1}$  is a smooth function on  $\phi_\alpha(U_\alpha \cap U_\beta)$ .

**Remark A.1.1** The sets  $U_\alpha$  in the definition are called *coordinate charts*. The mappings  $\phi_\alpha$  are called *coordinate functions* or *local coordinates*. A collection of charts satisfying both conditions is called an *atlas*.  $\square$

**Remark A.1.2** Refinements in the definition of a smooth manifold such as maximality conditions and equivalence classes of charts are ignored. (See [Wa1983, Le2003] for excellent discussions of these matters.)  $\square$

**Example A.1.1** Manifolds often arise as level sets

$$M = \{x \mid f_i(x) = 0, i = 1, \dots, k\},$$

for a given set of smooth functions  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, k$ .

If the gradients  $\nabla f_i$  are linearly independent, or more generally if the rank of  $\{\nabla f(x)\}$  is a constant  $r$  for all  $x$ , then  $M$  is a smooth manifold of dimension  $n - r$ . The proof uses the implicit function theorem to show that an  $(n - r)$ -dimensional coordinate chart may be defined in a neighbourhood of each point on  $M$ . In this situation, the set  $M$  is called an *implicit submanifold* of  $\mathbb{R}^n$  (see [Le2003, Ol2000]).

**Definition A.1.2 (Submersion)** If  $r = k = \dim M$ , then the map  $\{f_i\} : \mathbb{R}^n \rightarrow M$  is a **submersion**. (A submersion is a smooth map between smooth manifolds whose derivative is everywhere surjective.)

**Example A.1.2 (Stereographic projection of  $S^2 \rightarrow \mathbb{R}^2$ )** The unit sphere

$$S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$$

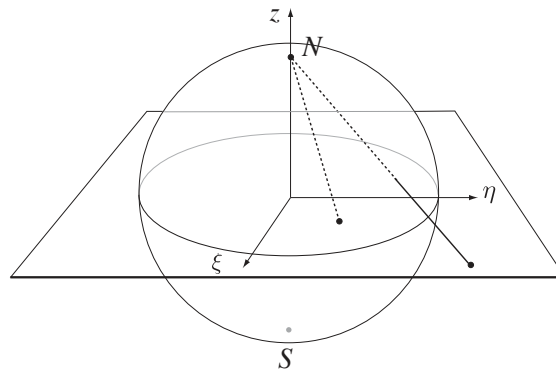
is a smooth two-dimensional manifold realised as a submersion in  $\mathbb{R}^3$ . Let

$$U_N = S^2 \setminus \{0, 0, 1\} \quad \text{and} \quad U_S = S^2 \setminus \{0, 0, -1\}$$

be the subsets obtained by deleting the north and south poles of  $S^2$ , respectively. Let

$$\chi_N : U_N \rightarrow (\xi_N, \eta_N) \in \mathbb{R}^2 \quad \text{and} \quad \chi_S : U_S \rightarrow (\xi_S, \eta_S) \in \mathbb{R}^2$$

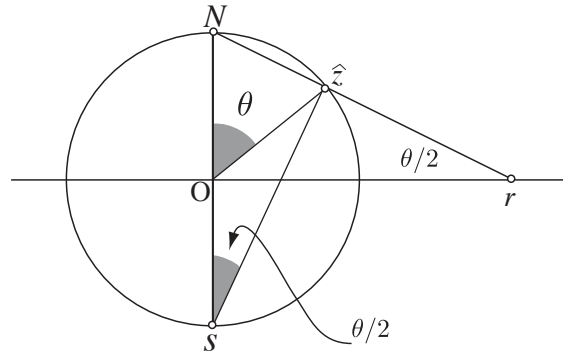
be stereographic projections from the north and south poles onto the equatorial plane,  $z = 0$ .



**Figure A.1.** Two Riemann projections from the north pole of the unit sphere onto the  $z = 0$  plane with coordinates  $(\xi_N, \eta_N)$ .

Thus, one may place two different coordinate patches in  $S^2$  intersecting everywhere except at the points along the  $z$ -axis at  $z = 1$  (north pole) and  $z = -1$  (south pole).

## 290 A : GEOMETRICAL STRUCTURE



**Figure A.2.** Riemann projection in the  $\xi_N - z$  plane at fixed azimuth  $\phi = 0$ . The projection through  $\hat{z} = (\sin \theta, \cos \theta)$  strikes the  $\xi_N$ -axis at distance  $r = \cot(\theta/2)$ .

In the equatorial plane  $z = 0$ , one may define two sets of (right-handed) coordinates,

$$\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^2 \setminus \{0\}, \quad \alpha = N, S,$$

obtained by the following two stereographic projections from the north and south poles:

- (valid everywhere except  $z = 1$ )

$$\phi_N(x, y, z) = (\xi_N, \eta_N) = \left( \frac{x}{1-z}, \frac{y}{1-z} \right),$$

- (valid everywhere except  $z = -1$ )

$$\phi_S(x, y, z) = (\xi_S, \eta_S) = \left( \frac{x}{1+z}, \frac{-y}{1+z} \right).$$

(The two complex planes are identified differently with the plane  $z = 0$ . An orientation-reversal is necessary to maintain consistent coordinates on the sphere.)

One may check directly that on the overlap  $U_N \cap U_S$ , the map

$$\phi_N \circ \phi_S^{-1} : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2 \setminus \{0\}$$

is a smooth diffeomorphism, given by the inversion

$$\phi_N \circ \phi_S^{-1}(x, y) = \left( \frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2} \right).$$

**Exercise.** Construct the mapping from  $(\xi_N, \eta_N) \rightarrow (\xi_S, \eta_S)$  and verify that it is a diffeomorphism in  $\mathbb{R}^2 \setminus \{0\}$ .

Hint:  $(1+z)(1-z) = 1 - z^2 = x^2 + y^2$ . ★

**Answer.**

$$(\xi_S, -\eta_S) = \frac{1-z}{1+z} (\xi_N, \eta_N) = \frac{1}{\xi_N^2 + \eta_N^2} (\xi_N, \eta_N).$$

The map  $(\xi_N, \eta_N) \rightarrow (\xi_S, \eta_S)$  is smooth and invertible except at  $(\xi_N, \eta_N) = (0, 0)$ . ▲

**Exercise.** Show that the circle of points on the sphere with polar angle (colatitude)  $\theta$  from the north pole project stereographically to another circle given in the complex plane by  $\zeta_N = \xi_N + i\eta_N = \cot(\theta/2)e^{i\phi}$ . ★

**Answer.** At a fixed azimuth  $\phi = 0$  (in the  $\xi_N, z$ -plane as in Figure A.2) a point on the sphere at colatitude  $\theta$  from the north pole has coordinates

$$\xi_N = \sin \theta, \quad z = \cos \theta.$$

## 292 A : GEOMETRICAL STRUCTURE

Its projection strikes the  $(\xi_N, \eta_N)$ -plane at distance  $r$  and angle  $\psi$ , given by

$$\cot \psi = r = \frac{r - \sin \theta}{\cos \theta}.$$

Thus  $\psi = \theta/2$ , since

$$\cot \psi = r = \frac{\sin \theta}{1 - \cos \theta} = \cot(\theta/2).$$

The stereographic projection of the sphere from its north pole at polar angle  $\theta$  thus describes a circle in the complex plane at

$$\zeta_N = \xi_N + i\eta_N = \frac{x + iy}{1 - z} = \cot(\theta/2)e^{i\phi}.$$

The corresponding stereographic projection of the sphere from its south pole describes the circle,

$$\zeta_S = \xi_S + i\eta_S = \frac{x - iy}{1 + z} = \tan(\theta/2)e^{-i\phi},$$

so that  $\zeta_S$  and  $\zeta_N$  are related by inversion,

$$\zeta_S = \frac{1}{\zeta_N} \quad \text{and} \quad \zeta_N = \frac{1}{\zeta_S}.$$



**Exercise.** Invert the previous relations to find the spherical polar coordinates  $(\theta, \phi)$  in terms of the planar variables  $\zeta_N = \xi_N + i\eta_N$ . ★

**Answer.** The stereographic formulas give

$$\zeta_N = \xi_N + i\eta_N = \frac{x + iy}{1 - z} = \cot(\theta/2)e^{i\phi},$$



with magnitude

$$|\zeta_N|^2 = \xi_N^2 + \eta_N^2 = \frac{1+z}{1-z} = \cot^2(\theta/2),$$

so that

$$1-z = \frac{2}{|\zeta_N|^2 + 1}.$$

Consequently,

$$x + iy = \frac{2\zeta_N}{|\zeta_N|^2 + 1} = \frac{2(\xi_N + i\eta_N)}{|\zeta_N|^2 + 1} = \sin(\theta)e^{i\phi},$$

and

$$z = \frac{|\zeta_N|^2 - 1}{|\zeta_N|^2 + 1} = \cos(\theta).$$

These are the usual invertible polar coordinate relations on the unit sphere. ▲

**Example A.1.3 (Torus  $T^2$ )** *If we start with two identical circles in the  $xz$ -plane, of radius  $r$  and centred at  $x = \pm 2r$ , then rotate them round the  $z$ -axis in  $\mathbb{R}^3$ , the result is a torus, written  $T^2$ . It is a manifold.*

**Exercise.** If we begin with a figure eight in the  $xz$ -plane, along the  $x$ -axis and centred at the origin, and spin it round the  $z$ -axis in  $\mathbb{R}^3$ , we get a pinched surface that looks like a sphere that has been pinched so that the north and south poles touch. Is this a manifold? Prove it. ★

**Answer.** The origin has a neighbourhood diffeomorphic to a double cone. This is not diffeomorphic to  $\mathbb{R}^2$ . A proof of this is that if the origin of the cone is removed, two components remain; while if the origin of  $\mathbb{R}^2$  is removed, only one component remains. ▲

**Exercise.** A sphere with unit radius is projected onto a plane, then projected back to a second sphere of different radius. Both projections are stereographic. Show that the polar angle (colatitude)  $\theta$  of the unit sphere is re-projected to the colatitude  $\theta'$  of the second sphere by

$$\tan(\theta'/2) = R^{-1} \tan(\theta/2),$$

where  $R$  is the radius of the second sphere. ★

**Definition A.1.3 (Immersed submanifolds)** An *immersed submanifold* of a manifold  $M$  is a subset  $S$  together with a topology and differential structure such that  $S$  is a manifold and the inclusion map  $i : S \hookrightarrow M$  is an *injective immersion*; that is, a smooth map between smooth manifolds whose derivative is everywhere injective. Hence, immersed submanifolds are the images of injective immersions.

**Remark A.1.3** Given any injective immersion  $f : N \rightarrow M$  the image of  $N$  in  $M$  uniquely defines an immersed submanifold so that  $f : N \rightarrow f(N)$  is a *diffeomorphism*; that is, a smooth invertible map with a smooth inverse. □

**Definition A.1.4 (Lie group and Lie subgroup manifolds)** A *Lie group* is a group that is also a manifold. A *Lie subgroup* is a submanifold that is invariant under group operations. That is, Lie subgroups are injective immersions.

**Definition A.1.5 (Tangent space to level sets)** Let the submersion defined by

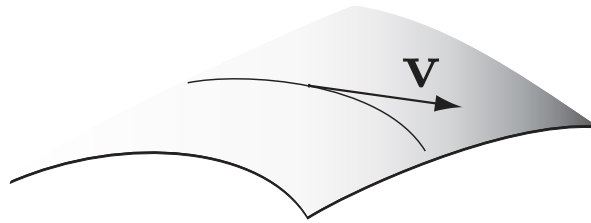
$$M = \{x \mid f_i(x) = 0, i = 1, \dots, k\}$$

be a submanifold of  $\mathbb{R}^n$ . The *tangent space* at each  $x \in M$  is defined by

$$T_x M = \left\{ v \in \mathbb{R}^n \mid \frac{\partial f_i}{\partial x^a}(x) v^a = 0, i = 1, \dots, k \right\}.$$

*Note:* We use the **summation convention**, i.e., repeated indices are summed over their range.

**Definition A.1.6 (Tangent vector)** The tangent space  $T_x M$  at a point  $x$  of a manifold  $M$  is a **vector space**. The elements of this space are called **tangent vectors** (Figure A.3).



**Figure A.3.** A tangent vector at a point on a manifold.

**Example A.1.4 (Tangent space to the sphere in  $\mathbb{R}^3$ )** The sphere  $S^2$  is the set of points  $(x, y, z) \in \mathbb{R}^3$  solving  $x^2 + y^2 + z^2 = 1$ . The tangent space to the sphere at such a point  $(x, y, z)$  is the plane containing vectors  $(u, v, w)$  satisfying  $xu + yv + zw = 0$ .

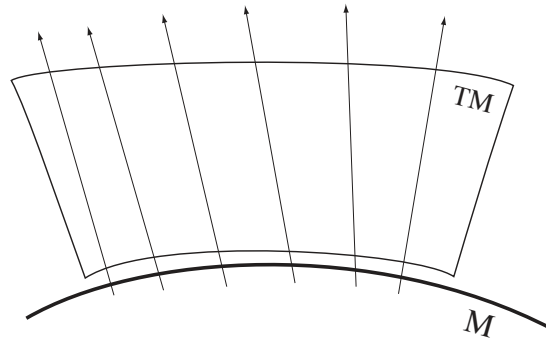
**Definition A.1.7 (Tangent bundle)** The **tangent bundle** of a manifold  $M$ , denoted by  $TM$ , is the smooth manifold whose underlying set is the disjoint union of the tangent spaces to  $M$  at the points  $x \in M$  (Figure A.4); that is,

$$TM = \bigcup_{x \in M} T_x M.$$

Thus, a single point of  $TM$  is  $(x, v)$  where  $x \in M$  and  $v \in T_x M$ .

**Remark A.1.4 (Dimension of tangent bundle  $TS^2$ )** Defining  $TS^2$  requires two independent conditions each of codimension 5 in  $\mathbb{R}^6$ ; so  $\dim TS^2 = 5 + 5 - 6 = 4$ .  $\square$

## 296 A : GEOMETRICAL STRUCTURE



**Figure A.4.** The tangent bundle  $TM$  of a manifold  $M$  is the disjoint union of tangent spaces over all points in  $M$ .

**Exercise.** Define the sphere  $S^{n-1}$  in  $\mathbb{R}^n$ . What is the dimension of its tangent space  $TS^{n-1}$ ? ★

## A.2 Motion: Tangent vectors and flows

Envisioning our later considerations of dynamical systems, we shall consider motion along curves  $c(t)$  parameterised by time  $t$  on a smooth manifold  $M$ . Suppose these curves are trajectories of a flow  $\phi_t$  along the tangent vectors of the manifold. We anticipate this means  $\phi_t(c(0)) = c(t)$  and  $\phi_t \circ \phi_s = \phi_{t+s}$  (flow property). The flow will be tangent to  $M$  along the curve. To deal with such flows, we will need to know more about tangent vectors.

Recall from Definition A.1.7 that the tangent bundle of  $M$  is

$$TM = \bigcup_{x \in M} T_x M.$$

We will now add a bit more to that definition. The tangent bundle is an example of a more general structure than a manifold.

**Definition A.2.1 (Bundle)** A *bundle* consists of a manifold  $E$ , another manifold  $B$  called the base space and a projection between them,  $\Pi : E \rightarrow B$ . The inverse images of the projection  $\Pi$  exist and are called the *fibres* of the bundle. Thus, subsets of the bundle  $E$  locally have the structure of a Cartesian product space  $B \times F$ , of the *base space*  $B$  with the *fibre space*  $F$ . An example is  $(E, B, \Pi)$  consisting of  $(\mathbb{R}^2, \mathbb{R}^1, \Pi : \mathbb{R}^2 \rightarrow \mathbb{R}^1)$ . In this case,  $\Pi : (x, y) \in \mathbb{R}^2 \rightarrow x \in \mathbb{R}^1$ . Likewise, the tangent bundle consists of  $M, TM$  and a map  $\tau_M : TM \rightarrow M$ .

Let  $x = (x^1, \dots, x^n)$  be local coordinates on  $M$ , and let  $v = (v^1, \dots, v^n)$  be components of a tangent vector.

$$T_x M = \left\{ v \in \mathbb{R}^n \mid \frac{\partial f_i}{\partial x} \cdot v = 0, i = 1, \dots, m \right\},$$

for

$$M = \{x \in \mathbb{R}^n \mid f_i(x) = 0, i = 1, \dots, m\}.$$

These  $2n$  numbers  $(x, v)$  give local coordinates on  $TM$ , whose dimension is  $\dim TM = 2 \dim M$ . The *tangent bundle projection* is a map  $\tau_M : TM \rightarrow M$  which takes a tangent vector  $v$  to a point  $x \in M$  where the tangent vector  $v$  is attached (that is,  $v \in T_x M$ ). The inverse of this projection  $\tau_M^{-1}(x)$  is called the fibre over  $x$  in the tangent bundle.

### A.2.1 Vector fields, integral curves and flows

**Definition A.2.2** A *vector field* on a manifold  $M$  is a map  $X : M \rightarrow TM$  that assigns a vector  $X(x)$  at each point  $x \in M$ . This implies that  $\tau_M \circ X = Id$ .

**Definition A.2.3** An *integral curve* of  $X$  with initial conditions  $x_0$  at  $t = 0$  is a differentiable map  $c : ]a, b[ \rightarrow M$ , where  $]a, b[$  is an open interval containing 0, such that  $c(0) = x_0$  and  $c'(t) = X(c(t))$  for all  $t \in ]a, b[$ .

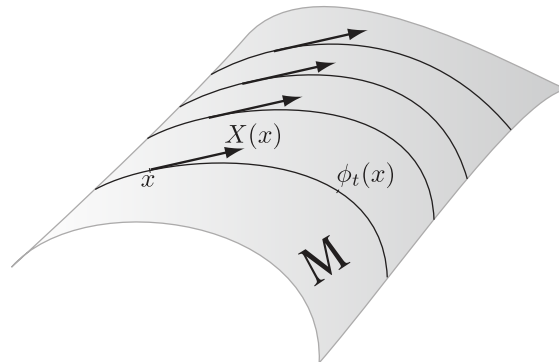
## 298 A : GEOMETRICAL STRUCTURE

**Remark A.2.1** A standard result from the theory of ordinary differential equations states that  $X$  being Lipschitz is sufficient for its integral curves to be unique and  $C^1$  [CoLe1984]. The integral curves  $c(t)$  are differentiable for smooth  $X$ .  $\square$

**Definition A.2.4** The *flow* of  $X$  is the collection of maps

$$\phi_t : M \rightarrow M,$$

where  $t \rightarrow \phi_t(x)$  is the integral curve of  $X$  with initial condition  $x$  (Figure A.5).



**Figure A.5.** A collection of maps  $\phi_t : M \rightarrow M$  shows the flow of a vector field  $X$  acting on a manifold  $M$ .

**Remark A.2.2**

- Existence and uniqueness results for solutions of  $c'(t) = X(c(t))$  guarantee that flow  $\phi$  of  $X$  is smooth in  $(x, t)$ , for smooth  $X$ .
- Uniqueness implies the flow property

$$\phi_{t+s} = \phi_t \circ \phi_s, \quad (\text{FP})$$

for initial condition  $\phi_0 = Id$ .

- The flow property (FP) generalises to the nonlinear case the familiar linear situation where  $M$  is a vector space,  $X(x) = Ax$  is a linear vector field for a bounded linear operator  $A$ , and  $\phi_t(x) = e^{At}x$  (exponentiation).

□

**Definition A.2.5 (Equivariance)** *A map between two spaces is **equivariant** when it respects group actions on these spaces.*

### A.2.2 Differentials of functions: The cotangent bundle

We are now ready to define differentials of smooth functions and the cotangent bundle.

Let  $f : M \rightarrow \mathbb{R}$  be a smooth function. We differentiate  $f$  at  $x \in M$  to obtain  $T_x f : T_x M \rightarrow T_{f(x)}\mathbb{R}$ . As is standard, we identify  $T_{f(x)}\mathbb{R}$  with  $\mathbb{R}$  itself, thereby obtaining a linear map  $df(x) : T_x M \rightarrow \mathbb{R}$ . The result  $df(x)$  is an element of the cotangent space  $T_x^* M$ , the dual space of the tangent space  $T_x M$ . The natural pairing between elements of the tangent space and the cotangent space is denoted as  $\langle \cdot, \cdot \rangle : T_x^* M \times T_x M \mapsto \mathbb{R}$ .

In coordinates, the linear map  $df(x) : T_x M \rightarrow \mathbb{R}$  may be written as the directional derivative,

$$\langle df(x), v \rangle = df(x) \cdot v = \frac{\partial f}{\partial x^i} \cdot v^i,$$

for all  $v \in T_x M$ .

(Reminder: The summation convention is applied over repeated indices.) Hence, elements  $df(x) \in T_x^* M$  are dual to vectors  $v \in T_x M$  with respect to the pairing  $\langle \cdot, \cdot \rangle$ .

**Definition A.2.6**  *$df$  is the **differential** of the function  $f$ .*

**Definition A.2.7** *The dual space of the tangent bundle  $TM$  is the **cotangent bundle**  $T^*M$ . That is,*

$$(T_x M)^* = T_x^* M \quad \text{and} \quad T^* M = \bigcup_x T_x^* M.$$

## 300 A : GEOMETRICAL STRUCTURE

Thus, replacing  $v \in T_x M$  with  $df \in T_x^* M$ , for all  $x \in M$  and for all smooth functions  $f : M \rightarrow \mathbb{R}$ , yields the cotangent bundle  $T^* M$ .

**Differential bases**

When the basis of vector fields is denoted as  $\frac{\partial}{\partial x^i}$  for  $i = 1, \dots, n$ , its dual basis is often denoted as  $dx^i$ . In this notation, the differential of a function at a point  $x \in M$  is expressed as

$$df(x) = \frac{\partial f}{\partial x^i} dx^i.$$

The corresponding pairing  $\langle \cdot, \cdot \rangle$  of bases is written in this notation as

$$\left\langle dx^j, \frac{\partial}{\partial x^i} \right\rangle = \delta_i^j.$$

Here  $\delta_i^j$  is the Kronecker delta, which equals unity for  $i = j$  and vanishes otherwise. That is, defining  $T^* M$  requires a pairing  $\langle \cdot, \cdot \rangle : T^* M \times TM \rightarrow \mathbb{R}$ .

**A.3 Tangent and cotangent lifts**

We next define derivatives of differentiable maps between manifolds (tangent lifts).

We expect that a smooth map  $f : U \rightarrow V$  from a chart  $U \subset M$  to a chart  $V \subset N$  will lift to a map between the tangent bundles  $TM$  and  $TN$  so as to make sense from the viewpoint of ordinary calculus,

$$\begin{aligned} U \times \mathbb{R}^m \subset TM &\longrightarrow V \times \mathbb{R}^n \subset TN \\ (q^1, \dots, q^m; X^1, \dots, X^m) &\longmapsto (Q^1, \dots, Q^n; Y^1, \dots, Y^n). \end{aligned}$$

Namely, the relations between the vector field components should be obtained from the differential of the map  $f : U \rightarrow V$ . Perhaps not unexpectedly, these vector field components will be related by

$$Y^i \frac{\partial}{\partial Q^i} = X^j \frac{\partial}{\partial q^j}, \quad \text{so} \quad Y^i = \frac{\partial Q^i}{\partial q^j} X^j,$$



in which the quantity called the *tangent lift*

$$Tf = \frac{\partial Q}{\partial q} \quad (\text{A.3.1})$$

of the function  $f$  arises from the chain rule and is equal to the Jacobian for the transformation  $Tf : TM \mapsto TN$ .

The dual of the tangent lift is the cotangent lift, which plays a major role in the definition of momentum maps in Chapter 11. Roughly speaking, the *cotangent lift* of the function  $f$ ,

$$T^*f = \frac{\partial q}{\partial Q}, \quad (\text{A.3.2})$$

arises from

$$\beta_i dQ^i = \alpha_j dq^j, \quad \text{so} \quad \beta_i = \alpha_j \frac{\partial q^j}{\partial Q^i},$$

and  $T^*f : T^*N \mapsto T^*M$ . Note the directions of these maps.

$$\begin{aligned} Tf &: q, X \in TM \mapsto Q, Y \in TN, \\ f &: q \in M \mapsto Q \in N, \\ T^*f &: Q, \beta \in T^*N \mapsto q, \alpha \in T^*M. \end{aligned}$$

### A.3.1 Summary of derivatives on manifolds

**Definition A.3.1 (Differentiable map)** A map  $f : M \rightarrow N$  from manifold  $M$  to manifold  $N$  is said to be **differentiable** (resp.  $C^k$ ) if it is represented in local coordinates on  $M$  and  $N$  by differentiable (resp.  $C^k$ ) functions.

**Definition A.3.2 (Derivative of a differentiable map)** The *derivative* of a differentiable map

$$f : M \rightarrow N$$

at a point  $x \in M$  is defined to be the linear map

$$T_x f : T_x M \rightarrow T_x N,$$

## 302 A : GEOMETRICAL STRUCTURE

constructed as follows. For  $v \in T_x M$ , choose a curve that maps an open interval  $(-\epsilon, \epsilon)$  around the point  $t = 0$  to the manifold  $M$ ,

$$\begin{aligned} c &: (-\epsilon, \epsilon) \longrightarrow M, \\ \text{with } c(0) &= x, \\ \text{and velocity vector } \frac{dc}{dt} \Big|_{t=0} &= v. \end{aligned}$$

Then  $T_x f \cdot v$  is the velocity vector at  $t = 0$  of the curve  $f \circ c : \mathbb{R} \rightarrow N$ . That is,

$$T_x f \cdot v = \frac{d}{dt} f(c(t)) \Big|_{t=0}. \quad (\text{A.3.3})$$

**Definition A.3.3** The union  $Tf = \bigcup_x T_x f$  of the derivatives  $T_x f : T_x M \rightarrow T_x N$  over points  $x \in M$  is called the **tangent lift** of the map  $f : M \rightarrow N$ .

**Remark A.3.1** The chain-rule definition of the derivative  $T_x f$  of a differentiable map at a point  $x$  depends on the function  $f$  and the vector  $v$ . Other degrees of differentiability are possible. For example, if  $M$  and  $N$  are manifolds and  $f : M \rightarrow N$  is of class  $C^{k+1}$ , then the tangent lift (Jacobian)  $T_x f : T_x M \rightarrow T_x N$  is  $C^k$ .  $\square$

**Exercise.** Let  $\phi_t : S^2 \rightarrow S^2$  rotate points on  $S^2$  about a fixed axis through an angle  $\psi(t)$ . Show that  $\phi_t$  is the flow of a certain vector field on  $S^2$ .  $\star$

**Exercise.** Let  $f : S^2 \rightarrow \mathbb{R}$  be defined by  $f(x, y, z) = z$ . Compute  $df$  using spherical coordinates  $(\theta, \phi)$ .  $\star$

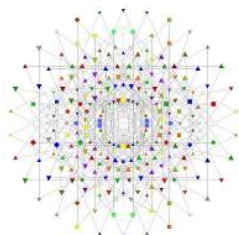
**Exercise.** Compute the tangent lifts for the two stereographic projections of  $S^2 \rightarrow \mathbb{R}^2$  in Example A.1.2. That is, assuming  $(x, y, z)$  depend smoothly on  $t$ , find

- How  $(\dot{\xi}_N, \dot{\eta}_N)$  depend on  $(\dot{x}, \dot{y}, \dot{z})$ . Likewise, for  $(\dot{\xi}_S, \dot{\eta}_S)$ .
- How  $(\dot{\xi}_N, \dot{\eta}_N)$  depend on  $(\dot{\xi}_S, \dot{\eta}_S)$ .

Hint: Recall  $(1+z)(1-z) = 1-z^2 = x^2 + y^2$  and use  $x\dot{x} + y\dot{y} + z\dot{z} = 0$  when  $(\dot{x}, \dot{y}, \dot{z})$  is tangent to  $S^2$  at  $(x, y, z)$ .







# B

## LIE GROUPS AND LIE ALGEBRAS

---

### Contents

B.1	Matrix Lie groups	306
B.2	Defining matrix Lie algebras	310
B.3	Examples of matrix Lie groups	312
B.4	Lie group actions	314
B.4.1	Left and right translations on a Lie group	316
B.5	Tangent and cotangent lift actions	317
B.6	Jacobi–Lie bracket	320
B.7	Lie derivative and Jacobi–Lie bracket	323
B.7.1	Lie derivative of a vector field	323
B.7.2	Vector fields in ideal fluid dynamics	325

---

## B.1 Matrix Lie groups

The main text refers to definitions and examples of Lie group properties that are collected in this appendix for ready reference. We follow [AbMa1978, MaRa1994, Ol2000], which should be consulted for further discussions as needed by the reader. See also [Ho2008] for another introductory discussion along similar lines.

**Definition B.1.1 (Group)** A *group*  $G$  is a set of elements possessing

- A binary product (multiplication),  $G \times G \rightarrow G$ , such that the following properties hold:
  - The product of  $g$  and  $h$  is written  $gh$ .
  - The product is associative,  $(gh)k = g(hk)$ .
- Identity element  $e$ :  $eg = g$  and  $ge = g$ , for all  $g \in G$ .
- Inverse operation  $G \rightarrow G$ , so that  $gg^{-1} = g^{-1}g = e$ .

**Definition B.1.2 (Lie group)** A *Lie group* is a smooth manifold  $G$  which is also a group and for which the group operations of multiplication,  $(g, h) \rightarrow gh$  for  $g, h \in G$ , and inversion,  $g \rightarrow g^{-1}$  with  $gg^{-1} = g^{-1}g = e$ , are smooth functions.

**Definition B.1.3** A *matrix Lie group* is a set of invertible  $n \times n$  matrices which is closed under matrix multiplication and which is a submanifold of  $\mathbb{R}^{n \times n}$ .

**Remark B.1.1** The conditions showing that a matrix Lie group is a Lie group are easily checked:

- A matrix Lie group is a manifold, because it is a submanifold of  $\mathbb{R}^{n \times n}$ .
- Its group operations are smooth, since they are algebraic operations on the matrix entries.

□

**Example B.1.1 (The general linear group  $GL(n, \mathbb{R})$ )** *The matrix Lie group  $GL(n, \mathbb{R})$  is the group of linear isomorphisms of  $\mathbb{R}^n$  to itself. The dimension of the  $n \times n$  matrices in  $GL(n, \mathbb{R})$  is  $n^2$ , the number of independent elements.*

**Remark B.1.2** Definition A.1.4 of Lie groups in terms of manifolds also defines Lie subgroups in terms of invariant submanifolds. This definition suggests a strategy for characterising Lie subgroups of  $GL(n, \mathbb{R})$  by imposing submanifold invariance conditions.  $\square$

**Proposition B.1.1 (Isotropy subgroup)** *Let  $K \in GL(n, \mathbb{R})$  be a symmetric matrix,  $K^T = K$ . The mapping*

$$S = \{U \in GL(n, \mathbb{R}) \mid U^T K U = K\}$$

*defines the subgroup  $S$  of  $GL(n, \mathbb{R})$ . Moreover, this subgroup is a submanifold of  $\mathbb{R}^{n \times n}$  of dimension  $n(n-1)/2$ .*

**Proof.** Is the mapping  $S$  in Proposition B.1.1 a subgroup? We check the following three defining properties:

- Identity:  $I \in S$  because  $I^T K I = K$ .
- Inverse:  $U \in S \implies U^{-1} \in S$  because

$$K = U^{-T}(U^T K U)U^{-1} = U^{-T}(K)U^{-1}.$$

- Closed under multiplication:  $U, V \in S \implies UV \in S$  because

$$(UV)^T K UV = V^T (U^T K U) V = V^T (K) V = K.$$

Hence,  $S$  is a subgroup of  $GL(n, \mathbb{R})$ .

Is  $S$  a submanifold of  $\mathbb{R}^{n \times n}$  of dimension  $n(n-1)/2$ ? Indeed,  $S$  is the zero locus of the mapping  $UKU^T - K$ . Its non-redundant entries, i.e., those on or above the diagonal, define a submersion and thus make  $S$  a submanifold.

For a submersion, the dimension of the level set is the dimension of the domain minus the dimension of the range space. In this case, this dimension is  $n^2 - n(n+1)/2 = n(n-1)/2$ .  $\blacksquare$

## 308 B : LIE GROUPS AND LIE ALGEBRAS

**Exercise.** Explain why one can conclude that the zero locus map for  $S$  is a submersion. In particular, pay close attention to establishing the constant rank condition for the linearisation of this map. ★

**Answer.** Here is the strategy for proving that  $S$  is a submanifold of  $R^{n \times n}$ .

(i) The mapping  $S$  is the zero locus of

$$U \rightarrow U^T K U - K \quad (\text{locus map}).$$

(ii) Let  $U \in S$ , and let  $\delta U$  be an arbitrary element of  $R^{n \times n}$ . Linearise the locus map to find

$$\begin{aligned} (U + \delta U)^T K (U + \delta U) - K & \quad (\text{B.1.1}) \\ = U^T K U - K + \delta U^T K U + U^T K \delta U + O(\delta U)^2. \end{aligned}$$

(iii) We may conclude that  $S$  is a submanifold of  $R^{n \times n}$  provided we can show that the linearisation of the locus map, namely the linear mapping defined by

$$L \equiv \delta U \rightarrow \delta U^T K U + U^T K \delta U, \quad R^{n \times n} \mapsto R^{n \times n}, \quad (\text{B.1.2})$$

has constant rank for all  $U \in S$ .

▲

To apply this strategy, one needs a lemma.

**Lemma B.1.1** *The linearisation map  $L$  in (B.1.2) is onto the space of  $n \times n$  symmetric matrices. Hence the original map is a submersion.*

**Proof.** Both the original locus map and the image of  $L$  lie in the subspace of  $n \times n$  symmetric matrices. Indeed, given  $U$  and any



symmetric matrix  $A$  we can find  $\delta U$  such that

$$\delta U^T K U + U^T K \delta U = A.$$

Namely

$$\delta U = K^{-1} U^{-T} A / 2.$$

Thus, the linearisation map  $L$  is onto the space of  $n \times n$  symmetric matrices and the original locus map  $U \rightarrow U K U^T - K$  to the space of symmetric matrices is a submersion. ■

**Remark B.1.3** The subgroup  $S$  leaves invariant a certain symmetric quadratic form under linear transformations,  $S \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  given by  $\mathbf{x} \rightarrow U\mathbf{x}$ , since

$$\mathbf{x}^T K \mathbf{x} = \mathbf{x}^T U^T K U \mathbf{x}.$$

So the matrices  $U \in S$  change the basis for this quadratic form, but they leave its value unchanged. This means  $S$  is the *isotropy subgroup* of the quadratic form associated with  $K$ . □

**Corollary B.1.1** ( *$S$  is a matrix Lie group*)  $S$  is both a subgroup and a submanifold of the general linear group  $GL(n, \mathbb{R})$ . Thus, by Definition B.1.3, the subgroup  $S$  is a matrix Lie group.

**Exercise.** What is the tangent space to  $S$  at the identity,  $T_I S$ ? ★

**Proposition B.1.2** *The linear space of matrices  $A$  satisfying*

$$A^T K + K A = 0$$

*defines  $T_I S$ , the tangent space at the identity of the matrix Lie group  $S$  defined in Proposition B.1.1.*

**Proof.** Near the identity the defining condition for  $S$  expands to

$$(I + \epsilon A^T + O(\epsilon^2)) K (I + \epsilon A + O(\epsilon^2)) = K, \quad \text{for } \epsilon \ll 1.$$

### 310 B : LIE GROUPS AND LIE ALGEBRAS

At linear order  $O(\epsilon)$  one finds

$$A^T K + K A = 0.$$

This relation defines the linear space of matrices  $A \in T_I S$ . ■

**Exercise.** If  $A, B \in T_I S$ , does it follow that  $[A, B] \in T_I S$ ?  
This is closure. ★

**Proposition B.1.3 (Closure)** For any pair of matrices  $A, B \in T_I S$ , the matrix commutator  $[A, B] \equiv AB - BA \in T_I S$ .

**Proof.** Using  $[A, B]^T = [B^T, A^T]$ , we may check closure by a direct computation,

$$\begin{aligned} [B^T, A^T]K + K[A, B] &= B^T A^T K - A^T B^T K + KAB - KBA \\ &= B^T A^T K - A^T B^T K - A^T K B + B^T K A = 0. \end{aligned}$$

Hence, the tangent space of  $S$  at the identity  $T_I S$  is closed under the matrix commutator  $[\cdot, \cdot]$ . ■

**Remark B.1.4** In a moment, we will show that the matrix commutator for  $T_I S$  also satisfies the Jacobi identity. This will imply that the condition  $A^T K + K A = 0$  defines a matrix Lie algebra. □

## B.2 Defining matrix Lie algebras

We are ready to prove the following, in preparation for defining matrix Lie algebras.

**Proposition B.2.1** Let  $S$  be a matrix Lie group, and let  $A, B \in T_I S$  (the tangent space to  $S$  at the identity element). Then  $AB - BA \in T_I S$ .

The proof makes use of a lemma.

**Lemma B.2.1** *Let  $R$  be an arbitrary element of a matrix Lie group  $S$ , and let  $B \in T_I S$ . Then  $RBR^{-1} \in T_I S$ .*

**Proof.** Let  $R_B(t)$  be a curve in  $S$  such that  $R_B(0) = I$  and  $R'_B(0) = B$ . Define  $S(t) = RR_B(t)R^{-1} \in S$  for all  $t$ . Then  $S(0) = I$  and  $S'(0) = RBR^{-1}$ . Hence,  $S'(0) \in T_I S$ , thereby proving the lemma. ■

**Proof of Proposition B.2.1.** Let  $R_A(s)$  be a curve in  $S$  such that  $R_A(0) = I$  and  $R'_A(0) = A$ . Define  $S(t) = R_A(t)BR_A(t)^{-1} \in T_I S$ . Then the lemma implies that  $S(t) \in T_I S$  for every  $t$ . Hence,  $S'(t) \in T_I S$ , and in particular,  $S'(0) = AB - BA \in T_I S$ . ■

**Definition B.2.1 (Matrix commutator)** *For any pair of  $n \times n$  matrices  $A, B$ , the **matrix commutator** is defined as*

$$[A, B] = AB - BA.$$

**Proposition B.2.2 (Properties of the matrix commutator)** *The matrix commutator has the following two properties:*

- Any two  $n \times n$  matrices  $A$  and  $B$  satisfy

$$[B, A] = -[A, B].$$

*(This is the property of skew-symmetry.)*

- Any three  $n \times n$  matrices  $A, B$  and  $C$  satisfy

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0.$$

*(This is known as the **Jacobi identity**.)*

**Definition B.2.2 (Matrix Lie algebra)** *A matrix Lie algebra  $\mathfrak{g}$  is a set of  $n \times n$  matrices which is a vector space with respect to the usual operations of matrix addition and multiplication by real numbers (scalars) and which is closed under the matrix commutator  $[\cdot, \cdot]$ .*

**Proposition B.2.3** *For any matrix Lie group  $S$ , the tangent space at the identity  $T_I S$  is a matrix Lie algebra.*

## 312 B : LIE GROUPS AND LIE ALGEBRAS

**Proof.** This follows by Proposition B.2.1 and because  $T_I S$  is a vector space. ■

### B.3 Examples of matrix Lie groups

**Example B.3.1 (The orthogonal group  $O(n)$ )** *The mapping condition  $U^T K U = K$  in Proposition B.1.1 specialises for  $K = I$  to  $U^T U = I$ , which defines the orthogonal group. Thus, in this case,  $S$  specialises to  $O(n)$ , the group of  $n \times n$  orthogonal matrices. The orthogonal group is of special interest in mechanics.*

**Corollary B.3.1 ( $O(n)$  is a matrix Lie group)** *By Proposition B.1.1 the orthogonal group  $O(n)$  is both a subgroup and a submanifold of the general linear group  $GL(n, \mathbb{R})$ . Thus, by Definition B.1.3, the orthogonal group  $O(n)$  is a matrix Lie group.*

**Example B.3.2 (The special linear group  $SL(n, \mathbb{R})$ )** *The subgroup of  $GL(n, \mathbb{R})$  with  $\det(U) = 1$  is called  $SL(n, \mathbb{R})$ .*

**Example B.3.3 (The special orthogonal group  $SO(n)$ )** *The special case of  $S$  with  $\det(U) = 1$  and  $K = I$  is called  $SO(n)$ . In this case, the mapping condition  $U^T K U = K$  specialises to  $U^T U = I$  with the extra condition  $\det(U) = 1$ .*

**Example B.3.4 (Tangent space of  $SO(n)$  at the identity)** *The special case with  $K = I$  of  $T_I SO(n)$  yields*

$$A^T + A = 0.$$

*These are antisymmetric matrices. Lying in the tangent space at the identity of a matrix Lie group, this linear vector space forms the matrix Lie algebra  $so(n)$ .*

**Example B.3.5 (The special unitary group  $SU(n)$ )** *The Lie group  $SU(n)$  comprises complex  $n \times n$  unitary matrices  $U$  with  $U^\dagger U = I$  and*

## B.3 EXAMPLES OF MATRIX LIE GROUPS 313

unit determinant  $\det U = 1$ . An element  $A$  in its tangent space at the identity satisfies

$$A + A^\dagger = 0 \quad \text{for } A \in T_I SU(n),$$

so that  $A \in \mathfrak{su}(n)$  is an  $n \times n$  traceless skew-Hermitian matrix.

**Example B.3.6 (The symplectic group  $Sp(l)$ )** Suppose  $n = 2l$  (that is, let  $n$  be even) and consider the nonsingular skew-symmetric matrix

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

where  $I$  is the  $l \times l$  identity matrix. One may verify that

$$Sp(l) = \{U \in GL(2l, \mathbb{R}) \mid U^T J U = J\}$$

is a group. This is called the symplectic group. Reasoning as before, the matrix algebra  $T_I Sp(l)$  is defined as the set of  $n \times n$  matrices  $A$  satisfying

$$J A^T + A J = 0.$$

This matrix Lie algebra is denoted as  $\mathfrak{sp}(l)$ .

**Example B.3.7 (The special Euclidean group  $SE(3)$ )** Consider the Lie group of  $4 \times 4$  matrices of the form

$$E(R, v) = \begin{bmatrix} R & v \\ 0 & 1 \end{bmatrix},$$

where  $R \in SO(3)$  and  $v \in \mathbb{R}^3$ . This is the special Euclidean group, denoted  $SE(3)$ . The special Euclidean group is of central interest in mechanics since it describes the set of rigid motions and coordinate transformations of three-dimensional space.

**Exercise.** A point  $P$  in  $\mathbb{R}^3$  undergoes a rigid motion associated with  $E(R_1, v_1)$  followed by a rigid motion associated with  $E(R_2, v_2)$ . What matrix element of  $SE(3)$  is associated with the composition of these motions in the given order?



## 314 B : LIE GROUPS AND LIE ALGEBRAS

**Exercise.** Multiply the special Euclidean matrices of  $SE(3)$ . Investigate their matrix commutators in their tangent space at the identity. (This is an example of a semidirect-product Lie group.) ★

**Exercise. (Tripos question)** When does a stone at the equator of the Earth weigh the most?

Two hints: (i) Assume the Earth's orbit is a circle around the Sun and ignore the declination of the Earth's axis of rotation. (ii) This is an exercise in using  $SE(2)$ . ★

**Exercise.** Suppose the  $n \times n$  matrices  $A$  and  $M$  satisfy

$$AM + MA^T = 0.$$

Show that  $\exp(At)M \exp(A^T t) = M$  for all  $t$ .

Hint:  $A^n M = M(-A^T)^n$ . This direct calculation shows that for  $A \in so(n)$  or  $A \in sp(l)$ , we have  $\exp(At) \in SO(n)$  or  $\exp(At) \in Sp(l)$ , respectively. ★

## B.4 Lie group actions

The action of a Lie group  $G$  on a manifold  $M$  is a group of transformations of  $M$  associated with elements of the group  $G$ , whose composition acting on  $M$  corresponds to group multiplication in  $G$ .

**Definition B.4.1 (Left action of a Lie group)** Let  $M$  be a manifold and let  $G$  be a Lie group. A **left action** of a Lie group  $G$  on  $M$  is a smooth mapping  $\Phi : G \times M \rightarrow M$  such that

- (i)  $\Phi(e, x) = x$  for all  $x \in M$ ;
- (ii)  $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$  for all  $g, h \in G$  and  $x \in M$ ; and
- (iii)  $\Phi(g, \cdot)$  is a diffeomorphism on  $M$  for each  $g \in G$ .

We often use the convenient notation  $gx$  for  $\Phi(g, x)$  and think of the group element  $g$  acting on the point  $x \in M$ . The associativity condition (ii) then simply reads  $(gh)x = g(hx)$ .

Similarly, one can define a **right action**, which is a map  $\Psi : M \times G \rightarrow M$  satisfying  $\Psi(x, e) = x$  and  $\Psi(\Psi(x, g), h) = \Psi(x, gh)$ . The convenient notation for right action is  $xg$  for  $\Psi(x, g)$ , the right action of a group element  $g$  on the point  $x \in M$ . Associativity  $\Psi(\Psi(x, g), h) = \Psi(x, gh)$  may then be expressed conveniently as  $(xg)h = x(gh)$ .

**Example B.4.1 (Properties of Lie group actions)** The action  $\Phi : G \times M \rightarrow M$  of a group  $G$  on a manifold  $M$  is said to be

- **transitive**, if for every  $x, y \in M$  there exists a  $g \in G$ , such that  $gx = y$ ;
- **free**, if it has no fixed points, that is,  $\Phi_g(x) = x$  implies  $g = e$ ; and
- **proper**, if whenever a convergent subsequence  $\{x_n\}$  in  $M$  exists, and the mapping  $g_n x_n$  converges in  $M$ , then  $\{g_n\}$  has a convergent subsequence in  $G$ .

**Definition B.4.2 (Group orbits)** Given a Lie group action of  $G$  on  $M$ , for a given point  $x \in M$ , the subset

$$\text{Orb } x = \{gx \mid g \in G\} \subset M$$

is called the **group orbit** through  $x$ .

## 316 B : LIE GROUPS AND LIE ALGEBRAS

**Remark B.4.1** In finite dimensions, it can be shown that group orbits are always smooth (possibly immersed) manifolds. Group orbits generalise the notion of orbits of a dynamical system.  $\square$

**Exercise.** The flow of a vector field on  $M$  can be thought of as an action of  $\mathbb{R}$  on  $M$ . Show that in this case the general notion of group orbit reduces to the familiar notion of orbit used in dynamical systems.  $\star$

**Theorem B.4.1** *Orbits of proper Lie group actions are embedded submanifolds.*

This theorem is stated in Chapter 9 of [MaRa1994], who refer to [AbMa1978] for the proof.

**Example B.4.2 (Orbits of  $SO(3)$ )** *A simple example of a group orbit is the action of  $SO(3)$  on  $\mathbb{R}^3$  given by matrix multiplication: The action of  $A \in SO(3)$  on a point  $\mathbf{x} \in \mathbb{R}^3$  is simply the product  $A\mathbf{x}$ . In this case, the orbit of the origin is a single point (the origin itself), while the orbit of any other point is the sphere through that point.*

**Example B.4.3 (Orbits of a Lie group acting on itself)** *The action of a group  $G$  on itself from either the left or the right also produces group orbits. This action sets the stage for discussing the tangent-lifted action of a Lie group on its tangent bundle.*

### B.4.1 Left and right translations on a Lie group

Left and right translations on the group are denoted  $L_g$  and  $R_g$ , respectively. For example,  $L_g : G \rightarrow G$  is the map given by  $h \rightarrow gh$ , while  $R_g : G \rightarrow G$  is the map given by  $h \rightarrow hg$ , for  $g, h \in G$ .

- **Left translation**  $L_g : G \rightarrow G; h \rightarrow gh$  defines a transitive and free action of  $G$  on itself. Right multiplication  $R_g : G \rightarrow G; h \rightarrow hg$



## B.5 TANGENT AND COTANGENT LIFT ACTIONS 317

defines a right action, while  $h \rightarrow hg^{-1}$  defines a left action of  $G$  on itself.

- $G$  acts on  $G$  by conjugation,  $g \rightarrow I_g = R_{g^{-1}} \circ L_g$ . The map  $I_g : G \rightarrow G$  given by  $h \rightarrow ghg^{-1}$  is the **inner automorphism** associated with  $g$ . Orbits of this action are called **conjugacy classes**.
- Differentiating conjugation at the identity  $e$  gives the **adjoint action** of  $G$  on  $\mathfrak{g}$ :

$$\text{Ad}_g := T_e I_g : T_e G = \mathfrak{g} \rightarrow T_e G = \mathfrak{g}.$$

Explicitly, the adjoint action of  $G$  on  $\mathfrak{g}$  is given by

$$\text{Ad} : G \times \mathfrak{g} \rightarrow \mathfrak{g}, \quad \text{Ad}_g(\xi) = T_e(R_{g^{-1}} \circ L_g)\xi.$$

We have already seen an example of adjoint action for matrix Lie groups acting on matrix Lie algebras, when we defined  $S(t) = R_A(t)BR_A(t)^{-1} \in T_1S$  as a key step in the proof of Proposition B.2.1.

- The **coadjoint action** of  $G$  on  $\mathfrak{g}^*$ , the dual of the Lie algebra  $\mathfrak{g}$  of  $G$ , is defined as follows. Let  $\text{Ad}_g^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$  be the dual of  $\text{Ad}_g$ , defined by

$$\langle \text{Ad}_g^* \alpha, \xi \rangle = \langle \alpha, \text{Ad}_g \xi \rangle$$

for  $\alpha \in \mathfrak{g}^*$ ,  $\xi \in \mathfrak{g}$  and pairing  $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$ . Then the map

$$\Phi^* : G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^* \quad \text{given by} \quad (g, \alpha) \mapsto \text{Ad}_{g^{-1}}^* \alpha$$

is the coadjoint action of  $G$  on  $\mathfrak{g}^*$ .

## B.5 Tangent and cotangent lift actions

**Definition B.5.1 (Tangent lift action)** Let  $\Phi : G \times M \rightarrow M$  be a left action, and write  $\Phi_g(m) = \Phi(g, m)$ . The **tangent lift action** of  $G$  on the tangent bundle  $TM$  is defined by

$$gv = T_x \Phi_g(v), \tag{B.5.1}$$

for every  $v \in T_x M$ .

## 318 B : LIE GROUPS AND LIE ALGEBRAS

**Definition B.5.2 (Cotangent lift action)** The *cotangent lift* action of  $G$  on  $T^*M$  is defined by

$$g\alpha = (T_m\Phi_{g^{-1}})(\alpha), \quad (\text{B.5.2})$$

for every  $\alpha \in T^*M$ .

**Remark B.5.1** In standard calculus notation, the expression for tangent lift may be written as

$$T_x\Phi \cdot v = \left. \frac{d}{dt}\Phi(c(t)) \right|_{t=0} = \left. \frac{\partial\Phi}{\partial c}c'(t) \right|_{t=0} =: D\Phi(x) \cdot v, \quad (\text{B.5.3})$$

with  $c(0) = x$  and  $c'(0) = v$ .  $\square$

**Definition B.5.3** If  $X$  is a vector field on  $M$  and  $\phi$  is a differentiable map from  $M$  to itself, then the **push-forward** of  $X$  by  $\phi$  is the vector field  $\phi_*X$  defined by

$$(\phi_*X)(\phi(x)) = T_x\phi(X(x)) = D\phi(x) \cdot X(x).$$

That is, the following diagram commutes:

$$\begin{array}{ccc} TM & \xrightarrow{T\phi} & TM \\ \uparrow X & & \uparrow \phi_*X \\ M & \xrightarrow{\phi} & M \end{array}$$

If  $\phi$  is a diffeomorphism then the **pull-back**  $\phi^*X$  is also defined:

$$(\phi^*X)(x) = T_{\phi(x)}\phi^{-1}(X(\phi(x))) = D\phi^{-1}(\phi(x)) \cdot X(\phi(x)).$$

This denotes the standard calculus operation of computing the Jacobian of the inverse map  $\phi^{-1}$  acting on a vector field.

**Definition B.5.4 (Group action on vector fields)** Let  $\Phi : G \times M \rightarrow M$  be a left action, and write  $\Phi_g(m) = \Phi(g, m)$ . Then  $G$  has a left action on  $\mathfrak{X}(M)$  (the set of vector fields on  $M$ ) by push-forwards:  $gX = (\Phi_g)_*X$ . (The definition for right actions is the same, but written  $Xg$  instead of  $gX$ .)

**Definition B.5.5 (Group-invariant vector fields)** Let  $G$  act on  $M$  on the left. A vector field  $X$  on  $M$  is *invariant* with respect to this action (we often say “ $G$ -invariant” if the action is understood) if  $gX = X$  for all  $g \in G$ ; equivalently (using all of the above definitions!)  $g(X(x)) = X(gx)$  for all  $g \in G$  and all  $x \in X$ . (Similarly for right actions.)

**Definition B.5.6 (Left-invariant vector fields)** Consider the left action of  $G$  on itself by left multiplication,

$$\Phi_g(h) = L_g(h) = gh.$$

A vector field on  $G$  that is invariant with respect to this action is called *left-invariant*. From Definition B.5.5, we see that  $X$  is left-invariant if and only if  $g(X(h)) = X(gh)$ , which in less compact notation means  $T_h L_g X(h) = X(gh)$ . The set of all such vector fields is written  $\mathfrak{X}^L(G)$ .

**Theorem B.5.1** Given a  $\xi \in T_e G$ , define  $X_\xi^L(g) = g\xi$  (recall  $g\xi \equiv T_e L_g \xi$ ). Then  $X_\xi^L$  is the unique left-invariant vector field such that  $X_\xi^L(e) = \xi$ .

**Proof.** To show that  $X_\xi^L$  is left-invariant, we need to show that  $g(X_\xi^L(h)) = X_\xi^L(gh)$  for every  $g, h \in G$ . This follows from the definition of  $X_\xi^L$  and the associativity property of Lie group actions,

$$g(X_\xi^L(h)) = g(h\xi) = (gh)\xi = X_\xi^L(gh).$$

We repeat the last line in less compact notation,

$$T_h L_g (X_\xi^L(h)) = T_h L_g (h\xi) = T_e L_{gh} \xi = X_\xi^L(gh).$$

For uniqueness, suppose  $X$  is left-invariant and  $X(e) = \xi$ . Then for any  $g \in G$ , we have  $X(g) = g(X(e)) = g\xi = X_\xi^L(g)$ . ■

The proof also yields the following.

**Corollary B.5.1** The map  $\xi \mapsto X_\xi^L$  is a vector space isomorphism from  $T_e G$  to  $\mathfrak{X}^L(G)$ .

## 320 B : LIE GROUPS AND LIE ALGEBRAS

**Remark B.5.2** Because of this vector space isomorphism, the left-invariant vector fields could have been used to define the Lie algebra of a Lie group. See [Ol2000] for an example of this approach using the corresponding isomorphism for the right-invariant vector fields.  $\square$

All of the above definitions and theorems have analogues for right actions. The definitions of right-invariant,  $\mathfrak{X}^R(G)$  and  $X_\xi^R$ , use the right action of  $G$  on itself defined by  $\Psi(g, h) = R_g(h) = hg$ . For example, the tangent lift for right action of  $G$  on the tangent bundle  $TM$  is defined by  $vg = T_x\Psi_g(v)$  for every  $v \in T_xM$ .

**Exercise.** Show that  $\Phi_g(h) = hg^{-1}$  defines a left action of  $G$  on itself.  $\star$

The map  $\xi \mapsto X_\xi^L$  may be used to relate the Lie bracket on  $\mathfrak{g}$ , defined as  $[\xi, \eta] = \text{ad}_\xi\eta$ , with the Jacobi–Lie bracket on vector fields.

## B.6 Jacobi–Lie bracket

**Definition B.6.1 (Jacobi–Lie bracket)** The *Jacobi–Lie bracket* on  $\mathfrak{X}(M)$  is defined in local coordinates by

$$[X, Y]_{J-L} \equiv (DX) \cdot Y - (DY) \cdot X,$$

which, in finite dimensions, is equivalent to

$$[X, Y]_{J-L} \equiv -(X \cdot \nabla)Y + (Y \cdot \nabla)X \equiv -[X, Y].$$

### Theorem B.6.1 (Properties of the Jacobi–Lie bracket)

- The Jacobi–Lie bracket satisfies

$$[X, Y]_{J-L} = \mathcal{L}_X Y \equiv \left. \frac{d}{dt} \right|_{t=0} \Phi_t^* Y,$$

where  $\Phi_t$  is the flow of  $X$ . (This is coordinate-free, and can be used as an alternative definition.)

- This bracket makes  $\mathfrak{X}^L(M)$  a Lie algebra with  $[X, Y]_{J-L} = -[X, Y]$ , where  $[X, Y]$  is the Lie algebra bracket on  $\mathfrak{X}(M)$ .
- $\phi_*[X, Y] = [\phi_*X, \phi_*Y]$  for any differentiable  $\phi : M \rightarrow M$ .

**Remark B.6.1** The first property of the Jacobi–Lie bracket is proved for matrices in Section B.7. The other two properties are proved below for the case that  $M$  is the Lie group  $G$ .  $\square$

**Theorem B.6.2**  $\mathfrak{X}^L(G)$  is a subalgebra of  $\mathfrak{X}(G)$ .

**Proof.** Let  $X, Y \in \mathfrak{X}^L(G)$ . Using the last item of the previous theorem, and then the  $G$ -invariance of  $X$  and  $Y$ , gives the push-forward relations

$$(L_g)_* [X, Y]_{J-L} = [(L_g)_* X, (L_g)_* Y]_{J-L},$$

for all  $g \in G$ . Hence  $[X, Y]_{J-L} \in \mathfrak{X}^L(G)$ . This is the second property in Theorem B.6.1.  $\blacksquare$

**Theorem B.6.3**

$$[X_\xi^L, X_\eta^L]_{J-L}(e) = [\xi, \eta],$$

for every  $\xi, \eta \in \mathfrak{g}$ , where the bracket on the right is the Jacobi–Lie bracket. (That is, the Lie bracket on  $\mathfrak{g}$  is the pull-back of the Jacobi–Lie bracket by the map  $\xi \mapsto X_\xi^L$ .)

**Proof.** The proof of this theorem for matrix Lie algebras is relatively easy: we have already seen that  $\text{ad}_A B = AB - BA$ . On the other hand, since  $X_A^L(C) = CA$  for all  $C$ , and this is linear in  $C$ , we have  $DX_B^L(I) \cdot A = AB$ , so

$$\begin{aligned} [A, B] &= [X_A^L, X_B^L]_{J-L}(I) \\ &= DX_B^L(I) \cdot X_A^L(I) - DX_A^L(I) \cdot X_B^L(I) \\ &= DX_B^L(I) \cdot A - DX_A^L(I) \cdot B \\ &= AB - BA. \end{aligned}$$

## 322 B : LIE GROUPS AND LIE ALGEBRAS

This is the third property of the Jacobi–Lie bracket listed in Theorem B.6.1. For the general proof, see [MaRa1994], Proposition 9.14. ■

**Remark B.6.2** This theorem, together with the second property in Theorem B.6.1, proves that the Jacobi–Lie bracket makes  $\mathfrak{g}$  into a Lie algebra. □

**Remark B.6.3** By Theorem B.6.2, the vector field  $[X_\xi^L, X_\eta^L]$  is left-invariant. Since  $[X_\xi^L, X_\eta^L]_{J-L}(e) = [\xi, \eta]$ , it follows that

$$[X_\xi^L, X_\eta^L] = X_{[\xi, \eta]}^L.$$

□

**Definition B.6.2 (Infinitesimal generator)** Let  $\Phi : G \times M \rightarrow M$  be a left action, and let  $\xi \in \mathfrak{g}$ . Let  $g(t)$  be a path in  $G$  such that  $g(0) = e$  and  $g'(0) = \xi$ . Then the **infinitesimal generator** of the action in the  $\xi$  direction is the vector field  $\xi_M$  on  $M$  defined by

$$\xi_M(x) = \left. \frac{d}{dt} \right|_{t=0} \Phi_{g(t)}(x).$$

**Remark B.6.4** Note: This definition does not depend on the choice of  $g(t)$ . For example, the choice in [MaRa1994] is  $\exp(t\xi)$ , where  $\exp$  denotes the exponentiation on Lie groups (not defined here). □

**Exercise.** Consider the action of  $SO(3)$  on the unit sphere  $S^2$  around the origin, and let  $\xi = (0, 0, 1)$ . Sketch the vector field  $\xi_M$ .

Hint: The vectors all point eastward. ★

**Theorem B.6.4** *For any left action of  $G$ , the Jacobi–Lie bracket of infinitesimal generators is related to the Lie bracket on  $\mathfrak{g}$  as follows (note the minus sign):*

$$[\xi_M, \eta_M] = -[\xi, \eta]_M.$$

For a proof, see [MaRa1994], Proposition 9.3.6.

**Exercise.** Express the statements and formulas of this appendix for the case of  $SO(3)$  action on its Lie algebra  $so(3)$ . (Hint: Look at the previous section.) Wherever possible, translate these formulas to  $\mathbb{R}^3$  by using the hat map  $\hat{\cdot} : so(3) \rightarrow \mathbb{R}^3$ .

Write the Lie algebra for  $so(3)$  using the Jacobi–Lie bracket in terms of linear vector fields on  $\mathbb{R}^3$ . What are the characteristic curves of these linear vector fields? ★

## B.7 Lie derivative and Jacobi–Lie bracket

### B.7.1 Lie derivative of a vector field

**Definition B.7.1 (Lie derivative)** *Let  $X$  and  $Y$  be two vector fields on the same manifold  $M$ . The Lie derivative of  $Y$  with respect to  $X$  is*

$$\mathcal{L}_X Y \equiv \left. \frac{d}{dt} \Phi_t^* Y \right|_{t=0},$$

where  $\Phi$  is the flow of  $X$ .

**Remark B.7.1** The Lie derivative  $\mathcal{L}_X Y$  is the derivative of  $Y$  in the direction given by  $X$ . Its definition is coordinate-independent. In contrast,  $DY \cdot X$  is also the derivative of  $Y$  in the  $X$  direction, but

## 324 B : LIE GROUPS AND LIE ALGEBRAS

its value depends on the coordinate system, and in particular does not usually equal  $\mathcal{L}_X Y$  in a given coordinate system.  $\square$

**Theorem B.7.1**

$$\mathcal{L}_X Y = [X, Y],$$

where the bracket on the right is the Jacobi–Lie bracket.

**Proof.** Assume that  $M$  is finite-dimensional and work in local coordinates. That is, consider everything as matrices, which allows the use of the product rule and the identities  $(M^{-1})' = -M^{-1}M'M^{-1}$  and  $\frac{d}{dt}(D\Phi_t(x)) = D\left(\frac{d}{dt}\Phi_t\right)(x)$ .

$$\begin{aligned} \mathcal{L}_X Y(x) &= \left. \frac{d}{dt} \Phi_t^* Y(x) \right|_{t=0} \\ &= \left. \frac{d}{dt} (D\Phi_t(x))^{-1} Y(\Phi_t(x)) \right|_{t=0} \\ &= \left[ \left( \frac{d}{dt} (D\Phi_t(x))^{-1} \right) Y(\Phi_t(x)) \right. \\ &\quad \left. + (D\Phi_t(x))^{-1} \frac{d}{dt} Y(\Phi_t(x)) \right]_{t=0} \\ &= \left[ - (D\Phi_t(x))^{-1} \left( \frac{d}{dt} D\Phi_t(x) \right) (D\Phi_t(x))^{-1} Y(\Phi_t(x)) \right. \\ &\quad \left. + (D\Phi_t(x))^{-1} \frac{d}{dt} Y(\Phi_t(x)) \right]_{t=0}. \end{aligned}$$

Regrouping yields

$$\begin{aligned} \mathcal{L}_X Y(x) &= \left[ - \left( \frac{d}{dt} D\Phi_t(x) \right) Y(x) + \frac{d}{dt} Y(\Phi_t(x)) \right]_{t=0} \\ &= -D \left( \left. \frac{d}{dt} \Phi_t(x) \right|_{t=0} \right) Y(x) + DY(x) \left( \left. \frac{d}{dt} \Phi_t(x) \right|_{t=0} \right) \\ &= -DX(x) \cdot Y(x) + DY(x) \cdot X(x) \\ &= [X, Y]_{J-L}(x). \end{aligned}$$

Therefore  $\mathcal{L}_X Y = [X, Y]_{J-L}$ .  $\blacksquare$



### B.7.2 Vector fields in ideal fluid dynamics

The Lie derivative formula for vector fields also applies in infinite dimensions, although the proof is more elaborate. For example, the equation for the vorticity dynamics of an Euler fluid with velocity  $\mathbf{u}$  (with  $\operatorname{div} \mathbf{u} = 0$ ) and vorticity  $\omega = \operatorname{curl} \mathbf{u}$  may be written as a Lie derivative formula,

$$\begin{aligned} \partial_t \omega &= -\mathbf{u} \cdot \nabla \omega + \omega \cdot \nabla \mathbf{u} \\ &= -[u, \omega]_{J-L} \\ &= \operatorname{ad}_u \omega \\ &= -\mathcal{L}_u \omega. \end{aligned}$$

These equations each express the invariance of the vorticity vector field  $\omega$  under the flow of its corresponding divergenceless velocity vector field  $u$ . This invariance is also encapsulated in the language of fluid dynamics as

$$\frac{d}{dt} \left( \omega \cdot \frac{\partial}{\partial \mathbf{x}} \right) = 0, \quad \text{along} \quad \frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}, t) = \operatorname{curl}^{-1} \omega.$$

Here, the curl-inverse operator is defined by the *Biot–Savart law*,

$$\mathbf{u} = \operatorname{curl}^{-1} \omega = \operatorname{curl}(-\Delta)^{-1} \omega,$$

which follows from the identity

$$\operatorname{curl} \operatorname{curl} \mathbf{u} = -\Delta \mathbf{u} + \nabla \operatorname{div} \mathbf{u}.$$

Thus, in coordinates,

$$\begin{aligned} \frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}, t) &\implies \mathbf{x}(t, \mathbf{x}_0) \quad \text{with} \quad \mathbf{x}(0, \mathbf{x}_0) = \mathbf{x}_0 \quad \text{at} \quad t = 0, \\ \text{and} \quad \omega^j(\mathbf{x}(t, \mathbf{x}_0), t) \frac{\partial}{\partial x^j(t, \mathbf{x}_0)} &= \omega^A(\mathbf{x}_0) \frac{\partial}{\partial x_0^A}, \end{aligned}$$

in which one sums over repeated indices. Consequently, one may write

$$\omega^j(\mathbf{x}(t, \mathbf{x}_0), t) = \omega^A(\mathbf{x}_0) \frac{\partial x^j(t, \mathbf{x}_0)}{\partial x_0^A}. \quad (\text{B.7.1})$$

## 326 B : LIE GROUPS AND LIE ALGEBRAS

This is Cauchy's (1859) solution of *Euler's equation for vorticity*,

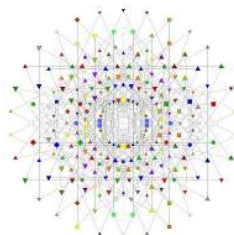
$$\partial_t \omega = - [\text{curl}^{-1} \omega, \omega]_{J-L} = \text{ad}_u \omega \quad \text{with} \quad \mathbf{u} = \text{curl}^{-1} \omega.$$

This type of Lie derivative equation often appears in the text. In it, the vorticity  $\omega$  evolves by the ad action of the right-invariant vector field  $\mathbf{u} = \text{curl}^{-1} \omega$ . The Cauchy solution is the tangent lift of this flow.

**Exercise.** Discuss the cotangent lift of this flow and how it might be related to the Kelvin circulation theorem,

$$\frac{d}{dt} \int_{c(\mathbf{u})} \mathbf{u} \cdot d\mathbf{x} = 0,$$

for a circuit  $c(\mathbf{u})$  moving with the flow. ★



# C

## ENHANCED COURSEWORK

---

### Contents

C.1	Variations on rigid-body dynamics	328
C.1.1	Two times	328
C.1.2	Rotations in complex space	334
C.1.3	Rotations in four dimensions: $SO(4)$	338
C.2	$\mathbb{C}^3$ oscillators	343
C.3	Momentum maps for $GL(n, \mathbb{R})$	348
C.4	Motion on the symplectic Lie group $Sp(2)$	354
C.5	Two coupled rigid bodies	359

---

## C.1 Variations on rigid-body dynamics

### C.1.1 Two times

**Scenario C.1.1** *The Bichrons are an alien life form who use two-dimensional time  $u = (s, t) \in \mathbb{R}^2$  for time travel. To decide whether we are an intelligent life form, they require us to define spatial and body angular velocity for free rigid rotation in their two time dimensions. What should we tell them?*

**Answer. (Bichrons)** Following the approach to rotating motion taken in Section 2, let's define a trajectory of a moving point  $\mathbf{x} = \mathbf{r}(u) \in \mathbb{R}^3$  as a smooth invertible map  $\mathbf{r} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  with "time"  $u \in \mathbb{R}^2$ , so that  $u = (s, t)$ . Suppose the components of the trajectory are given in terms of a fixed and a moving orthonormal frame by

$$\begin{aligned} \mathbf{r}(u) &= r_0^A(u) \mathbf{e}_A(0) && \text{fixed frame,} \\ &= r^a \mathbf{e}_a(u) && \text{moving frame,} \end{aligned}$$

with moving orthonormal frame defined by  $O : \mathbb{R}^2 \rightarrow SO(3)$ , so that

$$\mathbf{e}_a(u) = O(u) \mathbf{e}_a(0),$$

where  $O(u)$  is a *surface* in  $SO(3)$  parameterised by the two times  $u = (s, t)$ . The exterior derivative<sup>1</sup> of the moving frame relation above yields the infinitesimal spatial displacement,

$$d\mathbf{r}(u) = r^a d\mathbf{e}_a(u) = r^a dOO^{-1}(u) \mathbf{e}_a(u) = \widehat{\omega}(u) \mathbf{r},$$

in which  $\widehat{\omega}(u) = dOO^{-1}(u) \in so(3)$  is the one-form for spatial angular displacement. One denotes

$$dO = O' ds + \dot{O} dt,$$

<sup>1</sup>This subsection uses the notation of differential forms and wedge products. Readers unfamiliar with it may regard this subsection as cultural background.

so that the spatial angular displacement is the right-invariant  $so(3)$ -valued one-form

$$\widehat{\omega}(u) = dOO^{-1} = (O' ds + \dot{O} dt)O^{-1}.$$

Likewise, the body angular displacement is the left-invariant  $so(3)$ -valued one-form

$$\begin{aligned}\widehat{\Omega}(u) &= O^{-1}\widehat{\omega}(u)O = O^{-1}dO & (C.1.1) \\ &= O^{-1}(O' ds + \dot{O} dt) =: \widehat{\Omega}_s ds + \widehat{\Omega}_t dt,\end{aligned}$$

and  $\widehat{\Omega}_s$  and  $\widehat{\Omega}_t$  are its *two body angular velocities*.

This is the answer the Bichrons wanted: For them, free rotation takes place on a space-time surface in  $SO(3)$  and it has two body angular velocities because such a surface has two independent tangent vectors.  $\blacktriangle$

**Scenario C.1.2** *What would the Bichrons do with this information?*

**Answer.** To give an idea of what the Bichrons might do with our answer, let us define the *coframe* at position  $\mathbf{x} = \mathbf{r}(u)$  as the infinitesimal displacement in *body coordinates*,

$$\Xi = O^{-1}d\mathbf{r}. \quad (C.1.2)$$

Taking its exterior derivative gives the two-form,

$$d\Xi = -O^{-1}dO \wedge O^{-1}d\mathbf{r} = -\widehat{\Omega} \wedge \Xi, \quad (C.1.3)$$

in which the left-invariant  $so(3)$ -valued one-form  $\widehat{\Omega} = O^{-1}dO$  encodes the exterior derivative of the coframe as a rotation by the body angular displacement. In differential geometry,  $\widehat{\Omega}$  is called the *connection form* and Equation (C.1.3) is called *Cartan's first structure equation* for a moving orthonormal frame [Fl1963, Da1994]. Taking another exterior derivative gives zero (because  $d^2 = 0$ ) in the form of

$$0 = d^2\Xi = -d\widehat{\Omega} \wedge \Xi - \widehat{\Omega} \wedge d\Xi = -(d\widehat{\Omega} + \widehat{\Omega} \wedge \widehat{\Omega}) \wedge \Xi.$$

## 330 C : ENHANCED COURSEWORK

Hence we have *Cartan's second structure equation*,

$$d\widehat{\Omega} + \widehat{\Omega} \wedge \widehat{\Omega} = 0. \quad (\text{C.1.4})$$

The left-hand side of this equation is called the *curvature two-form* associated with the connection form  $\widehat{\Omega}$ . The interpretation of (C.1.4) is that the connection form  $\widehat{\Omega} = O^{-1}dO$  has zero curvature. This makes sense because the rotating motion takes place in Euclidean space,  $\mathbb{R}^3$ , which is flat.

Of course, one may also prove the *zero curvature relation* (C.1.4) directly from the definition  $\widehat{\Omega} = O^{-1}dO$  by computing

$$d\widehat{\Omega} = d(O^{-1}dO) = -O^{-1}dO \wedge O^{-1}dO = -\widehat{\Omega} \wedge \widehat{\Omega}.$$

Expanding this out using the two angular velocities  $\widehat{\Omega}_s = O^{-1}O'$  and  $\widehat{\Omega}_t = O^{-1}\dot{O}$  gives (by using antisymmetry of the wedge product,  $ds \wedge dt = -dt \wedge ds$ )

$$\begin{aligned} d\widehat{\Omega}(u) &= d(\widehat{\Omega}_s ds + \widehat{\Omega}_t dt) \\ &= -(\widehat{\Omega}_s ds + \widehat{\Omega}_t dt) \wedge (\widehat{\Omega}_s ds + \widehat{\Omega}_t dt) \\ &= -\widehat{\Omega} \wedge \widehat{\Omega} \\ &= \frac{\partial \widehat{\Omega}_s}{\partial t} dt \wedge ds + \frac{\partial \widehat{\Omega}_t}{\partial s} ds \wedge dt \\ &= -\widehat{\Omega}_s \widehat{\Omega}_t ds \wedge dt - \widehat{\Omega}_t \widehat{\Omega}_s dt \wedge ds \\ &= \left( \frac{\partial \widehat{\Omega}_t}{\partial s} - \frac{\partial \widehat{\Omega}_s}{\partial t} \right) ds \wedge dt \\ &= (\widehat{\Omega}_t \widehat{\Omega}_s - \widehat{\Omega}_s \widehat{\Omega}_t) ds \wedge dt \\ &=: [\widehat{\Omega}_t, \widehat{\Omega}_s] ds \wedge dt. \end{aligned}$$

Since  $ds \wedge dt \neq 0$ , this equality implies that the coefficients are equal. In other words, this calculation proves the following.

**Proposition C.1.1** *The zero curvature relation (C.1.4) may be expressed equivalently as*

$$\frac{\partial \widehat{\Omega}_t}{\partial s} - \frac{\partial \widehat{\Omega}_s}{\partial t} = \widehat{\Omega}_t \widehat{\Omega}_s - \widehat{\Omega}_s \widehat{\Omega}_t = [\widehat{\Omega}_t, \widehat{\Omega}_s], \quad (\text{C.1.5})$$

*in terms of the two angular velocities,  $\widehat{\Omega}_s = O^{-1}O'$  and  $\widehat{\Omega}_t = O^{-1}\dot{O}$ .* ▲

**Remark C.1.1** The component form (C.1.5) of the zero curvature relation (C.1.4) arises in Equation (10.3.10) of Chapter 11 in the derivation of the Euler–Poincaré equations for the dynamics of the classical spin chain. It also arises in many places in the theory of soliton solutions of completely integrable Hamiltonian partial differential equations, see, e.g., [FaTa1987]. □

**Exercise.** Why is  $\widehat{\Omega}$  called a connection form? ★

**Answer.** Consider the one-form Equation (C.1.2) written in components as

$$\Xi^j = \Xi_\alpha^j(r) dr^\alpha, \quad (\text{C.1.6})$$

in which the matrix  $\Xi_\alpha^j(r)$  depends on spatial location, and it need not be orthogonal. In the basis  $\Xi^j(r)$ , a one-form  $v$  may be expanded in components as

$$v = v_j \Xi^j. \quad (\text{C.1.7})$$

Its differential is computed in this basis as

$$\begin{aligned} dv &= d(v_j \Xi^j) \\ &= dv_j \wedge \Xi^j + v_j d\Xi^j. \end{aligned}$$

Substituting Equation (C.1.3) in components as

$$d\Xi^j = -\widehat{\Omega}_k^j \wedge \Xi^k \quad (\text{C.1.8})$$

## 332 C : ENHANCED COURSEWORK

then yields the differential two-form,

$$\begin{aligned} dv &= (dv_k - v_j \widehat{\Omega}_k^j) \wedge \Xi^k \\ &=: (dv_k - v_j \Gamma_{kl}^j \Xi^l) \wedge \Xi^k \\ &=: Dv_k \wedge \Xi^k. \end{aligned} \quad (\text{C.1.9})$$

The last equation defines the covariant exterior derivative operation  $D$  in the basis of one-form displacements  $\Xi(r)$ . The previous equation introduces the quantities  $\Gamma_{kl}^j$  defined as

$$\widehat{\Omega}_k^j = \Gamma_{kl}^j \Xi^l. \quad (\text{C.1.10})$$

$\Gamma_{kl}^j$  are the Christoffel coefficients in the *local coframe* given by Equation (C.1.6). These are the standard connection coefficients for curvilinear geometry. ▲

**Exercise.** Prove from their definition in formula (C.1.9) that the Christoffel coefficients are symmetric under the exchange of indices,  $\Gamma_{kl}^j = \Gamma_{lk}^j$ . ★

**Definition C.1.1 (Body covariant derivative)** *The relation in Equation (C.1.9)*

$$Dv_k := dv_k - v_j \widehat{\Omega}_k^j = dv_k - v_j \Gamma_{kl}^j \Xi^l \quad (\text{C.1.11})$$

*defines the components of the **covariant derivative** of the one-form  $v$  in the body frame; that is, in the  $\Xi$ -basis. Thus,  $\widehat{\Omega}$  is a connection form in the standard sense of differential geometry [F11963, Da1994].*

**Remark C.1.2 (Metric tensors)** The metric tensors in the two bases of infinitesimal displacements  $dr$  and  $\Xi$  are related by requiring that the element of length measured in either basis must be the same. That is,

$$ds^2 = g_{\alpha\beta} dr^\alpha \otimes dr^\beta = \delta_{jk} \Xi^j \otimes \Xi^k, \quad (\text{C.1.12})$$



where  $\otimes$  is the symmetric tensor product. This implies a relation between the metrics,

$$g_{\alpha\beta} = \delta_{jk} \Xi_{\alpha}^j \Xi_{\beta}^k, \quad (\text{C.1.13})$$

which, in turn, implies

$$\Gamma_{\beta\mu}^{\nu}(r) = \frac{1}{2} g^{\nu\alpha} \left[ \frac{\partial g_{\alpha\mu}(r)}{\partial r^{\beta}} + \frac{\partial g_{\alpha\beta}(r)}{\partial r^{\mu}} - \frac{\partial g_{\beta\mu}(r)}{\partial r^{\alpha}} \right]. \quad (\text{C.1.14})$$

This equation identifies  $\Gamma_{\beta\mu}^{\nu}(r)$  as the Christoffel coefficients in the *spatial basis*. Note that the spatial Christoffel coefficients are symmetric under the exchange of indices,  $\Gamma_{\beta\mu}^{\nu}(r) = \Gamma_{\mu\beta}^{\nu}(r)$ .  $\square$

**Definition C.1.2 (Spatial covariant derivative)** For the spatial metric  $g_{\alpha\mu}$ , the *covariant derivative* of the one-form  $v = v_{\beta} dr^{\beta}$  in the spatial coordinate basis  $dr^{\beta}$  is defined by the standard formula, cf. Equation (C.1.11),

$$Dv_{\beta} = dv_{\beta} - v_{\nu} \Gamma_{\alpha\beta}^{\nu} dr^{\alpha},$$

or, in components,

$$\nabla_{\alpha} v_{\beta} = \partial_{\alpha} v_{\beta} - v_{\nu} \Gamma_{\alpha\beta}^{\nu}.$$

**Remark C.1.3** Thus, in differential geometry, the connection one-form (C.1.10) in the local coframe encodes the Riemannian Christoffel coefficients for the spatial coordinates, via the equivalence of metric length (C.1.13) as measured in either set of coordinates. The left-invariant  $so(3)$ -valued one-form  $\hat{\Omega} = O^{-1}dO$  that the Bichrons need for keeping track of the higher-dimensional time components of their rotations in body coordinates in (C.1.1) plays the same role for their time surfaces in  $SO(3)$  as the connection one-form does for taking covariant derivatives in a local coframe. For more discussion of connection one-forms and their role in differential geometry, see, e.g., [Fl1963, Da1994].  $\square$

**Exercise.** Write the two-time version of the Euler–Poincaré equation for a left-invariant Lagrangian defined on  $so(3)$ .

Hint: Take a look at Chapter 10.



### C.1.2 Rotations in complex space

**Scenario C.1.3** *The Bers are another alien life form who use one-dimensional time  $t \in \mathbb{R}$  (thankfully), but their spatial coordinates are complex  $\mathbf{z} \in \mathbb{C}^3$ , while ours are real  $\mathbf{x} \in \mathbb{R}^3$ . They test us to determine whether we are an intelligent life form by requiring us to write the equations for rigid-body motion for body angular momentum coordinates  $\mathbf{L} \in \mathbb{C}^3$ .*

*Their definition of a rigid body requires its moment of inertia  $\mathbb{I}$ , rotational kinetic energy  $\frac{1}{2}\mathbf{L} \cdot \mathbb{I}^{-1}\mathbf{L}$  and magnitude of body angular momentum  $\sqrt{\mathbf{L} \cdot \mathbf{L}}$  all to be real. They also tell us these rigid-body equations must be invariant under the operations of parity  $\mathcal{P} \mathbf{z} \rightarrow -\mathbf{z}^*$  and time reversal  $\mathcal{T} : t \rightarrow -t$ . What equations should we give them? Are these equations the same as ours in real body angular momentum coordinates? Keep your approach general for as long as you like, but if you wish to simplify, work out your results with the simple example in which  $\mathbb{I} = \text{diag}(1, 2, 3)$ .*

**Answer.** Euler's equations for free rotational motion of a rigid body about its centre of mass may be expressed in real vector coordinates  $\mathbf{L} \in \mathbb{R}^3$  ( $\mathbf{L}$  is the body angular momentum vector) as

$$\dot{\mathbf{L}} = \frac{\partial C}{\partial \mathbf{L}} \times \frac{\partial E}{\partial \mathbf{L}}, \quad (\text{C.1.15})$$

where  $C$  and  $E$  are conserved quadratic functions defined by

$$C(\mathbf{L}) = \frac{1}{2}\mathbf{L} \cdot \mathbf{L}, \quad E(\mathbf{L}) = \frac{1}{2}\mathbf{L} \cdot \mathbb{I}^{-1}\mathbf{L}. \quad (\text{C.1.16})$$

Here,  $\mathbb{I}^{-1} = \text{diag}(I_1^{-1}, I_2^{-1}, I_3^{-1})$  is the inverse of the (real) moment of inertia tensor in principal axis coordinates. These equations are  $\mathcal{PT}$ -symmetric; they are invariant under spatial reflections of the angular momentum components in the body  $P : \mathbf{L} \rightarrow \mathbf{L}$  composed with time reversal  $T : \mathbf{L} \rightarrow -\mathbf{L}$ . The simplifying choice  $\mathbb{I}^{-1} = \text{diag}(1, 2, 3)$  reduces the dynamics (C.1.15) to

$$\dot{L}_1 = L_2 L_3, \quad \dot{L}_2 = -2L_1 L_3, \quad \dot{L}_3 = L_1 L_2, \quad (\text{C.1.17})$$

which may also be written equivalently as

$$\dot{\mathbf{L}} = \mathbf{L} \times \mathbf{KL}, \quad (\text{C.1.18})$$

with  $\mathbf{K} = \text{diag}(-1, 0, 1)$ .

Since  $\mathbf{L}$  is complex, we set  $\mathbf{L} = \mathbf{x} + i\mathbf{y}$  and obtain *four* conservation laws, namely the real and imaginary parts of  $C(\mathbf{L}) = \frac{1}{2}\mathbf{L} \cdot \mathbf{L}$  and  $H(\mathbf{L}) = \frac{1}{2}\mathbf{L} \cdot \mathbf{KL}$ , expressed as

$$C(\mathbf{L}) = \frac{1}{2}\mathbf{x} \cdot \mathbf{x} - \frac{1}{2}\mathbf{y} \cdot \mathbf{y} + i\mathbf{x} \cdot \mathbf{y}, \quad (\text{C.1.19})$$

$$H(\mathbf{L}) = \frac{1}{2}\mathbf{x} \cdot \mathbf{Kx} - \frac{1}{2}\mathbf{y} \cdot \mathbf{Ky} + i\mathbf{x} \cdot \mathbf{Ky}. \quad (\text{C.1.20})$$

The solutions to Euler's equations that have been studied in the past are the *real* solutions to (C.1.17), that is, the solutions for which  $\mathbf{y} = 0$ . For this case the phase space is three-dimensional and the two conserved quantities are

$$C = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2), \quad H = -\frac{1}{2}x_1^2 + \frac{1}{2}x_3^2. \quad (\text{C.1.21})$$

If we take  $C = \frac{1}{2}$ , then the phase-space trajectories are constrained to a sphere of radius 1. There are six critical points located at  $(\pm 1, 0, 0)$ ,  $(0, \pm 1, 0)$  and  $(0, 0, \pm 1)$ . These are the conventional trajectories that are discussed in standard textbooks on dynamical systems [MaRa1994].

**Exercise.** When  $H = 0$ , show that the resulting equation is a first integral of the simple pendulum problem. ★

Let us now examine the complex  $\mathcal{PT}$ -symmetric solutions to Euler's equations. The equation set (C.1.16) is six-dimensional. However, a reduction in dimension occurs because the requirement of  $\mathcal{PT}$  symmetry requires

## 336 C : ENHANCED COURSEWORK

the constants of motion  $C$  and  $H$  in (C.1.20) to be real. The vanishing of the imaginary parts of  $C$  and  $H$  gives the two equations

$$\mathbf{x} \cdot \mathbf{y} = 0, \quad \mathbf{x} \cdot \mathbf{K}\mathbf{y} = 0. \quad (\text{C.1.22})$$

These two bilinear constraints may be used to eliminate the  $\mathbf{y}$  terms in the complex Equations (C.1.17). When this elimination is performed using the definition  $\mathbf{K} = \text{diag}(-1, 0, 1)$ , one obtains the following real equations for  $\mathbf{x}$  on the  $\mathcal{PT}$  constraint manifolds (C.1.22):

$$\dot{\mathbf{x}} = \mathbf{x} \times \mathbf{K}\mathbf{x} + M(\mathbf{x}) \mathbf{x}. \quad (\text{C.1.23})$$

Here, the scalar function  $M = PN/D$ , where the functions  $P$ ,  $N$  and  $D$  are given by

$$P(\mathbf{x}) = 2x_1x_2x_3, \quad N(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - 1, \quad (\text{C.1.24})$$

$$D(\mathbf{x}) = \left| \text{Re} \left( \frac{\partial C}{\partial \mathbf{L}} \times \frac{\partial H}{\partial \mathbf{L}} \right) \right|^2 = x_1^2x_2^2 + x_2^2x_3^2 + 4x_1^2x_3^2.$$

The system (C.1.23) has nonzero divergence, so it cannot be Hamiltonian even though it arises from constraining a Hamiltonian system. Nonetheless, the system has two *additional* real conservation laws, and it reduces to the integrable form

$$\dot{x}_1 = x_2x_3 (1 + 2x_1^2N/D), \quad (\text{C.1.25})$$

$$\dot{x}_2 = -2x_1x_3 (1 - x_2^2N/D),$$

$$\dot{x}_3 = x_1x_2 (1 + 2x_3^2N/D), \quad (\text{C.1.26})$$

on level sets of two conserved quantities:

$$A = \frac{(N+1)^2N}{D}, \quad (\text{C.1.27})$$

$$B = \frac{x_1^2 - x_3^2}{D} (2x_2^2x_3^2 + 4x_1^2x_3^2 + x_2^4 + 2x_1^2x_2^2 - x_2^2).$$

Hence, the motion takes place in  $\mathbb{R}^3$  on the intersection of the level sets of these two conserved quantities. These

quantities vanish when either  $N = 0$  (the unit sphere) or  $x_3^2 - x_1^2 = 0$  (the degenerate hyperbolic cylinder). On these level sets of the conserved quantities the motion Equations (C.1.23) restrict to Equations (C.1.17) for the original real rigid body. ▲

**Remark C.1.4** We are dealing with rotations of the group of complex  $3 \times 3$  orthogonal matrices with unit determinant acting on complex three-vectors. These are the linear maps,  $SO(3, \mathbb{C}) \times \mathbb{C}^3 \mapsto \mathbb{C}^3$ .

Euler's Equations (C.1.15) for complex body angular momentum describe geodesic motion on  $SO(3, \mathbb{C})$  with respect to the metric given by the trace norm  $g(\Omega, \Omega) = \frac{1}{2} \text{trace}(\Omega^T \mathbb{I} \Omega)$  for the real symmetric moment of inertia tensor  $\mathbb{I}$  and left-invariant Lie algebra element  $\Omega(t) = g^{-1}(t)\dot{g}(t) \in so(3, \mathbb{C})$ . Because  $SO(3, \mathbb{C})$  is orthogonal,  $\Omega \in so(3, \mathbb{C})$  is a  $3 \times 3$  complex skew-symmetric matrix, which may be identified with complex vectors  $\hat{\Omega} \in \mathbb{C}^3$  by  $(\Omega)_{jk} = -\hat{\Omega}^i \epsilon_{ijk}$ . Euler's Equations (C.1.15) follow from Hamilton's principle in Euler-Poincaré or Lie-Poisson form:

$$\dot{\mu} = \text{ad}_{\Omega}^* \mu = \{\mu, H\}, \quad (\text{C.1.28})$$

where

$$\frac{\delta l}{\delta \Omega} = \mu, \quad g^{-1}\dot{g} = \Omega, \quad \Omega = \frac{\partial H}{\partial \mu}. \quad (\text{C.1.29})$$

These are Hamiltonian with the standard Lie-Poisson bracket defined on the dual Lie algebra  $so(3, \mathbb{C}^3)^*$ . Because of the properties of the trace norm, we may take  $\mu = \text{skew } \mathbb{I} \Omega$ . (Alternatively, we may set the preserved symmetric part of  $\mu$  initially to zero.) Hence,  $\mu$  may be taken as a skew-symmetric complex matrix, which again may be identified with the components of a complex three-vector  $\mathbf{z}$  as  $(\mu)_{jk} = -z^i \epsilon_{ijk}$ . On making this identification, Euler's Equations (C.1.15) emerge for  $\mathbf{z} \in \mathbb{C}^3$ , with real  $\mathbb{I}$ . The  $\mathcal{PT}$ -symmetric initial conditions on the real level sets of the preserved complex quantities  $C$  and  $H$  form an invariant manifold of this system of three complex ordinary differential equations. On this invariant manifold, the complex angular motion is completely integrable. By following the approach established by Manakov [Man1976] this reasoning may also extend to the rigid body on  $SO(n, \mathbb{C})$ . □

## 338 C : ENHANCED COURSEWORK

**Exercise.** The Bers left behind a toy monopole. This is a rigid body that rotates by complex angles and whose three moments of inertia are the complex cube roots of unity. What are the equations of motion for this toy monopole? For a hint, take a look at [Iv2006]. ★

### C.1.3 Rotations in four dimensions: $SO(4)$

**Scenario C.1.4** *The Tets are yet another alien life form who also use one-dimensional time  $t \in \mathbb{R}$  (we sigh with relief), but their spatial coordinates are  $\mathbf{X} \in \mathbb{R}^4$ , while ours are  $\mathbf{x} \in \mathbb{R}^3$ . They test us to determine whether we are an intelligent life form by requiring us to write the equations for rigid-body motion for four-dimensional rotations.*

*Hint: The angular velocity of rotation  $\widehat{\Psi} = O^{-1}\dot{O}(t)$  for rotations  $O(t) \in SO(4)$  in four dimensions will be represented by a  $4 \times 4$  skew-symmetric matrix. Write a basis for the  $4 \times 4$  skew-symmetric matrices by adding a row and column to the  $3 \times 3$  basis.*

**Answer.** Any  $4 \times 4$  skew-symmetric matrix may be represented as a linear combination of  $4 \times 4$  basis matrices with three-dimensional vector coefficients  $\Omega, \Lambda \in \mathbb{R}^3$  in the form

$$\begin{aligned} \widehat{\Psi} &= \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 & -\Lambda_1 \\ \Omega_3 & 0 & -\Omega_1 & -\Lambda_2 \\ -\Omega_2 & \Omega_1 & 0 & -\Lambda_3 \\ \Lambda_1 & \Lambda_2 & \Lambda_3 & 0 \end{pmatrix} \\ &= \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K} \\ &= \Omega_a \widehat{J}_a + \Lambda_b \widehat{K}_b. \end{aligned}$$

This is the formula for the angular velocity of rotation in four dimensions.

The  $4 \times 4$  basis set  $\widehat{J} = (J_1, J_2, J_3)^T$  and  $\widehat{K} = (K_1, K_2, K_3)^T$  consists of the following six linearly independent  $4 \times 4$  skew-symmetric matrices,  $\widehat{J}_a, \widehat{K}_b$  with  $a, b = 1, 2, 3$ :

$$\begin{aligned} \widehat{J}_1 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \widehat{K}_1 &= \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \\ \widehat{J}_2 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \widehat{K}_2 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \\ \widehat{J}_3 &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \widehat{K}_3 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \end{aligned}$$

The matrices  $\widehat{J}_a$  with  $a = 1, 2, 3$  embed the basis for  $3 \times 3$  skew-symmetric matrices into the  $4 \times 4$  matrices by adding a row and column of zeros. The skew matrices  $\widehat{K}_a$  with  $a = 1, 2, 3$  then extend the  $3 \times 3$  basis to  $4 \times 4$ .

### Commutation relations

The skew matrix basis  $\widehat{J}_a, \widehat{K}_b$  with  $a, b = 1, 2, 3$  satisfies the commutation relations,

$$\begin{aligned} [\widehat{J}_a, \widehat{J}_b] &= \widehat{J}_a \widehat{J}_b - \widehat{J}_b \widehat{J}_a = \epsilon_{abc} \widehat{J}_c, \\ [\widehat{J}_a, \widehat{K}_b] &= \widehat{J}_a \widehat{K}_b - \widehat{K}_b \widehat{J}_a = \epsilon_{abc} \widehat{K}_c, \\ [\widehat{K}_a, \widehat{K}_b] &= \widehat{K}_a \widehat{K}_b - \widehat{K}_b \widehat{K}_a = \epsilon_{abc} \widehat{J}_c. \end{aligned}$$

These commutation relations may be verified by a series of direct calculations, as  $[\widehat{J}_1, \widehat{J}_2] = \widehat{J}_3$ , etc.

## 340 C : ENHANCED COURSEWORK

**Hat map for  $4 \times 4$  skew matrices**

The map above for the  $4 \times 4$  skew matrix  $\widehat{\Psi}$  may be written as

$$\widehat{\Psi} = \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K} = \Omega_a \widehat{J}_a + \Lambda_b \widehat{K}_b, \text{ sum on } a, b = 1, 2, 3.$$

This map provides the  $4 \times 4$  version of the hat map, written now as  $(\cdot)^\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto so(4)$ . Here  $so(4)$  is the Lie algebra of the  $4 \times 4$  special orthogonal matrices, which consists of the  $4 \times 4$  skew matrices represented in the six-dimensional basis of  $\widehat{J}$ 's and  $\widehat{K}$ 's.

**Commutator as intertwined vector product**

The commutator of  $4 \times 4$  skew matrices corresponds to an intertwined vector product, as follows. For any vectors  $\Omega, \Lambda, \omega, \lambda \in \mathbb{R}^3$ , one has

$$\begin{aligned} & \left[ \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K}, \omega \cdot \widehat{J} + \lambda \cdot \widehat{K} \right] \\ &= \left( \Omega \times \omega + \Lambda \times \lambda \right) \cdot \widehat{J} + \left( \Omega \times \lambda - \Lambda \times \omega \right) \cdot \widehat{K}. \end{aligned}$$

Likewise, the matrix pairing  $\langle A, B \rangle = \text{tr}(A^T B)$  is related to the vector dot-product pairing in  $\mathbb{R}^3$  by

$$\left\langle \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K}, \omega \cdot \widehat{J} + \lambda \cdot \widehat{K} \right\rangle = \Omega \cdot \omega + \Lambda \cdot \lambda.$$

That is,

$$\left\langle \widehat{J}_a, \widehat{J}_b \right\rangle = \delta_{ab} = \left\langle \widehat{K}_a, \widehat{K}_b \right\rangle \quad \text{and} \quad \left\langle \widehat{J}_a, \widehat{K}_b \right\rangle = 0.$$

**Euler–Poincaré equation on  $so(4)$ \***

For

$$\Phi = O^{-1} \delta O(t) = \xi \cdot \widehat{J} + \eta \cdot \widehat{K} \in so(4),$$



Hamilton's principle  $\delta S = 0$  for  $S = \int_a^b \ell(\Psi) dt$  with

$$\Psi = O^{-1}\dot{O}(t) = \Omega \cdot \hat{J} + \Lambda \cdot \hat{K} \in so(4)$$

leads to

$$\delta S = \int_a^b \left\langle \frac{\delta \ell}{\delta \Psi}, \delta \Psi \right\rangle dt = \int_a^b \left\langle \frac{\delta \ell}{\delta \Psi}, \dot{\Phi} + \text{ad}_\Psi \Phi \right\rangle dt,$$

where

$$\begin{aligned} \text{ad}_\Psi \Phi &= [\Psi, \Phi] = [\Omega \cdot \hat{J} + \Lambda \cdot \hat{K}, \xi \cdot \hat{J} + \eta \cdot \hat{K}] \\ &= (\Omega \times \xi + \Lambda \times \eta) \cdot \hat{J} + (\Omega \times \eta - \Lambda \times \xi) \cdot \hat{K}. \end{aligned}$$

Thus,

$$\begin{aligned} \delta S &= \int_a^b \left\langle -\frac{d}{dt} \frac{\delta \ell}{\delta \Psi}, \Phi \right\rangle + \left\langle \frac{\delta \ell}{\delta \Psi}, \text{ad}_\Psi \Phi \right\rangle dt \\ &= \int_a^b \left\langle -\frac{d}{dt} \frac{\delta \ell}{\delta \Omega} \cdot \hat{J} - \frac{d}{dt} \frac{\delta \ell}{\delta \Lambda} \cdot \hat{K}, \xi \cdot \hat{J} + \eta \cdot \hat{K} \right\rangle dt \\ &\quad + \int_a^b \left\langle \frac{\delta \ell}{\delta \Omega} \cdot \hat{J} + \frac{\delta \ell}{\delta \Lambda} \cdot \hat{K}, \right. \\ &\quad \left. (\Omega \times \xi + \Lambda \times \eta) \cdot \hat{J} + (\Omega \times \eta - \Lambda \times \xi) \cdot \hat{K} \right\rangle dt \\ &= \int_a^b \left( -\frac{d}{dt} \frac{\delta \ell}{\delta \Omega} + \frac{\delta \ell}{\delta \Omega} \times \Omega - \frac{\delta \ell}{\delta \Lambda} \times \Lambda \right) \cdot \xi \\ &\quad + \left( -\frac{d}{dt} \frac{\delta \ell}{\delta \Lambda} + \frac{\delta \ell}{\delta \Lambda} \times \Omega + \frac{\delta \ell}{\delta \Omega} \times \Lambda \right) \cdot \eta dt. \end{aligned}$$

Hence,  $\delta S = 0$  yields

$$\begin{aligned} \frac{d}{dt} \frac{\delta \ell}{\delta \Omega} &= \frac{\delta \ell}{\delta \Omega} \times \Omega - \frac{\delta \ell}{\delta \Lambda} \times \Lambda \\ \text{and } \frac{d}{dt} \frac{\delta \ell}{\delta \Lambda} &= \frac{\delta \ell}{\delta \Lambda} \times \Omega + \frac{\delta \ell}{\delta \Omega} \times \Lambda. \quad (\text{C.1.30}) \end{aligned}$$

## 342 C : ENHANCED COURSEWORK

These are the  $\widehat{J}, \widehat{K}$  basis components of the Euler–Poincaré equation on  $so(4)^*$ ,

$$\frac{d}{dt} \frac{\delta \ell}{\delta \Psi} = \text{ad}_{\Psi}^* \frac{\delta \ell}{\delta \Psi},$$

written with  $\Psi = \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K}$  in this basis.

**Hamiltonian form on  $so(4)^*$** 

Legendre-transforming yields the pairs

$$\Pi = \frac{\delta \ell}{\delta \Omega}, \quad \Omega = \frac{\delta h}{\delta \Pi}, \quad \text{and} \quad \Xi = \frac{\delta \ell}{\delta \Lambda}, \quad \Lambda = \frac{\delta h}{\delta \Xi}.$$

Hence, these equations may be expressed in Hamiltonian form as

$$\frac{d}{dt} \begin{bmatrix} \Pi \\ \Xi \end{bmatrix} = \begin{bmatrix} \Pi \times & \Xi \times \\ \Xi \times & \Pi \times \end{bmatrix} \begin{bmatrix} \delta h / \delta \Pi \\ \delta h / \delta \Xi \end{bmatrix}. \quad (\text{C.1.31})$$

The corresponding Lie–Poisson bracket is given by

$$\begin{aligned} \{f, h\} &= -\Pi \cdot \left( \frac{\delta f}{\delta \Pi} \times \frac{\delta h}{\delta \Pi} + \frac{\delta f}{\delta \Xi} \times \frac{\delta h}{\delta \Xi} \right) \\ &\quad - \Xi \cdot \left( \frac{\delta f}{\delta \Pi} \times \frac{\delta h}{\delta \Xi} - \frac{\delta h}{\delta \Pi} \times \frac{\delta f}{\delta \Xi} \right). \end{aligned}$$

This Lie–Poisson bracket has an extra term proportional to  $\Pi$ , relative to the  $se(3)^*$  bracket (7.1.10) for the heavy top. Its Hamiltonian matrix has two null eigenvectors for the variational derivatives of  $C_1 = |\Pi|^2 + |\Xi|^2$  and  $C_2 = \Pi \cdot \Xi$ . The functions  $C_1, C_2$  are the Casimirs of the  $so(4)$  Lie–Poisson bracket. That is,  $\{C_1, H\} = 0 = \{C_2, H\}$  for every Hamiltonian  $H(\Pi, \Xi)$ .

The Hamiltonian matrix in Equation (C.1.31) is similar to that for the Lie–Poisson formulation of heavy-top dynamics, except for the one extra term  $\{\Xi, \Xi\} \neq 0$ .  $\blacktriangle$

## C.2 $\mathbb{C}^3$ oscillators

**Scenario C.2.1** For  $\mathbf{a} \in \mathbb{C}^3$  one may write the  $3 \times 3$  Hermitian matrix  $Q = \mathbf{a} \otimes \mathbf{a}^*$  as the sum  $Q = S + iA$  of a  $3 \times 3$  real symmetric matrix  $S$  plus  $i$  times a  $3 \times 3$  real antisymmetric matrix  $A$ :

$$\begin{aligned} Q &= \begin{bmatrix} M_1 & N_3 - iL_3 & N_2 + iL_2 \\ N_3 + iL_3 & M_2 & N_1 - iL_1 \\ N_2 - iL_2 & N_1 + iL_1 & M_3 \end{bmatrix} \\ &= \begin{bmatrix} M_1 & N_3 & N_2 \\ N_3 & M_2 & N_1 \\ N_2 & N_1 & M_3 \end{bmatrix} + i \begin{bmatrix} 0 & -L_3 & L_2 \\ L_3 & 0 & -L_1 \\ -L_2 & L_1 & 0 \end{bmatrix}. \end{aligned}$$

- (i) Compute the Poisson brackets of the  $L$ 's,  $M$ 's and  $N$ 's among themselves, given that  $\{a_j, a_k^*\} = -2i\delta_{jk}$  for  $j, k = 1, 2, 3$ .
- (ii) Transform into a rotating frame in which the real symmetric part of  $Q$  is diagonal. Write the Hamiltonian equations for the  $L$ 's,  $M$ 's and  $N$ 's in that rotating frame for a rotationally invariant Hamiltonian.

**Answer. (Oscillator variables in three dimensions)** The nine elements of  $Q$  are the  $S^1$ -invariants

$$Q_{jk} = a_j a_k^* = S_{jk} + iA_{jk}, \quad j, k = 1, 2, 3.$$

The Poisson brackets among these variables are evaluated from the canonical relation,

$$\{a_j, a_k^*\} = -2i\delta_{jk},$$

by using the Leibniz property (product rule) for Poisson brackets to find

$$\{Q_{jk}, Q_{lm}\} = 2i(\delta_{kl}Q_{jm} - \delta_{jm}Q_{kl}), \quad j, k, l, m = 1, 2, 3.$$

## 344 C : ENHANCED COURSEWORK

**Remark C.2.1** The quadratic  $S^1$ -invariant quantities in  $\mathbb{C}^3$  Poisson commute among themselves. This property of *closure* is to be expected for a simple reason. The Poisson bracket between two homogeneous polynomials of weights  $w_1$  and  $w_2$  produces a homogeneous polynomial of weight  $w = w_1 + w_2 - 2$  and  $2 + 2 - 2 = 2$ ; so the quadratic homogeneous polynomials Poisson-commute among themselves.

The result is also a simple example of Poisson reduction by symmetry, obtained by *transforming to quantities that are invariant* under the action of a Lie group. The action in this case is the (diagonal)  $S^1$  phase shift  $a_j \rightarrow a_j e^{i\phi}$  for  $j = 1, 2, 3$ .  $\square$

- (i) One defines  $L_a := -\frac{1}{2} \epsilon_{ajk} A_{jk} = (\mathbf{p} \times \mathbf{q})_a$  and finds the Poisson bracket relations,

$$\begin{aligned} \{L_a, L_b\} &= A_{ab} - A_{ba} = \epsilon_{abc} L_c, \\ \{L_a, Q_{jk}\} &= \frac{1}{2} [\epsilon_{ajc} Q_{ck} - \epsilon_{akc} Q_{jc}]. \end{aligned}$$

Thus, perhaps not unexpectedly, the Poisson bracket for quadratic  $S^1$ -invariant quantities in  $\mathbb{C}^3$  contains the angular momentum Poisson bracket among the variables  $L_a$  with  $a = 1, 2, 3$ . This could be expected, because the  $3 \times 3$  form of  $Q$  contains the  $2 \times 2$  form, which we know admits the Hopf fibration into quantities which satisfy Poisson bracket relations dual to the Lie algebra  $so(3) \simeq su(2)$ . Moreover, the imaginary part  $\text{Im } Q = \mathbf{L} \cdot \hat{\mathbf{J}} = L_a \hat{J}_a$ , where  $\hat{J}_a$  with  $a = 1, 2, 3$  is a basis set for  $so(3)$  as represented by the  $3 \times 3$  skew-symmetric real matrices.

Another interesting set of Poisson bracket relations among the  $M$ 's,  $N$ 's and  $L$ 's may be found. These

relations are

$$\begin{aligned}\{N_a - iL_a, N_b - iL_b\} &= 2i \epsilon_{abc}(N_c + iL_c), \\ \{M_a, M_b\} &= 0, \\ \{M_a, N_b - iL_b\} &= 2i \operatorname{sgn}(b - a)(-1)^{a+b} \\ &\quad (N_b - iL_b),\end{aligned}$$

where  $\operatorname{sgn}(b - a)$  is the sign of the difference  $(b - a)$ , which vanishes when  $b = a$ .

Additional Poisson bracket relations may also be read off from the Poisson commutators of the real and imaginary components of  $Q = S + iA$  among themselves as

$$\begin{aligned}\{S_{jk}, S_{lm}\} &= \delta_{jl}A_{mk} + \delta_{kl}A_{mj} - \delta_{jm}A_{kl} - \delta_{km}A_{jl}, \\ \{S_{jk}, A_{lm}\} &= \delta_{jl}S_{mk} + \delta_{kl}S_{mj} - \delta_{jm}S_{kl} - \delta_{km}S_{jl}, \\ \{A_{jk}, A_{lm}\} &= \delta_{jl}A_{mk} - \delta_{kl}A_{mj} + \delta_{jm}A_{kl} - \delta_{km}A_{jl}.\end{aligned}\tag{C.2.1}$$

These relations produce the following five tables of Poisson brackets in addition to  $\{M_a, M_b\} = 0$ :

$\{\cdot, \cdot\}$	$L_1$	$L_2$	$L_3$
$L_1$	0	$L_3$	$-L_2$
$L_2$	$-L_3$	0	$L_1$
$L_3$	$L_2$	$-L_1$	0

$\{\cdot, \cdot\}$	$N_1$	$N_2$	$N_3$
$N_1$	0	$-L_3$	$L_2$
$N_2$	$L_3$	0	$-L_1$
$N_3$	$-L_2$	$L_1$	0

$\{\cdot, \cdot\}$	$L_1$	$L_2$	$L_3$
$M_1$	0	$2N_2$	$-2N_3$
$M_2$	$-2N_1$	0	$2N_3$
$M_3$	$2N_1$	$-2N_2$	0

## 346 C : ENHANCED COURSEWORK

$\{\cdot, \cdot\}$	$N_1$	$N_2$	$N_3$
$M_1$	0	$-2L_2$	$2L_3$
$M_2$	$2L_1$	0	$-2L_3$
$M_3$	$-2L_1$	$2L_2$	0

$\{\cdot, \cdot\}$	$L_1$	$L_2$	$L_3$
$N_1$	$M_2 - M_3$	$-N_3$	$N_2$
$N_2$	$N_3$	$M_3 - M_1$	$-N_1$
$N_3$	$-N_2$	$N_1$	$M_1 - M_2$

As expected, the system is closed and it has the angular momentum Poisson bracket table as a closed subset. Next, we will come to understand that this is because the Lie algebra  $su(2)$  is a subalgebra of  $su(3)$ .

- (ii) The rotation group  $SO(3)$  is a subgroup of  $SU(3)$ . An element  $Q \in su(3)^*$  transforms under  $SO(3)$  by the coAdjoint action

$$\text{Ad}_R^* Q = R^{-1}QR = R^{-1}SR + iR^{-1}AR.$$

Choose  $R \in SO(3)$  so that  $R^{-1}SR = D = \text{diag}(d_1, d_2, d_3)$  is diagonal. (That is, rotate into principal axis coordinates for  $S$ .) The eigenvalues are unique up to their order, which one may fix as, say,  $d_1 \geq d_2 \geq d_3$ . While it diagonalises the symmetric part of  $Q$ , the rotation  $R$  takes the antisymmetric part from the spatial frame to the body frame, where  $S$  is diagonal. At the same time the spatial angular momentum matrix  $A$  is transformed to  $B = R^{-1}AR$ , which is the body angular momentum. Thus,

$$\text{Ad}_R^* Q = R^{-1}SR + iR^{-1}AR =: D + iB.$$

Define the body angular velocity  $\Omega = R^{-1}\dot{R} \in so(3)$ , which is left-invariant. The Hamiltonian

dynamical system obeys

$$\dot{Q} = \{Q, H(Q)\}.$$

For  $B = R^{-1}AR$ , this implies

$$\dot{B} + [\Omega, B] = R^{-1}\dot{A}R = R^{-1}\{A, H(Q)\}R.$$

However,  $H(Q)$  being rotationally symmetric means the spatial angular momentum  $A$  will be time-independent  $\dot{A} = \{A, H(Q)\} = 0$ . Hence,

$$\dot{B} + [\Omega, B] = 0.$$

Thus, the equation for the body angular momentum  $B$  is formally identical to Euler's equations for rigid-body motion. Physically, this represents conservation of spatial angular momentum, because of the rotational symmetry of the Hamiltonian.

Likewise, for  $D = R^{-1}SR$ , one finds

$$\dot{D} + [\Omega, D] = R^{-1}\dot{S}R = R^{-1}\{S, H(Q)\}R \neq 0.$$

The body angular momentum  $B$  satisfies Euler's rigid-body equations, but this body is not rigid! While the rotational degrees of freedom satisfy spatial angular momentum conservation, the shape of the body depends on the value of the Poisson bracket  $R^{-1}\{S, H(Q)\}R$  which is likely to be highly nontrivial! For example, the Hamiltonian  $H(Q)$  may be chosen to be a function of the following three rotationally invariant quantities:

$$\begin{aligned} \text{tr}(A^T A) &= \text{tr}(B^T B), \\ \text{tr}(A^T S A) &= \text{tr}(B^T D B), \\ \text{tr}(A^T S^2 A) &= \text{tr}(B^T D^2 B). \end{aligned}$$

Dependence of the Hamiltonian on these quantities will bring the complications of the Poisson bracket relations in (C.2.1) into the dynamics of the triaxial ellipsoidal shape represented by  $D$ .  $\blacktriangle$

348 C : ENHANCED COURSEWORK

**Remark C.2.2** The quantity

$$\tilde{Q} = \mathbf{a} \otimes \mathbf{a}^* - \frac{1}{3} \text{Id} |\mathbf{a}|^2 : \mathbb{C}^3 \mapsto su(3)^*$$

corresponds for the action of  $SU(3)$  on  $\mathbb{C}^3$  to the momentum map  $J : \mathbb{C}^2 \mapsto su(2)^*$  in Example 11.3.1 for the action of  $SU(2)$  on  $\mathbb{C}^2$ .  $\square$

---

### C.3 Momentum maps for $GL(n, \mathbb{R})$

**Scenario C.3.1** ( $GL(n, \mathbb{R})$  invariance) *Begin with the Lagrangian*

$$L = \frac{1}{2} \text{tr}(\dot{S} S^{-1} \dot{S} S^{-1}) + \frac{1}{2} \dot{\mathbf{q}}^T S^{-1} \dot{\mathbf{q}},$$

where  $S = S^T$  is an  $n \times n$  symmetric matrix and  $\mathbf{q} \in \mathbb{R}^n$  is an  $n$ -component column vector.

- (i) Legendre-transform to construct the corresponding Hamiltonian and canonical equations.
- (ii) Show that the system is invariant under the group action

$$\mathbf{q} \rightarrow A\mathbf{q} \quad \text{and} \quad S \rightarrow ASA^T,$$

for any constant invertible  $n \times n$  matrix,  $A$ .

- (iii) Compute the infinitesimal generator for this group action and construct its corresponding momentum map. Is this momentum map equivariant?
- (iv) Verify directly that this momentum map is a conserved  $n \times n$  matrix quantity by using the equations of motion.
- (v) Is this system completely integrable for any value of  $n > 2$ ?



**Answer. ( $GL(n, \mathbb{R})$  invariance)**

(i) Legendre-transform

$$P = \frac{\partial L}{\partial \dot{S}} = S^{-1} \dot{S} S^{-1} \quad \text{and} \quad \mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}} = S^{-1} \dot{\mathbf{q}}.$$

Thus,  $P = P^T$  is also a symmetric matrix. The Hamiltonian  $H(Q, P)$  and its canonical equations are

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}, S, P) &= \frac{1}{2} \text{tr}(PS \cdot PS) + \frac{1}{2} \mathbf{p} \cdot S \mathbf{p}, \\ \dot{S} = \frac{\partial H}{\partial P} &= SPS, \quad \dot{P} = -\frac{\partial H}{\partial S} = -\left(PSP + \frac{1}{2} \mathbf{p} \otimes \mathbf{p}\right), \\ \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} &= S \mathbf{p}, \quad \dot{\mathbf{p}} = \frac{\partial H}{\partial \mathbf{q}} = 0. \end{aligned}$$

(ii) Under the group action  $\mathbf{q} \rightarrow G\mathbf{q}$  and  $S \rightarrow GSG^T$  for any constant invertible  $n \times n$  matrix,  $G$ , one finds

$$\dot{S}S^{-1} \rightarrow G\dot{S}S^{-1}G^{-1} \quad \text{and} \quad \dot{\mathbf{q}} \cdot S^{-1}\dot{\mathbf{q}} \rightarrow \dot{\mathbf{q}} \cdot S^{-1}\dot{\mathbf{q}}.$$

Hence,  $L \rightarrow L$ . Likewise,  $P \rightarrow G^{-T}PG^{-1}$  so  $PS \rightarrow G^{-T}PSG^T$  and  $\mathbf{p} \rightarrow G^{-T}\mathbf{p}$  so that  $S\mathbf{p} \rightarrow GS\mathbf{p}$ . Hence,  $H \rightarrow H$ , as well; so both  $L$  and  $H$  for the system are invariant.

(iii) The infinitesimal actions for  $G(\epsilon) = Id + \epsilon A + O(\epsilon^2)$ , where  $A \in gl(n)$  are

$$\begin{aligned} X_A \mathbf{q} &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} G(\epsilon) \mathbf{q} = A \mathbf{q} \quad \text{and} \\ X_A S &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \left( G(\epsilon) S G(\epsilon)^T \right) = AS + SA^T. \end{aligned}$$

The defining relation for the corresponding momentum map yields

$$\begin{aligned} \langle J, A \rangle &= \langle (Q, P), X_A \rangle \\ &= \text{tr}(P^T X_A S) + \mathbf{p} \cdot X_A \mathbf{q} \\ &= \text{tr}(P^T (AS + SA^T)) + \mathbf{p} \cdot A \mathbf{q}. \end{aligned}$$

## 350 C : ENHANCED COURSEWORK

Hence,

$$\langle J, A \rangle := \text{tr}(JA^T) = \text{tr}\left((S(P + P^T) + \mathbf{q} \otimes \mathbf{p})A\right),$$

which implies

$$J = (P^T + P)S + \mathbf{p} \otimes \mathbf{q} = 2PS + \mathbf{p} \otimes \mathbf{q}.$$

This momentum map is a cotangent lift, so it is equivariant.

- (iv) Conservation of the momentum map is verified directly by

$$\begin{aligned} \dot{J} &= (2\dot{P}S + 2P\dot{S} + \mathbf{p} \otimes \dot{\mathbf{q}}) \\ &= -2PSPS - (\mathbf{p} \otimes \mathbf{p})S + 2PSPS + \mathbf{p} \otimes S\mathbf{p} = 0. \end{aligned}$$

- (v) Integrability would be a good question for further study in this problem. ▲

**Scenario C.3.2 (Ellipsoidal motions on  $GL(3, \mathbb{R})$ )** Choose the Lagrangian in three dimensions,

$$L = \frac{1}{2}\text{tr}(\dot{Q}^T \dot{Q}) - V(\text{tr}(Q^T Q), \det(Q)),$$

where  $Q(t) \in GL(3, \mathbb{R})$  is a  $3 \times 3$  matrix function of time and the potential energy  $V$  is an arbitrary function of  $\text{tr}(Q^T Q)$  and  $\det(Q)$ .

- (i) Legendre-transform this Lagrangian. That is, find the momenta  $P_{ij}$  canonically conjugate to  $Q_{ij}$ . Then construct the Hamiltonian  $H(Q, P)$  and write Hamilton's canonical equations of motion for this problem.
- (ii) Show that the Hamiltonian is invariant under left action  $Q \rightarrow UQ$  where  $U \in SO(3)$ . Construct the cotangent lift of this action on  $P$ . Hence, construct the momentum map of left action.
- (iii) Show that the Hamiltonian is also invariant under right action  $Q \rightarrow QU$  where  $U \in SO(3)$ . Construct the cotangent lift of this action on  $P$ . That is, construct the momentum map of right action.

C.3 MOMENTUM MAPS FOR  $GL(n, \mathbb{R})$  351

(iv) Find the Poisson bracket relations among the various Hamiltonians for the two types of momentum maps. Explain your results in terms of the Lie algebra homomorphism that preserves bracket relations for equivariant momentum maps.

**Answer. (Ellipsoidal motions on  $GL(3, \mathbb{R})$ )**

(i) Legendre-transform

$$P_{ij} = \frac{\partial L}{\partial \dot{Q}_{ij}^T} = \dot{Q}_{ij}.$$

Thus, the Hamiltonian  $H(Q, P)$  and its canonical equations are

$$\begin{aligned} H(Q, P) &= \frac{1}{2} \text{tr}(P^T P) + V(\text{tr}(Q^T Q), \det(Q)), \\ \dot{Q}_{ij} &= \frac{\partial H}{\partial P_{ij}^T} = P_{ij}, \\ \dot{P}_{ij} &= -\frac{\partial H}{\partial Q_{ij}^T} \\ &= -\left( \frac{\partial V}{\partial \text{tr}(Q^T Q)} 2Q_{ij} \right. \\ &\quad \left. + \frac{\partial V}{\partial \det(Q)} (\det Q) (Q^{-1})_{ij} \right) \end{aligned}$$

where one uses the identity  $d(\det Q) = (\det Q) \text{tr}(Q^{-1} dQ^T)$ . The corresponding canonical Poisson bracket is

$$\begin{aligned} \{F, H\} &= \frac{\partial F}{\partial Q_{ij}} \frac{\partial H}{\partial P_{ij}} - \frac{\partial F}{\partial P_{ij}} \frac{\partial H}{\partial Q_{ij}} \\ &= \text{tr} \left( \frac{\partial F}{\partial Q} \frac{\partial H}{\partial P^T} - \frac{\partial F}{\partial P} \frac{\partial H}{\partial Q^T} \right). \quad (\text{C.3.1}) \end{aligned}$$

(ii) Under  $Q \rightarrow UQ$ ,  $\dot{Q} \rightarrow U\dot{Q}$ , one finds  $P \rightarrow UP$ . For  $U \in SO(3)$ , this means

$$\begin{aligned} \det Q &\rightarrow \det UQ = \det U \det Q = \det Q, \\ Q^T Q &\rightarrow Q^T U^T U Q = Q^T Q, \quad \text{and} \\ P^T P &\rightarrow P^T U^T U P = P^T P. \end{aligned}$$

## 352 C : ENHANCED COURSEWORK

Hence,  $H(Q, P)$  is left-invariant under  $Q \rightarrow UQ$  for  $U \in SO(3)$ .

The infinitesimal generator of left action  $X_A^L$  is found from

$$\begin{aligned} X_A^L Q &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} U(\epsilon)Q \\ &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} (Id + \epsilon A + O(\epsilon^2))Q = AQ, \end{aligned}$$

with  $A \in so(3)$  a  $3 \times 3$  antisymmetric matrix; that is,  $A^T = -A$ .

Likewise,  $P \rightarrow UP$  with  $U \in SO(3)$  has infinitesimal action  $X_A^L P = AP$ ,  $A \in so(3)$ . The action

$$\begin{aligned} X_A^L Q &= AQ = \frac{\partial \langle J_L, A \rangle}{\partial P^T} \quad \text{and} \\ X_A^L P &= AP = -\frac{\partial \langle J_L, A \rangle}{\partial Q^T} \end{aligned}$$

is generated canonically by the Hamiltonian

$$\begin{aligned} \langle J_L, A \rangle &= -\text{tr}(J_L^T A) = \text{tr}(P^T A Q) \\ &= \text{tr}(Q P^T A) = -\text{tr}(P Q^T A). \end{aligned}$$

Hence, the momentum map for left action is

$$J_L = [QP^T] = -[PQ^T],$$

where the square brackets mean taking the anti-symmetric part.

- (iii) Likewise,  $H(Q, P)$  is right-invariant under  $(Q, P) \rightarrow (QU, PU)$  for  $U \in SO(3)$ . The corresponding infinitesimal generator is found from

$$\begin{aligned} X_A^R Q &= QA = \frac{\partial \langle J_R, A \rangle}{\partial P^T} \quad \text{and} \\ X_A^R P &= PA = -\frac{\partial \langle J_R, A \rangle}{\partial Q^T} \end{aligned}$$

which is generated by the Hamiltonian

$$\begin{aligned}\langle J_R, A \rangle &= -\operatorname{tr}(J_R^T A) = \operatorname{tr}(J_R A) \\ &= \operatorname{tr}(P^T Q A) = -\operatorname{tr}(Q^T P A).\end{aligned}$$

Hence, the momentum map for right action is

$$J_R = [P^T Q] = -[Q^T P].$$

Note that  $J_R = Q^{-1} J_L Q = \operatorname{Ad}_{Q^{-1}} J_L$ .

- (iv) Define Hamiltonians for the momentum maps as the functions

$$J_L^A(Q, P) := \langle J_L, A \rangle, \quad J_R^B(Q, P) := \langle J_R, B \rangle, \quad \text{etc.}$$

Then by using the canonical Poisson bracket (C.3.1) on  $SO(3)$  and antisymmetry of the matrices  $A, B \in so(3)$  we find

$$\begin{aligned}\{J_L^A, J_L^B\} &= J_L^{[A, B]}, \\ \{J_R^A, J_R^B\} &= -J_R^{[A, B]}, \\ \{J_L^A, J_R^B\} &= 0.\end{aligned}$$

The Poisson brackets for the left (resp., right) actions give plus (resp., minus) representations of the Lie algebra, which is  $so(3)$  in this case. This result is expected, because cotangent-lift momentum maps are equivariant, and thus are infinitesimally equivariant, which in turn implies the Lie algebra homomorphism (11.2.4) that preserves bracket relations. Likewise, since left and right actions commute, the two momentum maps commute and their Hamiltonians Poisson-commute. ▲

### C.4 Motion on the symplectic Lie group $Sp(2)$

Let the set of  $2 \times 2$  matrices  $M_i$  with  $i = 1, 2, 3$  satisfy the defining relation for the symplectic Lie group  $Sp(2)$ ,

$$M_i J M_i^T = J \quad \text{with} \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The corresponding elements of its Lie algebra  $m_i = \dot{M}_i M_i^{-1} \in sp(2)$  satisfy  $(J m_i)^T = J m_i$  for each  $i = 1, 2, 3$ . Thus,  $X_i = J m_i$  satisfying  $X_i^T = X_i$  is a set of three symmetric  $2 \times 2$  matrices. For definiteness, we may choose a basis given by

$$X_1 = J m_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad X_2 = J m_2 = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \quad X_3 = J m_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This basis corresponds to the momentum map  $\mathbb{R}^6 \rightarrow sp(2)^*$  of quadratic phase-space functions  $\mathbf{X} = (|\mathbf{q}|^2, |\mathbf{p}|^2, \mathbf{q} \cdot \mathbf{p})^T$ . One sees this by using the symmetric matrices  $X_1, X_2, X_3$  above to compute the following three quadratic forms defined using  $\mathbf{z} = (\mathbf{q}, \mathbf{p})^T$ :

$$\frac{1}{2} \mathbf{z}^T X_1 \mathbf{z} = |\mathbf{q}|^2 = X_1, \quad \frac{1}{2} \mathbf{z}^T X_2 \mathbf{z} = |\mathbf{p}|^2 = X_2, \quad \frac{1}{2} \mathbf{z}^T X_3 \mathbf{z} = \mathbf{q} \cdot \mathbf{p} = X_3.$$

**Exercise. (The Lie bracket)** For  $X = Jm$  and  $Y = Jn \in sym(2)$  with  $m, n \in sp(2)$ , prove

$$[X, Y]_J := XJY - YJX = -J(mn - nm) = -J[m, n].$$

Use this equality to show that the  $J$ -bracket  $[X, Y]_J$  satisfies the Jacobi identity. ★

**Answer.** The first part is a straightforward calculation using  $J^2 = -\text{Id}_{2 \times 2}$  with the definitions of  $X$  and  $Y$ .

The second part follows from the Jacobi identity for the symplectic Lie algebra and linearity in the definitions of  $X, Y \in \mathfrak{sym}(2)$  in terms of  $m, n \in \mathfrak{sp}(2)$ . ▲

**Exercise. (A variational identity)** If  $X = J\dot{M}M^{-1}$  for derivative  $\dot{M} = \partial M(s, \sigma)/\partial s|_{\sigma=0}$  and  $Y = JM'M^{-1}$  for variational derivative  $\delta M = M' = \partial M(s, \sigma)/\partial \sigma|_{\sigma=0}$ , show that equality of cross derivatives in  $s$  and  $\sigma$  implies the relation

$$\delta X = X' = \dot{Y} + [X, Y]_J. \quad \star$$

**Answer.** This relation follows from an important standard calculation in geometric mechanics, performed earlier in deriving Equation (9.1.4). It begins by computing the time derivative of  $MM^{-1} = Id$  along the curve  $M(s)$  to find  $(MM^{-1})' = 0$ , so that

$$(M^{-1})' = -M^{-1}\dot{M}M^{-1}.$$

Next, one defines  $m = \dot{M}M^{-1}$  and  $n = M'M^{-1}$ . Then the previous relation yields

$$\begin{aligned} m' &= \dot{M}'M^{-1} - \dot{M}M^{-1}M'M^{-1} \\ \dot{n} &= \dot{M}'M^{-1} - M'M^{-1}\dot{M}M^{-1} \end{aligned}$$

so that subtraction yields the relation

$$m' - \dot{n} = nm - mn =: -[m, n].$$

Then, upon substituting the definitions of  $X$  and  $Y$ , one finds

$$\begin{aligned} X' = Jm' &= J\dot{n} - J[m, n] \\ &= \dot{Y} + [X, Y]_J = \dot{Y} + 2\text{sym}(XJY). \end{aligned} \quad \blacktriangle$$

**Exercise. (Hamilton's principle for  $sp(2)$ )** Use the previous relation to compute the Euler–Poincaré equation for evolution resulting from Hamilton's principle,

$$0 = \delta S = \delta \int \ell(X(s)) ds = \int \text{tr} \left( \frac{\partial \ell}{\partial X} \delta X \right) ds.$$

★

**Answer.** Integrate by parts and rearrange as follows:

$$\begin{aligned} 0 = \delta S &= \int \text{tr} \left( \frac{\partial \ell}{\partial X} X' \right) ds \\ &= \int \text{tr} \left( \frac{\partial \ell}{\partial X} (\dot{Y} - YJX + XJY) \right) ds \\ &= \int \text{tr} \left( \left( -\frac{d}{ds} \frac{\partial \ell}{\partial X} - JX \frac{\partial \ell}{\partial X} + \frac{\partial \ell}{\partial X} XJ \right) Y \right) ds \\ &= \int \text{tr} \left( \left( -\frac{d}{ds} \frac{\partial \ell}{\partial X} - 2\text{sym} \left( JX \frac{\partial \ell}{\partial X} \right) \right) Y \right) ds, \end{aligned}$$

upon setting the boundary term  $\text{tr} \left( \frac{\partial \ell}{\partial X} Y \right) \Big|_{s_0}^{s_1}$  equal to zero. This results in the Euler–Poincaré equation,

$$\frac{d}{ds} \frac{\partial \ell}{\partial X} = -2\text{sym} \left( JX \frac{\partial \ell}{\partial X} \right) = 2\text{sym} \left( \frac{\partial \ell}{\partial X} XJ \right). \quad (\text{C.4.1})$$

▲

**Exercise. (Geodesic motion on  $sp(2)^*$ )** Specialise this evolution equation to the case that  $\ell(X) = \frac{1}{2} \text{tr}(X^2)$ , where  $\text{tr}$  denotes the trace of a matrix. (This is geodesic motion on the matrix Lie group  $Sp(2)$  with respect to the trace norm of matrices.) ★



**Answer.** When  $\ell(X) = \frac{1}{2}\text{tr}(X^2)$  we have  $\partial\ell/\partial X = X$ , so the Euler–Poincaré Equation (C.4.1) becomes

$$\dot{X} = -2\text{sym}(JX^2) = X^2J - JX^2 = [X^2, J]. \quad (\text{C.4.2})$$

This is called a *Bloch–Iserles equation* [BIs2006]. ▲

**Exercise. (Lie–Poisson Hamiltonian formulation)** Write the Hamiltonian form of the Euler–Poincaré equation on  $SP(2)$  and identify the associated Lie–Poisson bracket. ★

**Answer.** The Hamiltonian form of the Euler–Poincaré Equation (C.4.1) is found from the Legendre transform via the dual relations

$$\mu = \frac{\partial\ell}{\partial X} \quad \text{and} \quad X = \frac{\partial h}{\partial\mu} \quad \text{with} \quad h(\mu) = \text{tr}(\mu X) - \ell(X).$$

Thus,

$$\dot{\mu} = -2\text{sym}\left(J\frac{\partial h}{\partial\mu}\mu\right) = -J\frac{\partial h}{\partial\mu}\mu + \mu\frac{\partial h}{\partial\mu}J.$$

The Lie–Poisson bracket is obtained from

$$\begin{aligned} \frac{d}{ds}f(\mu) &= \text{tr}\left(\frac{\partial f}{\partial\mu}\frac{d\mu}{ds}\right) \\ &= -2\text{tr}\left(\mu\text{sym}\left(\frac{\partial f}{\partial\mu}J\frac{\partial h}{\partial\mu}\right)\right) \\ &= -\text{tr}\left(\mu\left[\frac{\partial f}{\partial\mu}, \frac{\partial h}{\partial\mu}\right]_J\right) \\ &=: \{f, h\}_J. \end{aligned}$$

The Jacobi identity for this Lie–Poisson bracket follows from that of the  $J$ -bracket discussed earlier.

The geodesic Bloch–Iserles Equation (C.4.2) is recovered when the Hamiltonian is chosen as  $h = \frac{1}{2}\text{tr}(\mu^2)$  and one sets  $\mu \rightarrow X$ . ▲

**Exercise. (A second Bloch–Iserles Poisson bracket)** Show that the geodesic Bloch–Iserles Equation (C.4.2) may also be written in Hamiltonian form with Hamiltonian  $h = \frac{1}{3}\text{tr}(\mu^3)$ . ★

**Answer.** Equation (C.4.2) may also be written as

$$\dot{\mu} = -2\text{sym}\left(J\frac{\partial h}{\partial\mu}\mu\right) = -J\frac{\partial h}{\partial\mu} + \frac{\partial h}{\partial\mu}J$$

with Hamiltonian  $h = \frac{1}{3}\text{tr}(\mu^3)$ . The corresponding Poisson bracket has constant coefficients,

$$\begin{aligned} \frac{d}{ds}f(\mu) &= \text{tr}\left(\frac{\partial f}{\partial\mu}\frac{d\mu}{ds}\right) \\ &= -2\text{tr}\left(\text{sym}\left(\frac{\partial f}{\partial\mu}J\frac{\partial h}{\partial\mu}\right)\right) \\ &= -\text{tr}\left(\left[\frac{\partial f}{\partial\mu}, \frac{\partial h}{\partial\mu}\right]_J\right) \\ &=: \{f, h\}_{J_2}. \end{aligned}$$



**Exercise. (A parallel with the rigid body)** The geodesic Bloch–Iserles Equation (C.4.2) may be written in a form reminiscent of the rigid body, as

$$\frac{d}{dt}X = [X, \Omega] \quad \text{with} \quad \Omega = JX + XJ = -\Omega^T.$$

This suggests the Manakov form

$$\frac{d}{dt}(X + \lambda J) = [X + \lambda J, JX + XJ + \lambda^2 J^2].$$

This seems dual to the Manakov form (2.4.28) for the rigid body, because the symmetric and antisymmetric matrices exchange roles.

Verify these equations and explain what the Manakov form means in determining the conservation laws for this problem. ★

**Exercise. (The Bloch–Iserles  $G$ -strand)** Refer to Chapter 10 and compute the  $G$ -strand equations for  $G = Sp(2)$ . ★

## C.5 Two coupled rigid bodies

### *Formulation of the problem*

In the centre of mass frame, the Lagrangian for the problem of two coupled rigid bodies may be written as depending only on the angular velocities of the two bodies  $\Omega_1 = A_1^{-1}\dot{A}_1(t)$ ,  $\Omega_2 = A_2^{-1}\dot{A}_2(t)$  and the relative angle  $A = A_1^{-1}A_2$  between the bodies [GrKrMa1988],

$$l(\Omega_1, \Omega_2, A) : \mathfrak{so}(3) \times \mathfrak{so}(3) \times SO(3) \rightarrow \mathbb{R},$$

which we write as

$$l(\Omega_1, \Omega_2, A) = \frac{1}{2} \begin{pmatrix} \Omega_1 \\ \Omega_2 \end{pmatrix}^T \cdot M(A) \begin{pmatrix} \Omega_1 \\ \Omega_2 \end{pmatrix},$$

where  $M(A)$  is a  $6 \times 6$  block matrix containing both  $A$  and the two inertia tensors of the bodies.

Upon identifying  $\mathbb{R}_3$  with  $\mathfrak{so}(3) = T_e SO(3)$  by the hat map, this Lagrangian becomes

$$l = l(\widehat{\Omega}_1, \widehat{\Omega}_2, A)$$

## 360 C : ENHANCED COURSEWORK

and we may identify  $SO(3)$  with its dual  $SO^*(3)$  through the matrix pairing  $SO(3) \times SO^*(3) \rightarrow \mathbb{R}$ .

The Lagrangian is then a function

$$l : \mathfrak{so}(3) \times \mathfrak{so}(3) \times SO^*(3) \rightarrow \mathbb{R}$$

which may be written as

$$l(\Omega, A) = \frac{1}{2} \langle M(A) \Omega, \Omega \rangle =: \frac{1}{2} \langle \Pi, \Omega \rangle,$$

where a nondegenerate matrix trace pairing is defined in components by

$$\langle \Pi, \Omega \rangle := \text{Tr} \left[ \begin{pmatrix} \hat{\Omega}_1 \\ \hat{\Omega}_2 \end{pmatrix}^T \cdot \begin{pmatrix} \tilde{\Pi}_1 \\ \tilde{\Pi}_2 \end{pmatrix} \right]$$

for all  $\Omega = (\hat{\Omega}_1, \hat{\Omega}_2) \in \mathfrak{so}(3) \times \mathfrak{so}(3)$ ,  $\Pi = (\tilde{\Pi}_1, \tilde{\Pi}_2) \in \mathfrak{so}^*(3) \times \mathfrak{so}^*(3)$ .

The Euler–Poincaré theory has been developed to treat Lagrangians of the form

$$l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$$

where  $V$  is a vector space on which the Lie algebra acts. This Appendix deals with the following.

**Problem C.5.1** *Formulate the Euler–Poincaré equations for the problem of two coupled rigid bodies.*

### Problem Solution

The direct-product Lie algebra

$$\mathfrak{g} = \mathfrak{so}(3) \times \mathfrak{so}(3)$$

is endowed with the product Lie bracket

$$\text{ad}_{\Omega} \Xi = [\Omega, \Xi] = \left[ \begin{pmatrix} \hat{\Omega}_1 \\ \hat{\Omega}_2 \end{pmatrix}, \begin{pmatrix} \hat{\Xi}_1 \\ \hat{\Xi}_2 \end{pmatrix} \right] = \begin{pmatrix} \llbracket \hat{\Omega}_1, \hat{\Xi}_1 \rrbracket \\ \llbracket \hat{\Omega}_2, \hat{\Xi}_2 \rrbracket \end{pmatrix} \quad (\text{C.5.1})$$

where  $[\cdot, \cdot]$  indicates the standard  $\mathfrak{so}(3)$  matrix commutator.

Formulating the Euler–Poincaré theorem for this problem will require a Lie algebra action of  $\mathfrak{so}(3) \times \mathfrak{so}(3)$  on  $SO^*(3)$ , which fortunately is readily available. Indeed, from the definitions of the two body angular velocities and relative angle  $A = A_1^{-1}A_2$ , one finds

$$\frac{dA}{dt} = -\widehat{\Omega}_1 A + A \widehat{\Omega}_2, \quad (\text{C.5.2})$$

which is the Lie algebra action we seek, abbreviated as

$$\frac{dA}{dt} = -\Omega(A). \quad (\text{C.5.3})$$

The Euler–Poincaré variational principle is then  $\delta S = 0$ , for

$$\begin{aligned} \delta \int_{t_0}^{t_1} l(\Omega, A) dt &= \int_{t_0}^{t_1} \left\langle \frac{\delta l}{\delta \Omega}, \delta \Omega \right\rangle + \left\langle \frac{\delta l}{\delta A}, \delta A \right\rangle dt \\ &= \int_{t_0}^{t_1} \left\langle \frac{\delta l}{\delta \Omega}, \frac{d\Xi}{dt} + \text{ad}_\Omega \Xi \right\rangle + \left\langle \frac{\delta l}{\delta A}, -\Xi(A) \right\rangle dt \\ &= \int_{t_0}^{t_1} \left\langle -\frac{d}{dt} \frac{\delta l}{\delta \Omega} + \text{ad}_\Omega^* \frac{\delta l}{\delta \Omega} + \frac{\delta l}{\delta A} \diamond A, \Xi \right\rangle dt \end{aligned}$$

with  $\delta \Omega = \dot{\Xi} + \text{ad}_\Omega \Xi$  and  $\delta A = -\Xi(A)$ . As a result, the (left-invariant) Euler–Poincaré equations may be written as

$$\frac{d}{dt} \frac{\delta l}{\delta \Omega} = \text{ad}_\Omega^* \frac{\delta l}{\delta \Omega} + \frac{\delta l}{\delta A} \diamond A. \quad (\text{C.5.4})$$

This, of course, is the general form of the Euler–Poincaré equations with advected quantities.

The Euler–Poincaré equations for the present problem of coupled rigid bodies will take their final form, once we have computed the diamond operation ( $\diamond$ ),

$$\diamond : SO(3) \times SO(3)^* \rightarrow \mathfrak{so}(3)^*. \quad (\text{C.5.5})$$

## 362 C : ENHANCED COURSEWORK

The Lie algebra action (C.5.3) yields the following definition of diamond for our case,

$$\begin{aligned} \left\langle \frac{\delta l}{\delta A} \diamond A, \Xi \right\rangle &:= - \left\langle A, \Xi \left( \frac{\delta l}{\delta A} \right) \right\rangle \\ &= - \left\langle A, \left( \widehat{\Xi}_1 \frac{\delta l}{\delta A} - \frac{\delta l}{\delta A} \widehat{\Xi}_2 \right) \right\rangle \\ &= - \left\langle A \frac{\delta l}{\delta A}, \widehat{\Xi}_1 \right\rangle + \left\langle \frac{\delta l}{\delta A} A, \widehat{\Xi}_2 \right\rangle, \end{aligned}$$

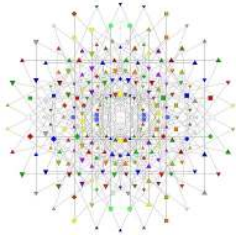
where the last step is justified by the cyclic property of the trace. Consequently, the components of the diamond operation are given by

$$\frac{\delta l}{\delta A} \diamond A = \left( -A \frac{\delta l}{\delta A}, \frac{\delta l}{\delta A} A \right)$$

and substituting them into the general form of the Euler–Poincaré equations in (C.5.4) gives the equations of motion of our problem:

$$\begin{aligned} \frac{d}{dt} \widetilde{\Pi}_1 &= \text{ad}_{\widehat{\Omega}_1}^* \widetilde{\Pi}_1 - A \frac{\delta l}{\delta A}, \\ \frac{d}{dt} \widetilde{\Pi}_2 &= \text{ad}_{\widehat{\Omega}_2}^* \widetilde{\Pi}_2 + \frac{\delta l}{\delta A} A. \end{aligned}$$

These along with the auxiliary Equation (C.5.2) comprise the Euler–Poincaré form of the Lie–Poisson equations that are derived for the motion of two coupled rigid bodies in [GrKrMa1988]. The corresponding Lie–Poisson equations in [GrKrMa1988] may be derived from the Euler–Poincaré equations here by applying a symmetry-reduced Legendre transform.



# D

## POINCARÉ'S 1901 PAPER

### H. Poincaré (1901)

**Sur une forme nouvelle des équations de la mécanique.**<sup>1</sup>  
*C.R. Acad. Sci.* **132**, 369-371.

Having had the opportunity to work on the rotational motion of hollow solid bodies filled with liquid, I have been led to cast the equations of mechanics into a new form that could be interesting to know. Assume there are  $n$  degrees of freedom and let  $\{x^1, \dots, x^n\}$  be the variables describing the state of the system. Let  $T$  and  $U$  be the kinetic and potential energy of the system.

Consider any continuous, transitive group (that is, its action covers the entire manifold). Let  $X_i(f)$  be any infinitesimal transformation of this group such that<sup>2</sup>

$$X_i(f) = \sum_{\mu=1}^n X_i(x^\mu) \frac{\partial f}{\partial x^\mu} = X_i^1 \frac{\partial f}{\partial x^1} + X_i^2 \frac{\partial f}{\partial x^2} + \dots + X_i^n \frac{\partial f}{\partial x^n}.$$

Since these transformations form a group, we must have

$$X_i X_k - X_k X_i = \sum_{s=1}^r c_{ik}^s X_s.$$

<sup>1</sup>Translation of [Po1901] into English by D. D. Holm and J. Kirsten

<sup>2</sup>For the finite dimensional case considered here, the  $\{X_i\}$  may be regarded as a set of  $r$  constant  $n \times n$  matrices that act linearly on the set of states  $\{x\}$ . Then, for example,  $X_i(x^\mu) = \sum_{\nu=1}^n [X_i]_{\nu}^{\mu} x^\nu$ . (Translator's note)

## 364 D : POINCARÉ'S 1901 PAPER

Since the group is transitive we can write

$$\dot{x}^\mu(t) = \frac{dx^\mu}{dt} = \sum_{i=1}^r \eta^i(t) X_i(x^\mu) = \eta^1(t) X_1^\mu + \eta^2(t) X_2^\mu + \cdots + \eta^r(t) X_r^\mu,$$

in such a way that we can go from the state  $(x^1, \dots, x^n)$  of the system to a state  $(x^1 + \dot{x}^1 dt, \dots, x^n + \dot{x}^n dt)$  by using the infinitesimal transformation of the group,  $\sum_{i=1}^r \eta^i X_i(f)$ .

$T$  instead of being expressed as a function of the  $x$  and  $\dot{x}$  can be written as a function of the  $\eta$  and  $x$ . If we increase the  $\eta$  and  $x$  by virtual displacements  $\delta\eta$  and  $\delta x$ , respectively, there will be resulting increases in  $T$  and  $U$

$$\delta T = \sum \frac{\delta T}{\delta \eta} \delta \eta + \sum \frac{\delta T}{\delta x} \delta x \quad \text{and} \quad \delta U = \sum \frac{\delta U}{\delta x} \delta x.$$

Since the group is transitive I will be able to write

$$\delta x^\mu = \omega^1 X_1^\mu + \omega^2 X_2^\mu + \cdots + \omega^r X_r^\mu,$$

in such a way that we can go from the state  $x^\mu$  of the system to the state  $x^\mu + \delta x^\mu$  by using the infinitesimal transformation of the group  $\delta x^\mu = \sum_{i=1}^r \omega^i X_i(x^\mu)$ . I will then write<sup>3</sup>

$$\delta T - \delta U = \sum_{i=1}^r \frac{\delta T}{\delta \eta^i} \delta \eta^i + \sum_{\mu=1}^n \left( \frac{\delta T}{\delta x^\mu} - \frac{\delta U}{\delta x^\mu} \right) \delta x^\mu = \sum_{i=1}^r \frac{\delta T}{\delta \eta^i} \delta \eta^i + \sum_{i=1}^r \Omega_i \omega^i.$$

Next, let the Hamilton integral be

$$J = \int (T - U) dt,$$

so we will have

$$\delta J = \int \left( \sum \frac{\delta T}{\delta \eta^i} \delta \eta^i + \sum \Omega_i \omega^i \right) dt,$$

<sup>3</sup>Here Poincaré's formula reveals that  $\Omega_i = \sum_{\mu, \nu=1}^n \frac{\delta L}{\delta x^\mu} [X_i]^\mu_\nu x^\nu$  with  $L = T - U$ , or equivalently  $\Omega = \frac{\partial L}{\partial x} \diamond x$  in the notation of Chapter 6. (Translator's note)



and can easily find

$$\delta\eta^i = \frac{d\omega^i}{dt} + \sum_{s,k=1}^r c_{sk}{}^i \eta^k \omega^s.$$

The principle of stationary action then gives

$$\frac{d}{dt} \frac{\delta T}{\delta \eta^s} = \sum c_{sk}{}^i \frac{\delta T}{\delta \eta^i} \eta^k + \Omega_s.$$

These equations encompass some particular cases:

1. The Lagrange equations, when the group is reduced to the transformations, all commuting amongst each other, which each shift one of the variables  $x$  by an infinitesimally small constant.
2. Also, Euler's equations for solid body rotations emerge, in which the role of the  $\eta_i$  is played by the components  $p, q, r$  of the rotations and the role of  $\Omega$  by the coupled external forces.

These equations will be of special interest where the potential  $U$  is zero and the kinetic energy  $T$  only depends on the  $\eta$  in which case  $\Omega$  vanishes.



# Bibliography

---

- [AbMa1978] Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics, 2nd ed.* Reading, MA: Addison-Wesley.
- [AbMaRa1988] Abraham, R., Marsden, J. E. and Ratiu, T. S. [1988] *Manifolds, Tensor Analysis, and Applications, 2nd Ed.* Applied Mathematical Sciences, Vol. 75. New York: Springer.
- [AcHoKoTi1997] Aceves, A., Holm, D. D., Kovacic, G. and Timofeyev, I. [1997] Homoclinic orbits and chaos in a second-harmonic generating optical cavity. *Phys. Lett. A* **233**, 203–208.
- [AlluMaRo1998] Alber, M. S., Luther, G. G., Marsden, J. E. and Robbins, J. M. [1998] Geometric phases, reduction and Lie–Poisson structure for the resonant three-wave interaction. *Physica D* **123**, 271–290.
- [AlluMaRo1999] Alber, M. S., Luther, G. G., Marsden, J. E. and Robbins, J. M. [1999] Geometry and control of three-wave interactions. In *The Arnoldfest (Toronto, ON, 1997)*, *Fields Inst. Commun.* **24**, 55–80.
- [AlEb1975] Allen, L. and Eberly, J. H. [1975] *Optical Resonance and Two-Level Atoms.* New York: Wiley.
- [AmCZ1996] Ambrosetti, A. and Coti Zelati, V. [1996] *Periodic Solutions of Singular Lagrangian Systems.* Boston: Birkhäuser.
- [Ap1900] Appel, P. [1900] Sur l’intégration des équations du mouvement d’un corps pesant de révolution roulant par une arête circulaire sur un plan horizontal; cas particulier du cerceau. *Rendiconti del circolo matematico di Palermo* **14**, 1–6.

## 368 BIBLIOGRAPHY

- [Ar350BC] Aristotle [350BC] *Physics*.
- [Ar1966] Arnold V. I. [1966] Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier, Grenoble* **16**, 319–361.
- [Ar1993] Arnold V. I. [1993] *Dynamical Systems III*. New York: Springer.
- [Ar1979] Arnold V. I. [1979] *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics, Vol. 60. New York: Springer.
- [ArKh1998] Arnold, V. I. and Khesin, B. A. [1998] *Topological Methods in Hydrodynamics*. New York: Springer.
- [Ba2006] Batista, M. [2006] Integrability of the motion of a rolling disk of finite thickness on a rough plane. *Int. J. Mech.* **41**, 850–859.
- [Be1986] Benci, V. [1986] Periodic solutions of Lagrangian systems on a compact manifold. *J. Differ. Equations* **63**, 135–161.
- [BeShWy1999] Bendik, J. J. Jr., Shaw, L. J. and Wyles, R. H. [1999] Spinning/rolling disc. U. S. Patent No. 5863235.
- [BeBuHo1994] Berman, G. P., Bulgakov, E. N. and Holm, D. D. [1994] *Crossover-Time in Quantum Boson and Spin Systems*. Springer Lecture Notes in Physics, Vol. 21. New York: Springer.
- [BlBr1992] Blanchard, P. and Bruning, E. [1992] *Variational Methods in Mathematical Physics*. New York: Springer.
- [Bl2004] Bloch, A. M. (with the collaboration of Baillieul, J., Crouch, P. E. and Marsden, J. E.) [2004] *Nonholonomic Mechanics and Control*. New York: Springer.
- [BlBrCr1997] Bloch, A. M., Brockett, R. W. and Crouch, P. E. [1997] Double bracket equations and geodesic flows on symmetric spaces. *Comm. Math. Phys.* **187**, 357–373.

- [BlCr1996] Bloch, A. M. and Crouch, P. E. [1996] Optimal control and geodesic flows. *Sys. Control Lett.* **28**, 65–72.
- [BlCrHoMa2001] Bloch, A. M., Crouch, P. E., Holm, D. D. and Marsden, J. E. [2001] An optimal control formulation for inviscid incompressible ideal fluid flow. In *Proc. of the 39th IEEE Conference on Decision and Control*. New Jersey: IEEE, pp. 1273–1279.
- [BlCrMaRa1998] Bloch, A. M., Crouch, P. E., Marsden, J. E. and Ratiu, T. S. [1998] Discrete rigid body dynamics and optimal control. In *Proc. 37th CDC IEEE*. New Jersey: IEEE. pp. 2249–2254.
- [BlIs2006] Bloch, A. M. and Iserles, A. [2006] On an isospectral Lie-Poisson system and its Lie algebra. *Found. Comput. Math.* **6**, 121–144.
- [BlKrMaMu1996] Bloch, A. M., Krishnaprasad, P. S., Marsden, J. E. and Murray, R. [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. Anal.* **136**, 21–99.
- [Bl1965] Bloembergen, N. [1965] *Nonlinear Optics*. New York: Benjamin.
- [Bo1986] Bobenko, A. I. [1986] Euler equations in the algebras  $e(3)$  and  $so(4)$ . Isomorphisms of integrable cases. *Funkts. Anal. Prilozh.* **20**, 1, 64–66. (English translation: *Funct. Anal. Appl.* **20**, 53–56.)
- [Bo1985] Bogoyavlensky, O. I. [1985] Integrable Euler equations on Lie algebras arising in problems of mathematical physics. *Math. USSR Izvestiya* **25**, 207–257.
- [BoMa2002] Borisov, A. V. and Mamaev, I. S. [2002] Rolling of a rigid body on a plane and sphere: Hierarchy of dynamics. *Regul. Chaotic Dyn.* **7**, 177–200.
- [BoMaKi2002] Borisov, A. V., Mamaev, I. S. and Kilin, A. A. [2002] Dynamics of a rolling disk. *Regul. Chaotic Dyn.* **8**, 201–212.

## 370 BIBLIOGRAPHY

- [BoWo1965] Born, M. and Wolf, E. [1965] *Principles of Optics*. Oxford: Pergamon Press, p. 31.
- [BoMa2009] Bou-Rabee, N. and Marsden, J. E. [2009] Hamilton-Pontryagin integrators on Lie groups part I: Introduction and structure-preserving properties. *Foundations of Computational Mathematics* **9**, 197–219.
- [BoMaRo2004] Bou-Rabee, N., Marsden, J. E. and Romero, L. A. [2004] Tippe top inversion as a dissipation-induced instability. *SIAM J. Appl. Dyn. Syst.* **3**, 352–377.
- [Bo1971] Bourbaki, N. [1971] *Variétés différentielles et analytiques. Fascicule de résultats*. Paris: Hermann.
- [Bo1989] Bourbaki, N. [1989] *Lie Groups and Lie Algebras: Chapters 1–3*. New York: Springer.
- [BrGr2011] Brody, D. C. and Graefe, E.-M. [2011] On complexified mechanics and coquaternions. *J. Phys. A: Math. Theor.* **44**, 072001.
- [Ca1995] Calogero, F. [1995] An integrable Hamiltonian system. *Phys. Lett. A* **201**, 306–310.
- [CaFr1996] Calogero, F. and Francoise, J.-P. [1996] An integrable Hamiltonian system. *J. Math. Phys.* **37**, 2863–2871.
- [CaHo1993] Camassa, R. and Holm, D. D. [1993] An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.* **71**, 1661–1664.
- [CeHoMaRa1998] Cendra, H., Holm, D. D., Marsden, J. E. and Ratiu T. S. [1998] Lagrangian reduction, the Euler–Poincaré equations, and semidirect products. *Arnol'd Festschrift Volume II*, **186**, Am. Math. Soc. Translations Series 2 (1999), 1–25, arXiv.chao-dyn/9906004.
- [CeMa1987] Cendra, H., and Marsden, J. E. [1987] Lin constraints, Clebsch potentials and variational principles. *Physica D* **27**, 63–89.

## BIBLIOGRAPHY 371

- [CeMaPeRa2003] Cendra, H., Marsden, J. E., Pekarsky, S., and Ratiu, T. S. [2003] Variational principles for Lie–Poisson and Hamilton–Poincaré equations. *Mos. Math. J.* **3**(3), 833–867.
- [Ch1897] Chaplygin, S. A. [1897] On motion of heavy rigid body of revolution on horizontal plane. *Proc. of the Physical Sciences section of the Society of Amateurs of Natural Sciences* **9**, 10–16.
- [Ch1903] Chaplygin, S. A. [1903] On the rolling of a sphere on a horizontal plane. *Mat. Sbornik* **XXIV**, 139–68 (in Russian).
- [ChMa1974] Chernoff, P. R. and Marsden, J. E. [1974] *Properties of Infinite Dimensional Hamiltonian Systems*. Lecture Notes in Mathematics, Vol. 425. New York: Springer.
- [CiLa2007] Ciocci, M. C. and Langerock, B. [2007] Dynamics of the tippe top via Routhian reduction. arXiv:0704.1221v1 [math.DS].
- [Co1848] Cockle, J. [1848] A new imaginary in algebra. *Phil. Mag.* **33**, 345–349,
- [CoLe1984] Coddington, E. A. and Levinson, N. [1984] *Theory of Ordinary Differential Equations*. New York: Krieger.
- [Co2003] Cordani, B. [2003] *The Kepler Problem: Group Theoretical Aspects, Regularization and Quantization, with Application to the Study of Perturbations*. Basel: Birkhäuser.
- [CoHo2006] Cotter, C. J. and Holm, D. D. [2006] Discrete momentum maps for lattice EPDiff. arXiv.math.NA/0602296.
- [CoHo2007] Cotter, C. J. and Holm, D. D. [2007] Continuous and discrete Clebsch variational principles. arXiv.math/0703495. *Found. Comput. Math.* **9**, 221–242. DOI: 10.1007/s10208-007-9022-9.
- [CoHo2009] Cotter, C. J. and Holm, D. D. [2009] Discrete momentum maps for lattice EPDiff. In *Computational Methods for the Atmosphere and the Ocean* (a special volume of the *Handbook of Numerical Analysis*), edited by R. Temam and J. Tribbia.

## 372 BIBLIOGRAPHY

- San Diego, CA: Academic Press, pp. 247–278. (Preprint at [arxiv.org/abs/0901.2025](http://arxiv.org/abs/0901.2025)).
- [Cu1998] Cushman, R. [1998] Routh’s sphere. *Rep. Math. Phys.* **42**, 47–70.
- [CuBa1997] Cushman, R. H. and Bates, L. M. [1997] *Global Aspects of Integrable Systems*. Basel: Birkhäuser.
- [CuHeKe1995] Cushman, R., Hermans, J. and Kemppainen, O. [1995] The rolling disc. In *Nonlinear Dynamical Systems and Chaos*, edited by H. W. Broer, S. A. van Gils, I. Haveijn and Fl. Takens. Basel: Birkhäuser, pp. 21–60.
- [DaBaBi1992] Dandoloff, R., Balakrishnan, R. and Bishop, A. R. [1992] Two-level systems: Space curve formalism, Berry’s phase and Gauss-Bonnet theorem. *J. Phys. A: Math. Gen.* **25**, L1105–L1110.
- [Da1994] Darling, R. W. R. [1994] *Differential Forms and Connections*. Cambridge: Cambridge University Press.
- [DeMe1993] Dellnitz, M. and Melbourne, I. [1993] The equivariant Darboux theorem. *Lectures in Appl. Math.* **29**, 163–169.
- [doCa1976] do Carmo, M. P. [1976] *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ: Prentice-Hall.
- [DuNoFo1995] Dubrovin, B., Novikov, S. P. and Fomenko, A. T. [1995] *Modern Geometry I, II, III*. Graduate Texts in Mathematics, Vols. 93, 104, 124. New York: Springer.
- [El2007] Ellis, D. C. P. [2007] The Hopf fibration and the 1-1 resonant optical traveling wave pulse. Unpublished manuscript.
- [EGHPR2010] Ellis, D. C. P., Gay-Balmaz, F., Holm, D. D., Putkaradze, V. and Ratiu, T. S. [2010] Dynamics of charged molecular strands. *Arch. Rat. Mech. Anal.* **197**(3), 811–902.
- [Eu1758] Euler, L. [1758] Du mouvement de rotation des corps solides autour d’un axe variable. *Mem. de l’acad. xi. Berlin* **14**, 154–193.



## BIBLIOGRAPHY 373

- [FaTa1987] Faddeev, L. D. and Takhtajan, L. A. [1987] *Hamiltonian Methods in the Theory of Solitons*. New York: Springer.
- [Fl1963] Flanders, H. [1963] *Differential Forms with Applications to the Physical Sciences*. New York: Academic Press.
- [Ga1632] Galilei, G. [1632] *Dialogue Concerning the Two Chief World Systems: the Ptolemaic and the Copernican*.
- [GaMi1995] Galper, A. and Miloh, T. [1995] Dynamic equations of motion for a rigid or deformable body in an arbitrary non-uniform potential flow field. *J. Fluid Mech.* **295**, 91–120.
- [GBRa2008] Gay-Balmaz, F. and Ratiu, T.S. [2008] The geometric structure of complex fluids. *Adv. in Appl. Math.* **42** (2), 176–275.
- [GaVi2010] Gay-Balmaz F. and Vizman, C. [2011] Dual pairs in fluid dynamics. *Ann. Glob. Anal. Geom.* **41**, 1–24.
- [GeDo1979] Gelfand, I. M. and Dorfman, I. Ya. R. [1979] Hamiltonian operators and algebraic structures related to them. *Funct. Anal. Appl.* **13**, 248 (in Russian).
- [GiHoKeRo2006] Gibbons, J. D., Holm, D. D., Kerr, R. M. and Roulstone, I. [2006] Quaternions and particle dynamics in the Euler fluid equations. *Nonlinearity* **19**, 1969–1983.
- [GiHoKu1982] Gibbons, J., Holm, D. D. and Kupershmidt, B. A. [1982] Gauge-invariant Poisson brackets for chromohydrodynamics. *Phys. Lett. A* **90**, 281–283.
- [GLPeWo2007] Glad, S. T., Petersson, D. and Rauch-Wojciechowski, S. [2007] Phase space of rolling solutions of the tippe top. *SIGMA* **3**, 14 pages.
- [Go1986] Gorringer, V. M. [1986] *Generalisations of the Laplace-Runge-Lenz Vector in Classical Mechanics*. Thesis, Faculty of Science, University of Witwatersrand, Johannesburg, Vol. I, II.

## 374 BIBLIOGRAPHY

- [GrNi2000] Gray, C. G. and Nickel, B. G. [2000] Constants of the motion for nonslipping tippe tops and other tops with round pegs. *Am. J. Phys.* **68**, 821–828.
- [GrKrMa1988] Grossman, R., Krishnaprasad, P. S. and Marsden, J. E. [1988] The dynamics of two coupled rigid bodies. In *Dynamical Systems Approaches to Nonlinear Problems in Systems and Circuits*, edited by F. M. Abdel Salam and M. Levi, Engineering Foundation, United States Army Research Office, and U.S. National Science Foundation. Philadelphia, PA: SIAM, pp. 373–378.
- [GuHo1983] Guckenheimer, J. and Holmes, P. [1983] *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York: Springer.
- [GuSt1984] Guillemin, V. and Sternberg, S. [1984] *Symplectic Techniques in Physics*. Cambridge: Cambridge University Press.
- [Ha2006] Hanson, A. J. [2006] *Visualizing Quaternions*. San Francisco: MK-Elsevier.
- [He1995] Hermans, J. [1995] A symmetric sphere rolling on a surface. *Nonlinearity* **8**, 493–515.
- [Ho1983] Holm, D. D. [1983] Magnetic tornadoes: Three-dimensional affine motions in ideal magnetohydrodynamics. *Physica D* **8**, 170–182.
- [Ho1991] Holm, D. D. [1991] Elliptical vortices and integrable Hamiltonian dynamics of the rotating shallow water equations. *J. Fluid Mech.* **227**, 393–406.
- [Ho2002] Holm, D. D. [2002] Euler–Poincaré dynamics of perfect complex fluids. In *Geometry, Mechanics, and Dynamics: In Honor of the 60th Birthday of Jerrold E. Marsden*, edited by P. Newton, P. Holmes and A. Weinstein. New York: Springer, pp. 113–167.
- [Ho2005] Holm, D. D. [2005] The Euler–Poincaré variational framework for modeling fluid dynamics. In *Geometric Mechanics and*

## BIBLIOGRAPHY 375

- Symmetry: The Peyresq Lectures*, edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306, Cambridge: Cambridge University Press.
- [Ho2008] Holm, D. D. [2008] *Geometric Mechanics I: Dynamics and Symmetry*. London: Imperial College Press.
- [HoKu1988] Holm, D. D. and Kupershmidt, B. A. [1988] The analogy between spin glasses and Yang–Mills fluids. *J. Math Phys.* **29**, 21–30.
- [HoMa1991] Holm, D. D. and Marsden, J. E. [1991] The rotor and the pendulum. In *Symplectic Geometry and Mathematical Physics*, edited by P. Donato, C. Duval, J. Elhadad, G. M. Tuynman. Prog. in Math., Vol. 99. Birkhäuser: Boston, pp. 189–203.
- [HoMa2004] Holm, D. D. and Marsden, J. E. [2004] Momentum maps and measure valued solutions (peakons, filaments, and sheets) for the EPDiff equation. In *The Breadth of Symplectic and Poisson Geometry*, edited by J. Marsden, and T. Ratiu. Boston: Birkhäuser. arXiv.nlin.CD/0312048.
- [HoMaRa1998] Holm, D. D., Marsden, J. E. and Ratiu, T. S. [1998] The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* **137**, 1–81.
- [HoScSt2009] Holm, D. D., Schmah, T. and Stoica, C. [2009] *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford: Oxford University Press.
- [HoVi2010] Holm, D. D. and Vizman, C. [2010] *The  $n:m$  resonance dual pair*. Preprint at <http://arxiv.org/abs/1102.4377>.
- [HoJeLe1998] Holmes, P., Jenkins, J. and Leonard, N. E. [1998] Dynamics of the Kirchhoff equations I: Coincident centres of gravity and buoyancy. *Physica D* **118**, 311–342.
- [Ho1931] Hopf, H. [1931] Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche. *Math. Ann.* **104**, 637–665.

## 376 BIBLIOGRAPHY

- [Is1999] Isham, C. J. [1999] *Differential Geometry for Physicists*. World Scientific Lecture Notes in Physics Vol. 61. Singapore: World Scientific.
- [Iv2006] Ivanov, R. [2006] Hamiltonian formulation and integrability of a complex symmetric nonlinear system. *Phys. Lett. A* **350**, 232–235.
- [Ja2011] Jachnik, J. [2011] Spinning and rolling of an unbalanced disk. Unpublished report.
- [Je1872] Jellett, J. H. [1872] *A Treatise on the Theory of Friction*. Dublin: Hodges, Foster and Co.
- [JoSa1998] José, J. V. and Saletan, E. J. [1998] *Classical Dynamics: A Contemporary Approach*. Cambridge: Cambridge University Press.
- [Jo1998] Jost, J. [1998] *Riemannian Geometry and Geometric Analysis, 2nd ed.* New York: Springer.
- [KaKoSt1978] Kazhdan, D., Kostant, B. and Sternberg, S. [1978] Hamiltonian group actions and dynamical systems of Calogero type. *Comm. Pure Appl. Math.* **31**, 481–508.
- [KhMi2003] Khesin, B. and Misiolek, G. [2003] Euler equations on homogeneous spaces and Virasoro orbits. *Adv. in Math.* **176**, 116–144.
- [Ki2001] Kilin, A. A. [2001] The dynamics of Chaplygin ball: The qualitative and computer analysis. *Regul. Chaotic Dyn.* **6**, 291–306.
- [Ki2011] Kim, B. [2011] *Constrained Dynamics of Rolling Balls and Moving Atoms*. PhD Thesis, Colorado State University.
- [KiPuHo2011] Kim, B., Putkaradze, V. and Holm, D. D. [2011] Conservation laws of imbalanced symmetric Chaplygin's ball. Unpublished notes.

## BIBLIOGRAPHY 377

- [KnHj2000] Knudsen, J. M. and Hjorth, P. G. [2000] *Elements of Newtonian Mechanics: Including Nonlinear Dynamics*. New York: Springer.
- [KoNo1996] Kobayashi, S. and Nomizu, K. [1996] *Foundations of Differential Geometry*, Vols. 1, 2. New York: Wiley-Interscience.
- [Ko1984] Koiller, J. [1984] A mechanical system with a “wild” horseshoe. *J. Math. Phys.* **25**, 1599–1604.
- [Ko1900] Korteweg, D. [1900] *Extrait d’une lettre à M. Appel: Rendiconti del circolo matematico di Palermo* **14**, 7–8.
- [Ko1966] Kostant, B. [1966] Orbits, symplectic structures and representation theory. *Proc. US-Japan Seminar on Diff. Geom.*, Kyoto. Nippon Hyronsha, Tokyo, p. 71.
- [Ko1985] Kozlov, V. V. [1985] On the integration theory of the equations in nonholonomic mechanics. *Adv. Mech.* **8**, 86–107.
- [Ku1999] Kuipers, J. B. [1999] *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. New Jersey: Princeton University Press.
- [Ku1978] Kummer, M. [1978] On resonant classical Hamiltonians with two equal frequencies. *Commun. Math. Phys.* **58**, 85–112.
- [KuSt1965] Kustaanheimo, P. and Stiefel, E. [1965] Perturbation theory of Kepler motion based on spinor regularization. *J. Reine Angew. Math.* **218**, 204–219.
- [La1999] Lang, S. [1999] *Fundamentals of Differential Geometry*. Graduate Texts in Mathematics, Vol. 191. New York: Springer.
- [LeFl2003] Leach, P. G. and Flessa, G. P. [2003] Generalisations of the Laplace–Runge–Lenz vector. *J. Nonlinear Math. Phys.* **10**, 340–423.
- [Le2003] Lee, J. [2003] *Introduction to Smooth Manifolds*. New York: Springer.

## 378 BIBLIOGRAPHY

- [LeLeGl2005] Le Saux, C., Leine, R. I. and Glocker, C. [2005] Dynamics of a rolling disk in the presence of dry friction. *J. Non-linear Sci.* **15**, 27–61.
- [LiMa1987] Libermann, P. and Marle, C.-M. [1987] *Symplectic Geometry and Analytical Mechanics*. Dordrecht: Reidel.
- [Li1890] Lie, S. [1890] *Theorie der Transformationsgruppen. Zweiter Abschnitt*. Leipzig: Teubner.
- [LuAlMaRo2000] Luther, G. G., Alber, M. S., Marsden, J. E. and Robbins, J. M. [2000] Geometric analysis of optical frequency conversion and its control in quadratic nonlinear media. *J. Opt. Soc. Am. B Opt. Phys.* **17**, 932–941.
- [LyBu2009] Lynch, P. and Bustamante, M. [2009] Precession and recession of the rock'n'roller. *J. Phys. A: Math. Theor.* **42**, 425203.
- [Man1976] Manakov, S. V. [1976] Note on the integration of Euler's equations of the dynamics of an  $n$ -dimensional rigid body. *Funct. Anal. Appl.* **10**, 328–329.
- [Mar1976] Marle, C.-M. [1976] Symplectic manifolds, dynamical groups, and Hamiltonian mechanics. In *Differential Geometry and Relativity*, edited by Cahen, M. and Flato, M. Boston: Reidel, pp. 249–269.
- [Ma1981] Marsden, J. E. [1981] *Lectures on Geometric Methods in Mathematical Physics*. CMBS, Vol. 37. Philadelphia: SIAM.
- [Ma1992] Marsden, J. E. [1992] *Lectures on Mechanics*. London Mathematical Society Lecture Note Series, Vol. 174. Cambridge: Cambridge University Press.
- [MaHu1983] Marsden, J. E. and Hughes, T. J. R. [1983] *Mathematical Foundations of Elasticity*. Englewood Cliffs, NJ: Prentice Hall. (Reprinted by Dover Publications, NY, 1994.)
- [MaMiOrPeRa2007] Marsden, J. E., Misiolek, G., Ortega, J.-P., Perlmutter, M. and Ratiu, T. S. [2007] *Hamiltonian Reduction by*

- Stages*. Lecture Notes in Mathematics, Vol. 1913. New York: Springer.
- [MaMoRa1990] Marsden, J. E., Montgomery, R. and Ratiu, T. S. [1990] Reduction, symmetry, and phases in mechanics. *Memoirs Amer. Math. Soc.* **88**(436), 1–110.
- [MaRa1994] Marsden, J. E. and Ratiu, T. S. [1994] *Introduction to Mechanics and Symmetry*. Texts in Applied Mathematics, Vol. 75. New York: Springer.
- [MaRaWe1984a] Marsden, J. E., Ratiu, T. S. and Weinstein, A. [1984a] Semidirect products and reduction in mechanics. *Trans. Amer. Math. Soc.* **281**(1), 147–177.
- [MaRaWe1984b] Marsden, J. E., Ratiu, T. S. and Weinstein, A. [1984b] Reduction and Hamiltonian structures on duals of semidirect product Lie algebras. *Contemporary Math.* **28**, 55–100.
- [MaWe1974] Marsden, J. E. and Weinstein, A. [1974] Reduction of symplectic manifolds with symmetry. *Rep. Math. Phys.* **5**, 121–130.
- [MaWe1983] Marsden, J. E. and Weinstein, A. [1983] Coadjoint orbits, vortices and Clebsch variables for incompressible fluids. *Physica D* **7**, 305–323.
- [MaWe2001] Marsden, J. E. and Weinstein, A. [2001] Comments on the history, theory, and applications of symplectic reduction. In *Quantization of Singular Symplectic Quotients*, edited by N. Landsman, M. Pflaum and M. Schlichenmaier. Boston: Birkhäuser, pp. 1–20.
- [MaWil1989] Mawhin, J. and Willem, M. [1989] *Critical Point Theory and Hamiltonian Systems*, 2nd ed. Applied Mathematical Sciences, Vol. 74. New York: Springer.
- [McDMcD2000] McDonald, A. J. and McDonald, K. T. [2000] The rolling motion of a disk on a horizontal plane. arXiv:physics/0008227v3.

## 380 BIBLIOGRAPHY

- [McSa1995] McDuff, D. and Salamon, D. [1995] *Introduction to Symplectic Topology*. Oxford: Clarendon Press.
- [MeDe1993] Melbourne, I. and Dellnitz, M. [1993] Normal forms for linear Hamiltonian vector fields commuting with the action of a compact Lie group. *Math. Proc. Cambridge* **114**, 235–268.
- [Mi1963] Milnor, J. [1963] *Morse Theory*. Princeton: Princeton University Press.
- [MiFo1978] Mishchenko, A. S. and Fomenko, A. T. [1978] Euler equations on finite-dimensional Lie groups. *Izv. Akad. Nauk SSSR, Ser. Mat.* **42**, 396–415 (Russian); English translation: *Math. USSR-Izv.* **12**, 371–389.
- [Mi2002] Misiolek, G. [2002] Classical solutions of the periodic Camassa–Holm equation, *Geom. Funct. Anal.* **12**, 1080–1104.
- [Mo2002] Montgomery, R. [2002] *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. Mathematical Surveys and Monographs, Vol. 91. Providence, RI: American Mathematical Society.
- [Na1973] Nambu, Y. [1973] Generalized Hamiltonian mechanics. *Phys. Rev. D* **7**, 2405–2412.
- [Ne1997] Needham, T. [1997] *Visual Complex Analysis*. Oxford: Clarendon Press.
- [NeFu1972] Neimark, J. I. and Fufaev, N. A. [1972] *Dynamics of Non-holonomic Systems*. Translations of Mathematical Monographs, Vol. 33. Providence, RI: American Mathematical Society.
- [No1918] Noether, E. [1918] 1861 *Invariante Variationsprobleme*. *Nachrichten Gesell. Wissenschaft. Göttingen* **2**, 235–257. See also C. H. Kimberling [1972], Emmy Noether. *Am. Math. Monthly* **79**, 136–149.
- [OcoRo1998] O’Connor, J. J. and Robertson, E. F. [1998] *Sir William Rowan Hamilton*. <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Hamilton.html>.



## BIBLIOGRAPHY 381

- [Ol2000] Olver, P. J. [2000] *Applications of Lie Groups to Differential Equations, 2nd ed.* New York: Springer.
- [Or1996] O'Reilly, O. M. [1996] The dynamics of rolling disks and sliding disks. *Nonlinear Dynamics* **10**, 287–305.
- [OrRa2004] Ortega, J.-P. and Ratiu, T. S. [2004] *Momentum Maps and Hamiltonian Reduction*. Progress in Mathematics, Vol. 222. Boston: Birkhäuser.
- [Ot1993] Ott, E. [1993] *Chaos in Dynamical Systems*. Cambridge: Cambridge University Press, p. 385.
- [Pa1968] Palais, R. [1968] *Foundations of Global Non-Linear Analysis*. New York: W. A. Benjamin.
- [PeHuGr2002] Petrie, D., Hunt, J. L. and Gray, C. G. [2002] Does the Euler Disk slip during its motion? *Am. J. Phys.* **70**, 1025–1028.
- [Po1892] Poincaré, H. [1892] *Théorie Mathématique de la Lumière*, Paris: Georges Carré, p. 275.
- [Po1901] Poincaré, H. [1901] Sur une forme nouvelle des équations de la mécanique. *C. R. Acad. Sci.* **132**, 369–371.
- [Ra1980] Ratiu, T. [1980] The motion of the free  $n$ -dimensional rigid body. *Indiana Univ. Math. J.* **29**, 609–627.
- [Ra1982] Ratiu, T. [1982] Euler–Poisson equations on Lie algebras and the  $N$ -dimensional heavy rigid body. *Am. J. Math.* **104**, 409–448.
- [RaVM1982] Ratiu, T. and van Moerbeke, P. [1982] The Lagrange rigid body motion. *Annales de l'institut Fourier* **32**(1), 211–234.
- [RaTuSbSoTe2005] Ratiu, T. S., Tudoran, R., Sbano, L., Sousa Dias, E. and Terra, G. [2005] A crash course in geometric mechanics. In *Geometric Mechanics and Symmetry: The Peyresq Lectures*, edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306. Cambridge: Cambridge University Press.

## 382 BIBLIOGRAPHY

- [Ro1860] Routh, E. J. [1860] *A Treatise on the Dynamics of a System of Rigid Bodies*. London: MacMillan.  
*Ibid* [1884] *Advanced Rigid Body Dynamics*. London: Macmillan. (Reprint: *A Treatise on the Dynamics of a System of Rigid Bodies*, parts I and II, Dover, New York, 1960.)
- [Ru2002] Ruina, A. [2002] Rolling and sliding of spinning things: Euler's disk, Jellett's Egg and Moffatt's Nature. [http://ruina.tam.cornell.edu/research/topics/miscellaneous/rolling\\_and\\_sliding/index.htm#Eulers\\_Disk](http://ruina.tam.cornell.edu/research/topics/miscellaneous/rolling_and_sliding/index.htm#Eulers_Disk).
- [Ru2000] Rumyantsev, V. V. [2000] Forms of Hamilton's principle for nonholonomic systems. *Facta Univ. Ser. Mech. Automat. Control and Robot.* **2**(10), 1035–1048.
- [SaCa1981] Sarlet, W. and Cantrijn, F. [1981] Generalizations of Noether's theorem in classical mechanics. *SIAM Rev.* **23**, 467–494.
- [Sc2002] Schneider, D. [2002] Non-holonomic Euler–Poincaré equations and stability in Chaplygin's sphere. *Dynam. Syst.* **17**, 87–130.
- [SeWh1968] Seliger, R. L. and Whitham, G. B. [1968] Variational principles in continuum mechanics. *Proc. R. Soc. A* **305**, 1–25.
- [Se1992] Serre, J.-P. [1992] *Lie Algebras and Lie Groups*. Lecture Notes in Mathematics, Vol. 1500. New York: Springer.
- [Se1959] Serrin, J. [1959] In *Mathematical Principles of Classical Fluid Mechanics*, Vol. VIII/1 of Encyclopedia of Physics, edited by S. Flügge. New York: Springer, sections 14–15, pp. 125–263.
- [ShWi1989] Shapere, A. and Wilcek, F. [1989] *Geometric Phases in Physics*. Singapore: World Scientific.
- [SiMaKr1988] Simo, J. C., Marsden, J. E. and Krishnaprasad, P. S. [1988] The Hamiltonian structure of nonlinear elasticity: The material and convective representations of solids, rods, and plates. *Arch. Rat. Mech. Anal.* **104**, 125–183.

- [Sk1961] Skyrme, T. H. R. [1961] A nonlinear field theory. *Proc. R. Soc. A* **260**, 127–138.
- [Sl1861] Slessor, G. M. [1861] Notes on rigid dynamics. *Q. J. Math.* **4**, 65–67.
- [Sm1970] Smale, S. [1970] Topology and mechanics. *Inv. Math.* **10**, 305–331; **11**, 45–64.
- [So1970] Souriau, J. M. [1970] *Structure des Systèmes Dynamiques*. Paris: Dunod.
- [Sp1982] Sparrow, C. [1982] *The Lorenz Equations: Bifurcations, Chaos and Strange Attractors*. New York: Springer, p. 269.
- [Sp1979] Spivak, M. [1979] *Differential Geometry, Volume I*. Houston, TX: Publish or Perish, Inc.
- [St1852] Stokes, G. G. [1852] On the composition and resolution of streams of polarized light from different sources. *Trans. Camb. Phil. Soc.* **9**, 399–423.
- [Sy1937] Synge, J. L. [1937] *Geometrical Optics: An Introduction to Hamilton's Method*. Cambridge: Cambridge University Press.  
*Ibid* [1937] Hamilton's method in geometrical optics. *J. Opt. Soc. Am.* **27**, 75–82.  
*Ibid* [1960] Classical dynamics. In *Handbuch der Physik*, edited by S. Flügge. Berlin: Springer-Verlag, Vol. III/1, p. 11.
- [Va1996] Vaisman, I. [1996] *Lectures on the Geometry of Poisson Manifolds*. Progress in Mathematics, Vol. 118. Boston: Birkhäuser.
- [Wa1685] Wallis, J. *Treatise of Algebra* (London, 1685), p. 272, quoted by William Rowan Hamilton in the Preface to his *Lectures on Quaternions* (Dublin, 1883), p. 34.
- [Wa1983] Warner, F. W. [1983] *Foundation of Differentiable Manifolds and Lie Groups*. Graduate Texts in Mathematics, Vol. 94. New York: Springer.
- [We1983a] Weinstein, A. [1983a] Sophus Lie and symplectic geometry. *Expo. Math.* **1**, 95–96.

**384** BIBLIOGRAPHY

- [We1983b] Weinstein, A. [1983b] The local structure of Poisson manifolds. *J. Differ. Geom.* **18**, 523–557.
- [We2002] Weinstein, A. [2002] Geometry of momentum (preprint). ArXiv:math/SG0208108 v1.
- [Wh1974] Whitham, G. B. [1974] *Linear and Nonlinear Waves*. New York: John Wiley and Sons.
- [Wi1984] Witten, E. [1984] Non-Abelian Bosonization in two dimensions. *Commun. Math. Phys.* **92**, 455–472.
- [Wi1988] Wiggins, S. [1988] *Global Bifurcations and Chaos: Analytical Methods*. New York: Springer.
- [YoMa2006] Yoshimura, H. and Marsden, J. E. [2006]. Dirac Structures and Lagrangian Mechanics, Part I: Implicit Lagrangian systems. *J. Geom. and Physics*, **57**, 133–156.
- [ZaMi1980] Zakharov, V.E. and Mikhailov, A.V. [1980] On the integrability of classical spinor models in two-dimensional space-time, *Commun. Math. Phys.* **74**, 21–40.
- [Ze1995] Zenkov, D. V. [1995] The geometry of the Routh problem. *J. Nonlinear Sci.* **5**, 503–519.

# Index

---

- action integral, 37
- AD operation
  - for  $SE(3)$ , 144
- Ad operation
  - for  $SE(2)$ , 155
  - for  $SE(3)$ , 144
- Adjoint action, 119
  - for  $SE(2)$ , 155
  - for  $SE(3)$ , 145
  - for  $SO(3)$ , 133
- Adjoint motion
  - equation, 120
- adjoint operation
  - Ad, 119
  - ad, 119
- alien life forms
  - Bers, 334
  - Bichrons, 328
  - Tets, 338
- alignment dynamics
  - Newton's second law, 86
  - quaternion, 88
- angular momentum
  - body, 184
  - body vs space, 28, 138
  - conservation, 28, 33
  - definition, 27
  - Poisson bracket table, 346
- angular velocity
  - body, 29
  - hat map, 25
  - space vs body, 30
  - spatial, 23
- Bloch–Iserles equation, 357
- body angular momentum, 33
- body angular velocity vector, 32
- body frame, 21
- bundle, 297
  - cotangent, 299
  - tangent, 295
- Cartan
  - structure equations, 329
- Chaplygin's ball, 248
  - Jellett's integral, 261
  - Routh's integral, 263
- Chaplygin's circular disk, 265
- Chaplygin's concentric sphere, 265
- chiral model, 217
- Christoffel coefficients, 332
- Christoffel symbols, 38
- Clebsch action principle
  - Euler–Poincaré, 199
  - Hamiltonian formulation, 206
  - heavy top, 187
  - rigid body, 69
- coAdjoint action, 122
  - for  $SE(2)$ , 155
  - for  $SE(3)$ , 145
  - for  $SO(3)$ , 135
- coadjoint motion
  - equation, 129, 163
  - Kirchhoff equations, 176
  - on  $se(3)^*$ , 170, 174, 175
  - relation, 122, 162
- coadjoint operation, 121
- commutation relations, 24
- commutator
  - group, 125
- complex fluid, 215
- conjugacy classes, 119
  - Lie group action, 317
  - quaternions, 117
- conservation
  - $se(3)^*$  momentum, 173

## 386 INDEX

- angular momentum, 28, 44, 90
- energy, 35, 42, 283
- Hamilton's vector, 91
- Jellet's integral, 261
- Lie symmetry, 39
- linear momentum, 44
- momentum map, 225
- Noether's theorem, 39
- Runge–Lenz vector, 92
- constant of motion, 40
- constrained reduced Lagrangian, 257, 278, 283
- constrained variational principle
  - Clebsch, 69
  - Hamilton–Pontryagin, 50, 198
  - implicit, 70, 203
  - Kaluza–Klein, 188
  - nonholonomic, 273
- coordinate charts, 288
- coquaternion, 108
  - conjugate, 109
- cotangent bundle, 299
- cotangent lift, 207, 300, 318
  - momentum map, 227
- covariant derivative, 332
- curvature two-form, 330
- diamond operation, 202
  - definition, 147
- differentiable map, 301
  - derivative, 301
- dual pairs, 240
- equivariance, 228, 230, 299
- Euler's disk, 268
  - Jellet's integral, 275
- Euler's equations
  - rigid body, 35, 46
  - vorticity, 326
- Euler–Poincaré equation, 194
  - continuum spin chain, 212
  - for  $SE(3)$ , 197, 218
  - for  $Sp(2)$ , 219
  - heavy top, 183, 259
  - on  $se(3)^*$ , 171
- Euler–Poincaré theorem, 194
  - nonholonomic, 278
- Euler–Rodrigues formula
  - rotation, 96, 112
- fibre bundle, 105
- fibre derivative, 53
- flows, 296
- force
  - noninertial, 29
- frame
  - body, 21
  - spatial, 22
- free rigid rotation, 21
- G-strand, 210, 359
- Galilean group, 4, 12, 17, 164
  - adjoint actions, 164
  - Lie group actions, 164
  - Lie algebra, 17
  - Lie group actions, 12
- Galilean relativity, 3
- Galilean transformations
  - boosts, 4
  - definition, 3
  - invariance, 3
  - Lie group property, 4
  - matrix representation, 14, 165
  - subgroups, 8
- geodesic equations, 38
- group
  - action, 8
  - automorphism, 126
  - commutator, 125
  - definition, 4, 306
  - Euclidean  $E(3)$ , 8
  - Galilean  $G(3)$ , 4
  - Heisenberg  $H$ , 124
  - matrix, 306
  - orthogonal  $O(3)$ , 8, 21
  - orthogonal  $O(n)$ , 312
  - semidirect  $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$ , 157
  - special Euclidean  $SE(2)$ , 153
  - special Euclidean  $SE(3)$ , 8, 15, 142, 197, 218, 313
  - special linear  $SL(n, \mathbb{R})$ , 312
  - special orthogonal  $SO(3)$ , 8, 21

- special orthogonal  $SO(4)$ , 338
- special orthogonal  $SO(n)$ , 312
- special unitary  $SU(2)$ , 101
- special unitary  $SU(n)$ , 313
- symplectic  $Sp(l)$ , 313
- group action
  - homogeneous, 11
  - vector space, 11
- group commutator, 125
- group composition law
  - for  $SE(3)$ , 9
- group orbit, 315
- Hamilton's principle, 37
- Hamilton–Pontryagin principle, 252
  - Euler–Poincaré, 198
  - heavy top, 259
  - nonholonomic, 273
  - rigid body, 50
- Hamiltonian form, 62
  - continuum spin chain, 213
  - Lie–Poisson, 173, 206
- hat map, 25, 132, 137, 323
  - isomorphism, 25
- heavy top, 182
  - Casimirs, 171
  - Clebsch action principle, 187
  - Euler–Poincaré equation, 259
  - Hamilton–Pontryagin, 259
  - Kaluza–Klein, 188
  - Legendre transform, 184
  - Lie–Poisson bracket, 178, 186
  - variational principle, 183
- Heisenberg group, 124
  - adjoint actions, 126
  - coadjoint actions, 127
  - coadjoint orbits, 130
- Hermitian
  - Pauli matrices, 101, 229
- Hopf fibration, 105, 228
- immersed submanifold, 294
- infinitesimal generator, 134
- infinitesimal transformation, 40
- injective immersion, 294
- inner automorphism, 119
- inverse Adjoint motion
  - identity, 121
- isotropy subgroup, 307
- iterated action
  - semidirect product, 167
- Jacobi identity, 62, 133
  - Lie algebra, 311
  - matrix commutator, 311
- Jacobi–Lie bracket, 320
- Jellett's integral
  - Chaplygin's ball, 261
  - Euler's disk, 275
- Kelvin–Noether theorem
  - for  $SE(3)$ , 176
  - for an arbitrary Lie group, 197
- Kepler problem
  - Hamilton's vector, 91
  - monopole, 93
  - quaternionic variables, 90
  - Runge–Lenz vector, 91
- Kirchhoff equations
  - underwater vehicle, 176
- Kustaanheimo–Stiefel map, 101
- Lagrange–d'Alembert, 275, 279
- Lagrangian
  - constrained reduced, 257, 278, 282, 283
  - reduced, 256, 278
- Legendre transform, 60
  - $se(3) \rightarrow se(3)^*$ , 172
  - Clebsch variables, 206
  - constrained reduced, 283
  - heavy top, 184
- Leibniz, 62
- Lie algebra
  - Jacobi identity, 311
  - matrix, 310
- Lie bracket
  - matrix commutator, 119
  - semidirect product, 187
- Lie derivative, 323
  - fluid dynamics, 325
- Lie group, 4, 306

## 388 INDEX

- action, 314
- as manifold, 294
- matrix, 118, 306
- semidirect product, 12
- Lie group action, 314
  - adjoint, 317
  - coadjoint, 317
  - conjugacy classes, 317
  - cotangent lift, 318
  - infinitesimal generator, 322
  - inner automorphism, 317
  - invariant vector fields, 319
  - left action, 314
  - left and right translations, 316
  - on vector fields, 318
  - properties, 315
  - tangent lift, 317
- Lie symmetry, 39
- Lie symmetry reduction, 242
- Lie–Poisson
  - Hamiltonian form, 206
- Lie–Poisson bracket, 64, 184
  - Casimirs, 173
  - continuum spin chain, 214
  - distinguished functions, 173
  - heavy top, 178, 186
  - momentum maps, 185
  - on  $se(3)^*$ , 172
  - spin chain, 210
- lower triangular matrices, 128
- Manakov
  - commutator form, 54, 170, 359
  - heavy top, 190
- manifold
  - coordinate charts, 288
  - definition, 288
  - smooth, 287
  - submersion, 289
- matrix commutator
  - Jacobi identity, 311
  - properties, 311
- matrix Lie algebra, 310
- matrix Lie group, 118, 306
  - $GL(n, \mathbb{R})$ , 307
- matrix representation
  - $G(3)$ , 13, 14, 165
  - $SE(3)$ , 9
  - $SE(3)$ , 10, 142
  - $SL(2, \mathbb{R}) \otimes \mathbb{R}^2$ , 157
  - for quaternions, 80
  - hat map, 15
  - Heisenberg group, 124
- mechanics on Lie groups, 79
- moment of inertia tensor, 27, 32
  - principal axis frame, 33
- momentum map, 71, 236
  - $GL(n, \mathbb{R})$  invariance, 348
  - $SU(2)$  action on  $\mathbb{C}^2$ , 227
  - $SU(3)$  action on  $\mathbb{C}^3$ , 348
  - $\mathbb{C}^2 \mapsto u(2)^*$ , 236
  - $\mathbb{R}^6 \rightarrow sp(2)^*$ , 354
  - 1:1 resonance, 236
  - canonical symplectic form, 236
- Clebsch variational principle, 208
- cotangent lift, 208, 224, 227
- equivariant, 230
- examples, 231
- Hopf fibration, 229
- infinite dimensional, 240
- Noether’s theorem, 225
- Poincaré sphere, 228, 229
- standard, 222
- motion equation
  - coadjoint, 129, 163
- Nambu bracket, 65
- Nambu form
  - geometric interpretation, 68
  - three-wave equations, 68
- narrow-hat notation, 23
- Noether’s theorem, 39
  - angular momentum, 45
  - centre of mass, 46
  - energy and momentum, 42
  - Euler–Poincaré motion, 195
  - Galilean boosts, 46
  - Hamilton–Pontryagin, 199
  - nonholonomic, 282
  - rigid body, 49, 53
  - space and time translations, 42



- nonholonomic
  - Chaplygin's ball, 248
  - Chaplygin's circular disk, 265
  - Chaplygin's concentric sphere, 265
  - Euler's disk, 268
  - Euler–Poincaré theorem, 278
  - Hamilton–Pontryagin, 252
  - Lagrange–d'Alembert, 279
  - Noether's theorem, 282
  - rolling constraint, 250
  - semidirect product, 276
- notation
  - narrow hat, 23
  - wide hat, 23
- orthogonal group, 21
- orthogonal transformation, 4
- oscillators
  - on  $\mathbb{C}^2$ , 228
  - on  $\mathbb{C}^3$ , 343
  - Poincaré sphere, 236
- pairing
  - on  $SE(3)$ , 145
- Pauli matrices, 80, 101, 229
  - anticommutator relation, 81
  - commutator relation, 81
  - Hermitian, 229
  - quaternions, 80
- Poincaré
  - 1901 paper, 55, 65, 194, 363
- Poincaré sphere, 104, 228
  - Hopf fibration, 105
  - momentum map, 229
  - oscillators, 236
- Poisson bracket, 62
  - rigid body, 63
- principal axis frame, 33
- product rule, 62
- prolongation, 41
- pull-back, 318
- push-forward, 318
- quaternion, 78
  - action on a vector, 112
- alignment dynamics, 88
- basis, 79
- conjugacy class, 95
- conjugate, 83
- conjugation, 93
- dot product, 83
- inner product, 83
- inverse, 84
- magnitude, 84
- multiplication rule, 80
- Pauli matrices, 80
- pure vector, 85
- tetrad, 79
- reconstruction
  - formula, 50, 69, 152, 200
- reduced Lagrangian, 256, 257, 278
- reflection, 4
- Riemann sphere, 289
- rigid body, 21
  - Cayley–Klein parameters, 112
  - Clebsch approach, 69
  - coadjoint orbit of  $SO(3)$ , 135
  - coupled, 359
  - eigenvalue problem, 56
  - Euler's equations, 46
  - Hamiltonian form, 60
  - isospectral problem, 56
  - Lagrangian form, 46
  - Lie–Poisson bracket, 64
  - Manakov's formulation, 54
  - matrix Euler equations, 51
  - Newtonian form, 30
  - Noether's theorem, 53
  - Poincaré's treatment, 365
  - quaternionic form, 112, 116
  - spatial and body frames, 138
  - symmetric equations, 70
  - variations, 328
- Rodrigues formula, 96
- rolling constraint, 250
- rotating motion
  - free rigid rotation, 21
  - in complex space, 334
  - in two times, 328
  - kinetic energy, 31

## 390 INDEX

- quaternions, 112
- rotation, 4
  - Cayley–Klein parameters, 97
  - Euler parameters, 97
  - Euler–Rodrigues formula, 96, 113
  - in complex space, 334
  - in four dimensions, 338
- Routh’s integral
  - Chaplygin ball, 263
- semidirect product
  - Adjoint action, 146
  - adjoint action, 148
  - coAdjoint action, 146
  - coadjoint action, 148
  - definition, 12
  - iterated, 167
  - Lie bracket, 149
    - vector notation, 149
  - Lie group, 143
- semidirect-product group
  - $G(3)$ , 14
  - $SE(3)$ , 12
  - definition, 12, 143
- semidirect-product structure
  - nonholonomic, 276
- sigma model, 217
- simple mechanical system, 37
- Skyrme model, 217
- sliding, 258
- space-time translations, 3
- spatial frame, 22, 31
- special orthogonal group, 21
- spin glass, 215
- stereographic projection, 289
- strand
  - $so(3)$ -valued spins, 210
- structure constants, 24
- subgroup, 8
  - invariant, 11
  - isotropy, 307
  - normal, 11
- submanifold, 288
- submersion, 289
- symplectic Lie group
  - $Sp(2)$ , 354
- tangent bundle, 295
- tangent lift, 41, 300, 317
- tangent space, 294
- tangent vector, 295, 296
- three-wave equations
  - Nambu form, 68
- tilde map, 102
  - Hopf fibration, 106
- translations
  - space-time, 3
- transpose
  - of the Lie-derivative, 203
- twist, 87
- underwater vehicle
  - Kirchhoff equations, 176
  - variational principle, 176
- uniform rectilinear motion, 3
- upper triangular matrices, 124
- variational derivative, 36, 47
  - left-invariant Lagrangian, 170
- variational principle, 49
  - chiral model, 217
  - Clebsch, 199
  - Clebsch Euler–Poincaré, 202
  - constrained, 72, 279
  - continuum spin chain, 212
  - Euler–Poincaré, 194, 278
  - Hamilton–Pontryagin, 198, 252
  - heavy top, 183
  - implicit, 70, 200
  - Lagrange–d’Alembert, 279
  - Noether’s theorem, 39
  - rigid body, 47
  - sigma model, 217
  - underwater vehicle, 176
- vector field, 297
  - differential basis, 300
  - fluid dynamics, 325
- wide-hat notation, 23
- zero curvature relation, 330
  - continuum spin chain, 212