Applications of Poisson geometry to physical problems

DARRYL D HOLM

These being lecture notes for a summer school, one should not seek original material in them. Rather, the most one could hope to find would be the insight arising from incorporating a unified approach (based on reduction by symmetry of Hamilton's principle) with some novel applications. I hope the reader will find insight in the lecture notes, which are meant to be informal, more like stepping stones than a proper path.

37K05, 53Z05, 70S05; 37K10, 37K65, 70S10

Preface

Many excellent encyclopedic texts have already been published on the foundations of this subject and its links to symplectic and Poisson geometry. See, for example, Abraham and Marsden [1], Arnold [3], Guillemin and Sternberg [21], José and Saletan [35], Libermann and Marle [39], Marsden and Ratiu [44], McDuff and Salamon [50] and many more. In fact, the scope encompassed by the modern literature on this subject is a bit overwhelming. In following the symmetry-reduction theme in geometric mechanics from the Euler–Poincaré viewpoint, I have tried to select only the material the student will find absolutely necessary for solving the problems and exercises, at the level of a beginning postgraduate student. The primary references are Marsden [42], Marsden and Ratiu [44], Lee [38], Bloch [5], and Ratiu, Tudoran, Sbano, Sousa Dias and Terra [57]. Other very useful references are Arnold and Khesin [4] and Olver [53]. The reader may see the strong influences of all these references in these lecture notes, but expressed at a considerably lower level of mathematical sophistication than the originals.

The scope of these lectures is quite limited: a list of the topics in geometric mechanics not included in these lectures would fill volumes! The necessary elements of calculus on smooth manifolds and the basics of Lie group theory are only briefly described here, because these topics were discussed in more depth by other lecturers at the summer school. Occasional handouts are included that add a bit more depth in certain key topics. The main subject of these lecture notes is the use of Lie symmetries in Hamilton's principle to derive symmetry-reduced equations of motion and to analyze their solutions.

Published: 20 April 2011 DOI: 10.2140/gtm.2011.17.221

The Legendre transformation provides the Hamiltonian formulation of these equations in terms of Lie–Poisson brackets.

For example, we consider Lagrangians in Hamilton's principle defined on the tangent space TG of a Lie group G. Invariance of such a Lagrangian under the action of G leads to the symmetry-reduced Euler–Lagrange equations called the Euler–Poincaré equations. In this case, the invariant Lagrangian is defined on the Lie algebra of the group and its Euler–Poincaré equations are defined on the dual Lie algebra, where dual is defined by the operation of taking variational derivative. On the Hamiltonian side, the Euler–Poincaré equations are Lie–Poisson and they possess accompanying momentum maps, which encode both their conservation laws and the geometry of their solution space.

The standard Euler–Poincaré examples are treated, including particle dynamics, the rigid body, the heavy top and geodesic motion on Lie groups. Additional topics deal with Fermat's principle, the $\mathbb{R}^3$ Poisson bracket, polarized optical traveling waves, deformable bodies (Riemann ellipsoids) and shallow water waves, including the integrable shallow water wave system known as the Camassa–Holm equation. The lectures end with the semidirect-product Euler–Poincaré reduction theorem for ideal fluid dynamics. This theorem introduces the Euler–Poincaré variational principle for incompressible and compressible motions of ideal fluids, with applications to geophysical fluids. It also leads to their Lie–Poisson Hamiltonian formulation.

Some of these lectures were first given at the MASIE (Mechanics and Symmetry in Europe) summer school in 2000 [24]. I am grateful to the MASIE participants for their helpful remarks and suggestions which led to many improvements in those lectures. For their feedback and comments, I am also grateful to my colleagues at Imperial College London, especially Colin Cotter, Matthew Dixon, J D Gibbon, J Gibbons, G Gottwald, J T Stuart, J-L Thiffeault, Cesare Tronci and the students who attended these thirty three lectures in my classes at Imperial College in Spring 2005. After each class, the students were requested to turn in a response sheet on which they answered two questions. These questions were, "What was this class about?" and "What question would you like to see pursued in the class?" The answers to these questions helped keep the lectures on track with the interests and understanding of the students and it enfranchised the students because they themselves selected the material in several of the lectures. Over the past few years, these lectures have evolved into the textbooks [25; 26; 33]. However, the lecture form here conveys the immediacy of the original course.

I am enormously grateful to many friends and colleagues whose encouragement, advice and support have helped sustain my interest in this field. I am particularly grateful to

J E Marsden, T S Ratiu and A Weinstein for their faithful camaraderie in many research endeavors.

1 Introduction

1.1 Road map for the course

- Spaces – Smooth Manifolds
- Motion – Flows $\phi_t \circ \phi_s = \phi_{t+s}$ of Lie groups acting on smooth manifolds
- Laws of Motion and discussion of solutions
- Newton's Laws
  - Newton: $dp/dt = F$, for momentum p and prescribed force F (on $\mathbb{R}^n$ historically)
  - Optimal motion
    * Euler–Lagrange equations – optimal "action" (Hamilton's principle)
    * Geodesic motion – optimal with respect to kinetic energy metric
- Lagrangian and Hamiltonian Formalism
  - Newton's Law of motion
  - Euler–Lagrange theorem
  - Noether theorem
  - Euler–Poincaré theorem
  - Kelvin–Noether theorem
- Applications and examples
  - Geodesic motion on a Riemannian manifold
  - Rigid body – geodesic motion on $SO(3)$
  - Other geodesic motion, for example, Riemann ellipsoids on $GL(3, R)$
  - Heavy top
- Lagrangian mechanics on Lie groups and Euler–Poincaré (EP) equations
  - $\mathrm{EP}(G)$, EP equations for geodesics on a Lie group G
  - $\mathrm{EPDiff}(\mathbb{R})$ for geodesics on $\mathrm{Diff}(\mathbb{R})$
  - Pulsons, the singular solutions of $\mathrm{EPDiff}(\mathbb{R}))$ with respect to any norm
  - Peakons, the singular solitons for $\mathrm{EPDiff}(\mathbb{R}, H^1)$, with respect to the H^1 norm
  - $\mathrm{EPDiff}(\mathbb{R}^n)$ and singular geodesics
  - Diffeons and momentum maps for $\mathrm{EPDiff}(\mathbb{R}^n)$
- Euler–Poincaré (EP) equations for continua
  - EP semidirect-product reduction theorem
  - Kelvin–Noether circulation theorem
  - EP equations with advected parameters for geophysical fluid dynamics

1.2 Hamilton's principle of stationary action

Lagrangians on $T\mathbb{R}^{3N}$	**G–invariant Lagrangians on TG**
Euler–Lagrange equations	Euler–Poincaré equations
Noether's theorem	Kelvin–Noether theorem
Symmetry $\implies$ cons. laws	Cons. laws are built-in
Legendre transformation	Legendre transformation
Hamilton's canonical equations	Lie–Hamilton equations
Poisson brackets	Lie–Poisson brackets
Symplectic manifold	Poisson manifold
Momentum map	Momentum map
Reduction by symmetry	Reduction to coadjoint orbits

1.3 Motivation for the geometric approach

We begin with a series of outline sketches to motivate the geometric approach taken in the course and explain more about its content.

Why is the geometric approach useful?

- Defines problems on manifolds
  - coordinate-free
    - don't have to re-do calculations when changing coordinates
    - more compact
    - unified framework for expressing ideas and using symmetry
- "First principles" approach
  - variational principles
  - systematic – unified approach
    for example, similarity between tops and fluid dynamics (semi-direct product), and MHD, and …
  - POWER
    Geometric constructions can give useful answers without us having to find and work with complicated explicit solutions. For example, stability of rigid body equilibria.

1.4 Course outline

- Geometrical Structure of Classical Mechanics
  - Smooth manifolds
    - calculus

  * tangent vectors
  * action principles
- Lie groups
  * flow property $\phi_{t+s} = \phi_t \circ \phi_s$
  * symmetries encode conservation laws into geometry
  * richer than vector spaces
- Variational principles with symmetries
  * Euler–Lagrange equations $\to$ Euler–Poincaré equations (more compact)
  * Two main formulations:

Lagrangian side:	Hamiltonian side:
Hamilton's principle	Lie–Poisson brackets
Noether's theorem	cons. laws $\Longleftrightarrow$ symmetries
symmetry $\Longrightarrow$ cons. laws	momentum maps
momentum maps	Jacobi identity

(These two views are mutual beneficial!)

- Applications and Modelling
  - oscillators and resonance (for example, LASER)
  - tops – integrable case
  - fluids
  - waves $\begin{cases}\text{shallow water waves}\\ \text{optical pulses}\\ \text{solitons}\end{cases}$

1.5 Range of topics

Rigid body

- Euler–Lagrange and Euler–Poincaré equations
- Kelvin–Noether theorem
- Lie–Poisson bracket, Casimirs and coadjoint orbits
- Reconstruction and momentum maps
- The symmetric form of the rigid body equations ($\dot{Q} = Q\Omega,\ \dot{P} = P\Omega$)
- $\mathbb{R}^3$ bracket and intersecting level surfaces

$$\dot{\mathbf{x}} = \nabla C \times \nabla H = \nabla(\alpha C + \beta H) \times \nabla(\gamma C + \epsilon H), \quad \text{for} \quad \alpha\epsilon - \beta\gamma = 1$$

Examples:
(1) Conversion: rigid body $\Longleftrightarrow$ pendulum,
(3) Fermat's principle and ray optics,
(2) Self-induced transparency.

- Nonlinear oscillators: the $n : m$ resonance
- $SU(2)$ rigid body, Cayley–Klein parameters and Hopf fibration
- The Poincaré sphere for polarization dynamics
- 3–wave resonance, Maxwell–Bloch equations, cavity resonators, symmetry reduction and the Hopf fibration
- 4–wave resonance, coupled Hopf fibrations, coupled Poincaré spheres and coupled rigid bodies
- Higher dimensional rigid bodies
  - Manakov integrable top on $O(n)$ and its spectral problem
- Semi-rigid bodies – geodesic motion on $GL(3)$ and Riemann ellipsoids
- Reduction with respect to subgroups of $GL(3)$ and Calogero equations

Heavy top

- Euler–Poincaré variational principle for the heavy top
- Kaluza–Klein formulation of the heavy top

Utility

- Kirchhoff elastica, underwater vehicles, liquid crystals, stratified flows, polarization dynamics of telecom optical pulses

General theory

- Euler–Poincaré semidirect-product reduction theorem
- Semidirect-product Lie–Poisson formulation

Shallow water waves

- CH equation – peakons (geodesics)
- EPDiff equation – (also geodesics)

Fluid dynamics

- Euler–Poincaré variational principle for incompressible ideal fluids
- Euler–Poincaré variational principle for compressible ideal fluids

Outlook The variational principles and the Poisson brackets for the rigid body and the heavy top provide models of a general construction associated to Euler–Poincaré reduction with respect to any Lie group. The Hamiltonian counterpart will be the semidirect-product Lie–Poisson formulation. We will often refer to the rigid body and the heavy top for interpretation and enhanced understanding of the general results.

2 Review: Newton, Lagrange and Hamilton

2.1 Newton's Law

$m\ddot{q} = F(q, \dot{q})$, inertial frames, uniform motion, etc.

2.2 Lagrange's equations

$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}$ for Lagrangian $L(q, \dot{q}, t)$.

Defined on the tangent bundle[1] TQ of the configuration space Q with coordinates $(q, \dot{q}) \in TQ$, the solution is a curve (or trajectory) in Q parametrized by time t. The tangent vector of the curve $q(t)$ through each point $q \in Q$ is the velocity $\dot{q}$ along the trajectory that passes though the point q at time t. This vector is written $\dot{q} \in T_q Q$.

Lagrange's equations may be expressed compactly in terms of vector fields and one-forms (differentials). Namely, the Lagrangian vector field $X_L = \dot{q}\frac{\partial}{\partial q} + F(q, \dot{q})\frac{\partial}{\partial \dot{q}}$ acts on the one-form $(\frac{\partial L}{\partial \dot{q}}\, dq)$ just as a time-derivative does, to yield

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}\, dq\Big) = \Big(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\Big)\, dq + \Big(\frac{\partial L}{\partial \dot{q}}\Big)\, d\dot{q} = dL \quad \Longrightarrow \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}$$

2.3 Hamiltonian $H(p \cdot q) = p\dot{q} - L$ and Hamilton's canonical equations

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}$$

The *configuration space* Q has coordinates $q \in Q$. Its phase space, or cotangent bundle T^*Q has coordinates $(q, p) \in T^*Q$.

[1] The terms tangent bundle and cotangent bundle are defined in Section 5. For now, we may think of the tangent bundle as the space of positions and velocities. Likewise, the cotangent bundle is the space of positions and momenta.

Hamilton's canonical equations are associated to the *canonical Poisson bracket* for functions on phase space, by

$$\dot{p} = \{p, H\}, \quad \dot{q} = \{q, H\} \Longleftrightarrow \dot{F}(q, p) = \{F, H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q}$$

The canonical Poisson bracket has the following familiar properties, which may be readily verified:

(1) It is bilinear,

(2) skew symmetric, $\{F, H\} = -\{H, F\}$,

(3) satisfies the Leibniz rule (chain rule),

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

for the product of any two phase space functions F and G,

(4) and satisfies the Jacobi identity

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$$

for any three phase space functions F, G and H.

Its Leibniz property (chain rule) property means the canonical Poisson bracket is a type of derivative. This derivation property of the Poisson bracket allows its use in defining the *Hamiltonian vector field* X_H, by

$$X_H = \{\cdot, H\} = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p},$$

for any phase space function H. The action of X_H on phase space functions is given by

$$\dot{p} = X_H p, \quad \dot{q} = X_H q, \quad \text{and} \quad X_H(FG) = (X_H F)G + FX_H G = \dot{F}G + F\dot{G}.$$

Thus, solutions of Hamilton's canonical equations are the characteristic paths of the first order linear partial differential operator X_H. That is, X_H corresponds to the time derivative along these characteristic paths, given by

$$dt = \frac{dq}{\partial H/\partial p} = \frac{dp}{-\partial H/\partial q} \tag{1}$$

The union of these paths in phase space is called the *flow* of the Hamiltonian vector field X_H.

Proposition 2.1 (Poisson bracket as commutator of Hamiltonian vector fields) *The Poisson bracket* $\{F, H\}$ *is associated to the commutator of the corresponding Hamiltonian vector fields* X_F *and* X_H *by*

$$X_{\{F,H\}} = X_H X_F - X_F X_H =: -[X_F, X_H]$$

Proof Verified by direct computation. □

Corollary 2.2 *Thus, the Jacobi identity for the canonical Poisson bracket* $\{\cdot, \cdot\}$ *is associated to the Jacobi identity for the commutator* $[\cdot, \cdot]$ *of the corresponding Hamiltonian vector fields,*

$$[X_F, [X_G, X_H] + [X_G, [X_H, X_F] + [X_H, [X_F, X_G] = 0.$$

Proof This is the Lie algebra property of Hamiltonian vector fields, as verified by direct computation. □

2.4 Differential forms

The *differential*, or *exterior derivative* of a function F on phase space is written

$$dF = F_q dq + F_p dp,$$

in which subscripts denote partial derivatives. For the Hamiltonian itself, the exterior derivative and the canonical equations yield

$$dH = H_q dq + H_p dp = -\dot{p} dq + \dot{q} dp.$$

The action of a Hamiltonian vector field X_H on a phase space function F commutes with its differential, or exterior derivative. Thus,

$$d(X_H F) = X_H (dF).$$

This means X_H may also act as a time derivative on differential forms defined on phase space. For example, it acts on the time-dependent one-form $p\,dq(t)$ along solutions of Hamilton's equations as

$$\begin{aligned} X_H\big(p\,dq\big) = \frac{d}{dt}\big(p\,dq\big) &= \dot{p}\,dq + p\,d\dot{q} \\ &= \dot{p}\,dq - \dot{q}\,dp + d(p\dot{q}) \\ &= -H_q dq - H_p dp + d(p\dot{q}) \\ &= d(-H + p\dot{q}) =: dL(q, p) \end{aligned}$$

upon substituting Hamilton's canonical equations.

The exterior derivative of the one-form pdq yields the canonical, or symplectic two-form[2]

$$d(pdq) = dp \wedge dq$$

Here we have used the chain rule for the exterior derivative and its property that $d^2 = 0$. (The latter amounts to equality of cross derivatives for continuous functions.) The result is written in terms of the wedge product $\wedge$, which combines two one-forms (the line elements dq and dp) into a two-form (the oriented surface element $dp \wedge dq = -dq \wedge dp$). As a result, the two-form $\omega = dq \wedge dp$ representing area in phase space is conserved along the Hamiltonian flows:

$$X_H\big(dq \wedge dp\big) = \frac{d}{dt}\big(dq \wedge dp\big) = 0$$

This proves

Theorem 2.3 (Poincaré's theorem) *Hamiltonian flows preserve area in phase space.*

Definition 2.4 (Symplectic two-form) The phase space area $\omega = dq \wedge dp$ is called the symplectic two-form.

Definition 2.5 (Symplectic flows) Flows that preserve area in phase space are said to be symplectic.

Remark 2.6 (Poincaré's theorem) Hamiltonian flows are symplectic.

3 Handout on exterior calculus, symplectic forms and Poincaré's theorem in higher dimensions

Exterior calculus on symplectic manifolds is the geometric language of Hamiltonian mechanics. As an introduction and motivation for more detailed study, we begin with a preliminary discussion.

In differential geometry, the operation of *contraction* denoted as $\,\lrcorner\,$ introduces a pairing between vector fields and differential forms. Contraction is also called *substitution* of a vector field into a differential form. For example, there are the dual relations,

$$\partial_q \,\lrcorner\, dq = 1 = \partial_p \,\lrcorner\, dp, \quad \text{and} \quad \partial_q \,\lrcorner\, dp = 0 = \partial_p \,\lrcorner\, dq$$

[2] The properties of differential forms are summarized in the handouts in sections 3 and 15.

A Hamiltonian vector field
$$X_H = \dot{q}\frac{\partial}{\partial q} + \dot{p}\frac{\partial}{\partial p} = H_p\partial_q - H_q\partial_p = \{\cdot, H\}$$
satisfies
$$X_H \lrcorner\, dq = H_p \quad \text{and} \quad X_H \lrcorner\, dp = -H_q.$$
The rule for contraction or substitution of a vector field into a differential form is to sum the substitutions of X_H over the permutations of the factors in the differential form that bring the corresponding dual basis element into its leftmost position. For example, substitution of the Hamiltonian vector field X_H into the symplectic form $\omega = dq \wedge dp$ yields
$$X_H \lrcorner\, \omega = X_H \lrcorner\, (dq \wedge dp) = (X_H \lrcorner\, dq)\, dp - (X_H \lrcorner\, dp)\, dq$$
In this example, $X_H \lrcorner\, dq = H_p$ and $X_H \lrcorner\, dp = -H_q$, so
$$X_H \lrcorner\, \omega = H_p dp + H_q dq = dH$$
which follows because $\partial_q \lrcorner\, dq = 1 = \partial_p \lrcorner\, dp$ and $\partial_q \lrcorner\, dp = 0 = \partial_p \lrcorner\, dq$. This calculation proves

Theorem 3.1 (Hamiltonian vector field) *The Hamiltonian vector field $X_H = \{\cdot, H\}$ satisfies*

$$X_H \lrcorner\, \omega = dH \quad \textit{with} \quad \omega = dq \wedge dp \tag{2}$$

Relation (2) may be taken as the *definition* of a Hamiltonian vector field.

As a consequence of this formula, the flow of X_H preserves the closed exact two form ω for any Hamiltonian H. This preservation may be verified by a formal calculation using (2). Along $(dq/dt, dp/dt) = (\dot{q}, \dot{p}) = (H_p, -H_q)$, we have
$$\begin{aligned}\frac{d\omega}{dt} &= d\dot{q} \wedge dp + dq \wedge d\dot{p} = dH_p \wedge dp - dq \wedge dH_q \\ &= d(H_p\, dp + H_q\, dq) = d(X_H \lrcorner\, \omega) = d(dH) = 0\end{aligned}$$
The first step uses the chain rule for differential forms and the third and last steps use the property of the exterior derivative d that $d^2 = 0$ for continuous forms. The latter is due to equality of cross derivatives $H_{pq} = H_{qp}$ and antisymmetry of the wedge product: $dq \wedge dp = -dp \wedge dq$.

Consequently, the relation $d(X_H \lrcorner\, \omega) = d^2 H = 0$ for Hamiltonian vector fields shows the following.

Theorem 3.2 (Poincaré's theorem for one degree of freedom) *The flow of a Hamiltonian vector field is* symplectic, *which means it preserves the phase-space area, or two-form,* $\omega = dq \wedge dp$.

Definition 3.3 (Cartan's formula for the Lie derivative) The operation of *Lie derivative* of a differential form ω by a vector field X_H is defined by

$$\pounds_{X_H}\omega := d(X_H \lrcorner\, \omega) + X_H \lrcorner\, d\omega \tag{3}$$

Corollary 3.4 *Because* $d\omega = 0$, *the symplectic property* $d\omega/dt = d(X_H \lrcorner\, \omega) = 0$ *in Poincaré's Theorem 3.2 may be rewritten using Lie derivative notation as*

$$0 = \frac{d\omega}{dt} = \pounds_{X_H}\omega := d(X_H \lrcorner\, \omega) + X_H \lrcorner\, d\omega =: (\operatorname{div} X_H)\omega. \tag{4}$$

The last equality defines the divergence *of the vector field* X_H *in terms of the Lie derivative.*

Remark 3.5

- Relation (4) associates Hamiltonian dynamics with the symplectic flow in phase space of the Hamiltonian vector field X_H, which is divergenceless with respect to the symplectic form ω.
- The Lie derivative operation defined in (4) is equivalent to the time derivative along the characteristic paths (flow) of the first order linear partial differential operator X_H, which are obtained from its characteristic equations in (1). This is the *dynamical meaning* of the Lie derivative $\pounds_{X_H}$ in (3) for which invariance $\pounds_{X_H}\omega = 0$ gives the geometric definition of symplectic flows in phase space.

Theorem 3.6 (Poincaré's theorem for N degrees of freedom) *For a system of N particles, or N degrees of freedom, the flow of a Hamiltonian vector field preserves each subvolume in the phase space $T^*\mathbb{R}^N$. That is, let $\omega_n \equiv dq_n \wedge dp_n$ be the symplectic form expressed in terms of the position and momentum of the nth particle. Then*

$$\frac{d\omega_M}{dt} = 0, \quad \textit{for} \quad \omega_M = \Pi_{n=1}^M \omega_n, \quad \forall M \le N.$$

The proof of the preservation of these *Poincaré invariants* ω_M with $M = 1, 2, \dots, N$ follows the same pattern as the verification above for a single degree of freedom. Basically, this is because each factor $\omega_n = dq_n \wedge dp_n$ in the wedge product of symplectic forms is preserved by its corresponding Hamiltonian flow in the sum

$$X_H = \sum_{n=1}^M \left(\dot{q}_n \frac{\partial}{\partial q_n} + \dot{p}_n \frac{\partial}{\partial p_n}\right) = \sum_{n=1}^M \left(H_{p_n}\partial_{q_n} - H_{q_n}\partial_{p_n}\right) = \sum_{n=1}^M X_{H_n} = \{\cdot, H\}$$

That is, $\pounds_{X_{H_n}}\omega_M$ vanishes for each term in the sum $\pounds_{X_H}\omega_M = \sum_{n=1}^M \pounds_{X_{H_n}}\omega_M$ since $\partial_{q_m} \lrcorner\, dq_n = \delta_{mn} = \partial_{p_m} \lrcorner\, dp_n$ and $\partial_{q_m} \lrcorner\, dp_n = 0 = \partial_{p_m} \lrcorner\, dq_n$.

4 Fermat's theorem in geometrical ray optics

4.1 Fermat's principle: Rays take paths of least optical length

In geometrical optics, the ray path is determined by Fermat's principle of least optical length,

$$\delta \int n(x,y,z)\, ds = 0.$$

Here $n(x,y,z)$ is the index of refraction at the spatial point (x,y,z) and ds is the element of arc length along the ray path through that point. Choosing coordinates so that the z–axis coincides with the optical axis (the general direction of propagation), gives

$$ds = [(dx)^2 + (dy)^2 + (dz)^2]^{1/2} = [1 + \dot{x}^2 + \dot{y}^2]^{1/2}\, dz,$$

with $\dot{x} = dx/dz$ and $\dot{y} = dy/dz$. Thus, Fermat's principle can be written in Lagrangian form, with z playing the role of time,

$$\delta \int L(x,y,\dot{x},\dot{y},z)\, dz = 0.$$

Here, the optical Lagrangian is,

$$L(x,y,\dot{x},\dot{y},z) = n(x,y,z)[1 + \dot{x}^2 + \dot{y}^2]^{1/2} =: n/\gamma,$$

or, equivalently, in two-dimensional vector notation with $\mathbf{q} = (x,y)$,

$$L(\mathbf{q},\dot{\mathbf{q}},z) = n(\mathbf{q},z)[1 + |\dot{\mathbf{q}}|^2]^{1/2} =: n/\gamma \quad \text{with} \quad \gamma = [1 + |\dot{\mathbf{q}}|^2]^{-1/2} \le 1.$$

Consequently, the vector Euler–Lagrange equation of the light rays is

$$\frac{d}{ds}\Big(n\frac{d\mathbf{q}}{ds}\Big) = \gamma\frac{d}{dz}\Big(n\gamma\frac{d\mathbf{q}}{dz}\Big) = \frac{\partial n}{\partial \mathbf{q}}.$$

The momentum p canonically conjugate to the ray path position q in an "image plane", or on an "image screen", at a fixed value of z is given by

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}} = n\gamma\dot{\mathbf{q}}$$

which satisfies $|\mathbf{p}|^2 = n^2(1-\gamma^2)$. This implies the velocity $\dot{\mathbf{q}} = \mathbf{p}/(n^2 - |\mathbf{p}|^2)^{1/2}$.

Hence, the momentum is real-valued and the Lagrangian is hyperregular, provided $n^2 - |\mathbf{p}|^2 > 0$. When $n^2 = |\mathbf{p}|^2$, the ray trajectory is vertical and has *grazing incidence* with the image screen.

Defining $\sin\theta = dz/ds = \gamma$ leads to $|\mathbf{p}| = n\cos\theta$, and gives the following geometrical picture of the ray path. Along the optical axis (the z–axis) each image plane normal to the axis is pierced at a point $\mathbf{q} = (x, y)$ by a vector of magnitude $n(\mathbf{q}, z)$ tangent to the ray path. This vector makes an angle θ to the plane. The projection of this vector onto the image plane is the canonical momentum $\mathbf{p}$. This picture of the ray paths captures all but the rays of grazing incidence to the image planes. Such grazing rays are ignored in what follows.

Passing now via the usual Legendre transformation from the Lagrangian to the Hamiltonian description gives

$$H = \mathbf{p}\cdot\dot{\mathbf{q}} - L = n\gamma|\dot{\mathbf{q}}|^2 - n/\gamma = -n\gamma = -\left[n(\mathbf{q}, z)^2 - |\mathbf{p}|^2\right]^{1/2}$$

Thus, in the geometrical picture, the component of the tangent vector of the ray-path along the optical axis is (minus) the Hamiltonian, that is, $n(\mathbf{q}, z)\sin\theta = -H$.

The phase space description of the ray path now follows from Hamilton's equations,

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{-1}{H}\mathbf{p}, \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} = \frac{-1}{2H}\frac{\partial n^2}{\partial \mathbf{q}}.$$

Remark 4.1 (Translation invariant media) If $n = n(\mathbf{q})$, so that the medium is translation invariant along the optical axis, z, then $H = -n\sin\theta$ is conserved. (Conservation of H at an interface is Snell's law.) For translation-invariant media, the vector ray-path equation simplifies to

$$\ddot{\mathbf{q}} = -\frac{1}{2H^2}\frac{\partial n^2}{\partial \mathbf{q}},$$

Newtonian dynamics for $\mathbf{q} \in \mathbb{R}^2$.

Thus, in this case geometrical ray tracing reduces to "Newtonian dynamics" in z, with potential $-n^2(\mathbf{q})$ and with "time" rescaled along each path by the value of $\sqrt{2}H$ determined from the initial conditions for each ray.

4.2 Axisymmetric, translation invariant materials

In axisymmetric, translation invariant media, the index of refraction is a function of the radius alone. Axisymmetry implies an additional constant of motion and, hence, reduction of the Hamiltonian system for the light rays to phase plane analysis. For such media, the index of refraction satisfies

$$n(\mathbf{q}, z) = n(r), \qquad r = |\mathbf{q}|.$$

Passing to polar coordinates (r, ϕ) with $\mathbf{q} = (x, y) = r(\cos\phi, \sin\phi)$ leads in the usual way to

$$|\mathbf{p}|^2 = p_r^2 + p_\phi^2/r^2.$$

Consequently, the optical Hamiltonian,

$$H = -\big[n(r)^2 - p_r^2 - p_\phi^2/r^2\big]^{1/2}$$

is independent of the azimuthal angle ϕ; so its canonically conjugate "angular momentum" p_ϕ is conserved.

Using the relation $\mathbf{q}\cdot\mathbf{p} = rp_r$ leads to an interpretation of p_ϕ in terms of the image-screen phase space variables $\mathbf{p}$ and $\mathbf{q}$. Namely,

$$|\mathbf{p}\times\mathbf{q}|^2 = |\mathbf{p}|^2|\mathbf{q}|^2 - (\mathbf{p}\cdot\mathbf{q})^2 = p_\phi^2$$

The conserved quantity $p_\phi = \mathbf{p}\times\mathbf{q} = yp_x - xp_y$ is called the skewness function, or the Petzval invariant for axisymmetric media. Vanishing of p_ϕ occurs for *meridional rays*, for which $\mathbf{p}$ and $\mathbf{q}$ are collinear in the image plane. On the other hand, p_ϕ takes its maximum value for *sagittal rays*, for which $\mathbf{p}\cdot\mathbf{q} = 0$, so that $\mathbf{p}$ and $\mathbf{q}$ are orthogonal in the image plane.

Exercise 4.2 (Axisymmetric, translation invariant materials) Write Hamilton's canonical equations for axisymmetric, translation invariant media. Solve these equations for the case of an optical fiber with radially graded index of refraction in the following form:

$$n^2(r) = \lambda^2 + (\mu - \nu r^2)^2, \quad \lambda, \mu, \nu = \text{constants},$$

by reducing the problem to phase plane analysis. How does the phase space portrait differ between $p_\phi = 0$ and $p_\phi \neq 0$? Show that for $p_\phi \neq 0$ the problem reduces to a Duffing oscillator in a rotating frame, up to a rescaling of time by the value of the Hamiltonian on each ray "orbit."

4.3 The Petzval invariant and its Poisson bracket relations

The skewness function

$$S = p_\phi = \mathbf{p}\times\mathbf{q} = yp_x - xp_y$$

generates rotations of phase space, of $\mathbf{q}$ and $\mathbf{p}$ jointly, each in its plane, around the optical axis. Its square, S^2 (called the Petzval invariant) is conserved for ray optics in axisymmetric media. That is, $\{S^2, H\} = 0$ for optical Hamiltonians of the form,

$$H = -\left[n(|\mathbf{q}|^2)^2 - |\mathbf{p}|^2\right]^{1/2}.$$

We define the axisymmetric invariant coordinates by the map $T^*\mathbb{R}^2 \mapsto \mathbb{R}^3$ $(\mathbf{q}, \mathbf{p}) \mapsto (X, Y, Z)$,

$$X = |\mathbf{q}|^2 \geq 0, \quad Y = |\mathbf{p}|^2 \geq 0, \quad Z = \mathbf{p} \cdot \mathbf{q}.$$

The following Poisson bracket relations hold

$$\{S^2, X\} = 0, \quad \{S^2, Y\} = 0, \quad \{S^2, Z\} = 0,$$

since rotations preserve dot products. In terms of these invariant coordinates, the Petzval invariant and optical Hamiltonian satisfy

$$S^2 = XY - Z^2 \geq 0, \quad \text{and} \quad H^2 = n^2(X) - Y \geq 0.$$

The level sets of S^2 are hyperboloids of revolution around the $X = Y$ axis, extending up through the interior of the $S = 0$ cone, and lying between the X– and Y–axes. The level sets of H^2 depend on the functional form of the index of refraction, but they are Z–independent.

4.4 $\mathbb{R}^3$ Poisson bracket for ray optics

The Poisson brackets among the axisymmetric variables X, Y and Z close among themselves,

$$\{X, Y\} = 4Z, \quad \{Y, Z\} = -2Y, \quad \{Z, X\} = -2X.$$

These Poisson brackets derive from a single $\mathbb{R}^3$ Poisson bracket for $\mathbf{X} = (X, Y, Z)$ given by

$$\{F, H\} = -\nabla S^2 \cdot \nabla F \times \nabla H$$

Consequently, we may re-express the equations of Hamiltonian ray optics in axisymmetric media with $H = H(X, Y)$ as

$$\dot{\mathbf{X}} = \nabla S^2 \times \nabla H.$$

with Casimir S^2, for which $\{S^2, H\} = 0$, for every H. Thus, the flow preserves volume ($\operatorname{div} \dot{\mathbf{X}} = 0$) and the evolution takes place on intersections of level surfaces of the axisymmetric media invariants S^2 and $H(X, Y)$.

4.5 Recognition of the Lie–Poisson bracket for ray optics

The Casimir invariant $S^2 = XY - Z^2$ is quadratic. In such cases, one may write the $\mathbb{R}^3$ Poisson bracket in the suggestive form

$$\{F, H\} = -C^k_{ij} X_k \frac{\partial F}{\partial X_i} \frac{\partial H}{\partial X_j}.$$

In this particular case, $C^3_{12} = 4$, $C^2_{23} = 2$ and $C^1_{31} = 2$ and the rest either vanish, or are obtained from antisymmetry of C^k_{ij} under exchange of any pair of its indices. These values are the structure constants of any of the Lie algebras $sp(2,\mathbb{R})$, $so(2,1)$, $su(1,1)$, or $sl(2,\mathbb{R})$. Thus, the reduced description of Hamiltonian ray optics in terms of axisymmetric $\mathbb{R}^3$ variables is said to be "Lie–Poisson" on the dual space of any of these Lie algebras, say, $sp(2,\mathbb{R})^*$ for definiteness. We will have more to say about Lie–Poisson brackets later, when we reach the Euler–Poincaré reduction theorem.

Exercise 4.3 Consider the $\mathbb{R}^3$ Poisson bracket

$$\{f, h\} = -\nabla c \cdot \nabla f \times \nabla h \tag{5}$$

Let $c = \mathbf{x}^T \cdot \mathbb{C}\mathbf{x}$ be a quadratic form on $\mathbb{R}^3$, and let $\mathbb{C}$ be the associated symmetric 3×3 matrix. Show that this is the Lie–Poisson bracket for the Lie algebra structure

$$[\mathbf{u}, \mathbf{v}]_{\mathbb{C}} = \mathbb{C}(\mathbf{u} \times \mathbf{v})$$

What is the underlying matrix Lie algebra? What are the coadjoint orbits of this Lie algebra?

Remark 4.4 (Coadjoint orbits) As one might expect, the coadjoint orbits of the group $SP(2,\mathbb{R})$ are the hyperboloids corresponding to the level sets of S^2.

Remark 4.5 As we shall see later, the map $T^*\mathbb{R}^2 \mapsto sp(2,\mathbb{R})^*$ taking $(\mathbf{q}, \mathbf{p}) \mapsto (X, Y, Z)$ is an example of a *momentum map*.

5 Geometrical structure of classical mechanics

5.1 Manifolds

Configuration space: coordinates $q \in M$, where M is a smooth manifold.

The composition $\phi_\beta \circ \phi_\alpha^{-1}$ is a smooth change of variables.

For later, smooth coordinate transformations: $q \to Q$ with $dQ = \frac{\partial Q}{\partial q} dq$

Definition 5.1 A smooth 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 $\phi_\alpha \colon U_\alpha \to \mathbb{R}^n$ such that

(1) $\bigcup_\alpha U_\alpha = M$

(2) For every nonempty intersection $U_\alpha \cap U_\beta$, the set $\phi_\alpha\left(U_\alpha \cap U_\beta\right)$ 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 $\phi_\alpha\left(U_\alpha \cap U_\beta\right)$.

Remark 5.2 The sets U_α in the definition are called *coordinate charts*. The mappings ϕ_α are called *coordinate functions* or *local coordinates*. A collection of charts satisfying 1 and 2 is called an *atlas*. Condition 3 allows the definition of manifold to be made independently of a choice of atlas. A set of charts satisfying 1 and 2 can always be extended to a maximal set; so, in practice, conditions 1 and 2 define the manifold.

Example 5.3 Manifolds often arise as intersections of zero level sets

$$M = \left\{x \,\middle|\, f_i(x) = 0,\ i = 1, \dots, k\right\},$$

for a given set of functions $f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$. If the gradients ∇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 neighborhood of each point on M. In this situation, the set M is called a *submanifold* of $\mathbb{R}^n$ (see Lee [38]).

Definition 5.4 If $r = k$, then the map $\{f_i\}$ is called a *submersion*.

Exercise 5.5 Prove that all submersions are submanifolds (see Lee [38]).

Definition 5.6 (Tangent space to level sets) Let

$$M = \left\{x \,\middle|\, f_i(x) = 0,\ i = 1, \dots, k\right\}$$

be a manifold in $\mathbb{R}^n$. The *tangent space* at each $x \in M$, is defined by

$$T_x M = \left\{ v \in \mathbb{R}^n \,\middle|\, \frac{\partial f_i}{\partial x^a}(x) v^a = 0,\ i = 1, \dots, k \right\}.$$

Note: we use the *summation convention*, that is, repeated indices are summed over their range.

Remark 5.7 The tangent space is a linear vector space.

Example 5.8 (Tangent space to the sphere in $\mathbb{R}^3$)

Example 5.9 (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 5.10 (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 $x \in M$; that is,

$$TM = \bigcup_{x \in M} T_x M$$

Thus, a point of TM is a vector v which is tangent to M at some point $x \in M$.

Example 5.11 (Tangent bundle TS^2 of S^2) The tangent bundle TS^2 of $S^2 \in \mathbb{R}^3$ is the union of the tangent spaces of S^2:

$$TS^2 = \left\{(x, y, z; u, v, w) \in \mathbb{R}^6 \;\middle|\; x^2 + y^2 + z^2 = 1 \text{ and } xu + yv + zw = 0\right\}.$$

Remark 5.12 (Dimension of tangent bundle TS^2) Defining TS^2 requires two independent conditions in $\mathbb{R}^6$; so $\dim TS^2 = 4$.

Exercise 5.13 Define the sphere S^{n-1} in $\mathbb{R}^n$. What is the dimension of its tangent space TS^{n-1}?

Example 5.14 (The two stereographic projections of $S^2 \to \mathbb{R}^2$) The unit sphere

$$S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$$

is a smooth two-dimensional manifold realized as a submersion in $\mathbb{R}^3$. Let

$$U_N = S^2 \backslash \{0, 0, 1\}, \quad \text{and} \quad U_S = S^2 \backslash \{0, 0, -1\}$$

be the subsets obtained by deleting the North and South poles of S^2, respectively. Let

$$\chi_N \colon U_N \to (\xi_N, \eta_N) \in \mathbb{R}^2, \quad \text{and} \quad \chi_S \colon U_S \to (\xi_S, \eta_S) \in \mathbb{R}^2$$

be stereographic projections from the North and South poles onto the equatorial plane, $z = 0$.

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).

In the equatorial plane $z = 0$, one may define two sets of (right-handed) coordinates,

$$\phi_\alpha \colon U_\alpha \to \mathbb{R}^2 \backslash \{0\}, \quad \alpha = N, S,$$

obtained by the following two stereographic projections from the North and South poles:

(1) (valid everywhere except $z = 1$)

$$\phi_N(x, y, z) = (\xi_N, \eta_N) = \Big(\frac{x}{1-z}, \frac{y}{1-z} \Big),$$

(2) (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,

$$\phi_N \circ \phi_S^{-1} \colon \mathbb{R}^2 \backslash \{0\} \to \mathbb{R}^2 \backslash \{0\}$$

is a smooth diffeomorphism, given by the inversion

$$\phi_N \circ \phi_S^{-1}(x, y) = \Big(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2} \Big).$$

Exercise 5.15 Construct the mapping from $(\xi_N, \eta_N) \to (\xi_S, \eta_S)$ and verify that it is a diffeomorphism in $\mathbb{R}^2 \backslash \{0\}$. Hint: $(1 + z)(1 - z) = 1 - z^2 = x^2 + y^2$.

Answer 5.16

$$(\xi_S, -\eta_S) = \frac{1-z}{1+z} (\xi_N, \eta_N) = \frac{1}{\xi_N^2 + \eta_N^2} (\xi_N, \eta_N).$$

The map $(\xi_N, \eta_N) \to (\xi_S, \eta_S)$ is smooth and invertible except at $(\xi_N, \eta_N) = (0, 0)$.

Example 5.17 If we start with two identical circles in the xz–plane, of radius r and centered at $x = \pm 2r$, then rotate them round the z axis in $\mathbb{R}^3$, we get a torus, written T^2. It's a manifold.

Exercise 5.18 If we begin with a figure eight in the xz–plane, along the x axis and centered 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 5.19 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.

Remark 5.20 The sphere will appear in several examples as a reduced space in which motion takes place after applying a symmetry. Reduction by symmetry is associated with a classical topic in celestial mechanics known as normal form theory. Reduction may be "singular," in which case it leads to "pointed" spaces that are smooth manifolds except at one or more points. For example different resonances of coupled oscillators correspond to the following reduced spaces: 1:1 resonance – sphere; 1:2 resonance – pinched sphere with one cone point; 1:3 resonance – pinched sphere with one cusp point; 2:3 resonance – pinched sphere with one cone point and one cusp point.

5.2 Motion: tangent vectors and flows

Envisioning our later considerations of dynamical systems, we shall consider motion along curves $c(t)$ parametrized by time t on a smooth manifold M. Suppose these curves are trajectories of a flow ϕ_t of a vector field. 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 the concept of *tangent vectors*.

Recall from Definition 5.10 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 5.21 (Bundle) A *bundle* consists of a manifold B, another manifold M called the "base space" and a projection between them $\Pi\colon B \to M$. Locally, in small enough regions of x the inverse images of the projection Π exist. These are called the *fibers* of the bundle. Thus, subsets of the bundle B locally have the structure of a Cartesian product. An example is (B, M, Π) consisting of $(\mathbb{R}^2, \mathbb{R}^1, \Pi\colon \mathbb{R}^2 \to \mathbb{R}^1)$. In this case, $\Pi\colon (x, y) \in \mathbb{R}^2 \to x \in \mathbb{R}^1$. Likewise, the tangent bundle consists of M, TM and a map $\tau_M\colon TM \to M$.

Let $x = \left(x^1, \dots, x^n\right)$ be local coordinates on M, and let $v = \left(v^1, \dots, v^n\right)$ be components of a tangent vector.

$$T_x M = \left\{ v \in \mathbb{R}^n \;\middle|\; \frac{\partial f_i}{\partial x} \cdot v = 0, i = 1, \dots, m \right\}$$

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, where $\dim TM = 2 \dim M$. The *tangent bundle projection* is a map $\tau_M \colon TM \to M$ 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 *fiber* over x in the tangent bundle.

5.3 Vector fields, integral curves and flows

Definition 5.22 A *vector field* on a manifold M is a map $X \colon M \to TM$ that assigns a vector $X(x)$ at each point $x \in M$. This implies that $\tau_M \circ X = \mathrm{Id}$.

Definition 5.23 An *integral curve* of X with initial conditions x_0 at $t = 0$ is a differentiable map $c \colon]a, b[\to M$, where $]a, b[$ is an open interval containing 0, such that $c(0) = 0$ and $c'(t) = X\,(c(t))$ for all $t \in]a, b[$.

Remark 5.24 A standard result from the theory of ordinary differential equations states that X being Lipschitz implies its integral curves are unique and C^1 (see Coddington and Levinson [12]). The integral curves $c(t)$ are differentiable for smooth X.

5.4 Summary

Definition 5.25 The *flow* of X is the collection of maps $\phi_t \colon M \to M$, where $t \to \phi_t(x)$ is the integral curve of X with initial condition x.

Remark 5.26

(1) Existence and uniqueness results for solutions of $c'(t) = X(c(t))$ guarantee that flow ϕ of X is smooth in (x, t), for smooth X.

(2) Uniqueness implies the flow property

$$\phi_{t+s} = \phi_t \circ \phi_s, \tag{FP}$$

for initial condition $\phi_0 = \mathrm{Id}$.

(3) The flow property (FP) generalizes 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 $\phi_t(x) = e^{At} x$.

5.5 Differentials of functions and the cotangent bundle

We are now ready to define differentials of smooth functions and the cotangent bundle.

Let $f\colon M \to \mathbb{R}$ be a smooth function. We differentiate f at $x \in M$ to obtain $T_x f\colon T_x M \to T_{f(x)}\mathbb{R}$. As is standard, we identify $T_{f(x)}\mathbb{R}$ with $\mathbb{R}$ itself, thereby obtaining a linear map $df(x)\colon T_x M \to \mathbb{R}$. 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 $\langle\cdot,\cdot\rangle\colon T_x^* M \times T_x M \mapsto \mathbb{R}$.

In coordinates, the linear map $df(x)\colon T_x M \to \mathbb{R}$ 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 intended 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 5.27 df is the *differential* of the function f.

Definition 5.28 The dual space of the tangent bundle TM is the *cotangent bundle* $T^* M$. That is,

$$(T_x M)^* = T_x^* M \quad \text{and} \quad T^* M = \bigcup_x T_x^* M.$$

Thus, replacing $v \in T_x M$ with $df \in T_x^* M$, for all $x \in M$ and for all smooth functions $f\colon M \to \mathbb{R}$, 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 δ_i^j is the Kronecker delta, which equals unity for $i = j$ and vanishes otherwise. That is, defining $T^* M$ requires a pairing $\langle\cdot,\cdot\rangle : T^* M \times TM \to \mathbb{R}$.

(Different pairings exist for curvilinear coordinates, Riemannian manifolds, etc.)

6 Derivatives of differentiable maps – the tangent lift

We next define derivatives of differentiable maps between manifolds (tangent lifts).

We expect that a smooth map $f\colon U \to V$ 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,

$$U \times \mathbb{R}^m \subset TM \longrightarrow V \times \mathbb{R}^n \subset TN$$
$$\left(q^1, \dots, q^m; X^1, \dots, X^m\right) \longmapsto \left(Q^1, \dots, Q^n; Y^1, \dots, Y^n\right)$$

Namely, the relations between the vector field components should be obtained from the differential of the map $f\colon U \to V$. 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*

$$Tf = \frac{\partial Q}{\partial q}$$

of the function f arises from the chain rule and is equal to the Jacobian for the transformation $Tf\colon TM \mapsto TN$.

The dual of the tangent lift is the cotangent lift, explained later in Definition 16.11. Roughly speaking, the *cotangent lift* of the function f,

$$T^* f = \frac{\partial q}{\partial Q}$$

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 $T^* f\colon T^* N \mapsto T^* M$. Note the directions of these maps:

$$\begin{aligned} Tf\colon & q, X \in TM \mapsto Q, Y \in TN \\ f\colon & q \in M \mapsto Q \in N \\ T^* f\colon & Q, \beta \in T^* N \mapsto q, \alpha \in T^* M \qquad \text{(map goes the other way)} \end{aligned}$$

6.1 Summary remarks about derivatives on manifolds

Definition 6.1 (Differentiable map) A map $f\colon M \to N$ 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 6.2 (Derivative of a differentiable map) The *derivative* of a differentiable map

$$f\colon M \to N$$

at a point $x \in M$ is defined to be the linear map

$$T_x f\colon T_x M \to T_x N$$

constructed, as follows. For $v \in T_x M$, choose a curve $c(t)$ that maps an open interval $t \in (-\epsilon, \epsilon)$ around the point $t = 0$ to the manifold M:

$$c\colon (-\epsilon, \epsilon) \to M$$

with $c(0) = x$ and velocity vector $c'(0) := \frac{dc}{dt}\big|_{t=0} = v$.
Then $T_x f \cdot v$ is the velocity vector at $t = 0$ of the curve $f \circ c\colon \mathbb{R} \to N$. That is,

$$T_x f \cdot v = \frac{d}{dt} f(c(t))\Big|_{t=0} = \frac{\partial f}{\partial c}\frac{d}{dt}c(t)\Big|_{t=0}$$

Definition 6.3 The union $Tf = \bigcup_x T_x f$ of the derivatives $T_x f\colon T_x M \to T_x N$ over points $x \in M$ is called the *tangent lift* of the map $f\colon M \to N$.

Remark 6.4 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 $f\colon M \to N$ is of class C^{k+1}, then the tangent lift (Jacobian) $T_x f\colon T_x M \to T_x N$ is C^k.

Exercise 6.5 Let $\phi_t\colon S^2 \to S^2$ rotate points on S^2 about a fixed axis through an angle $\psi(t)$. Show that ϕ_t is the flow of a certain vector field on S^2.

Exercise 6.6 Let $f\colon S^2 \to \mathbb{R}$ be defined by $f(x, y, z) = z$. Compute df using spherical coordinates (θ, ϕ).

Exercise 6.7 Compute the tangent lifts for the two stereographic projections of $S^2 \to \mathbb{R}^2$ in Example 5.14. That is, assuming (x, y, z) depend smoothly on t, find:

(1) How $(\dot\xi_N, \dot\eta_N)$ depend on $(\dot x, \dot y, \dot z)$. Likewise for $(\dot\xi_S, \dot\eta_S)$.

(2) 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).

7 Lie groups and Lie algebras

7.1 Matrix Lie groups

Definition 7.1 A *group* is a set of elements with:

(1) A binary product (multiplication), $G \times G \to G$, such that
- the product of g and h is written gh, and
- the product is associative: $(gh)k = g(hk)$.

(2) An identity element e such that $eg = g$ and $ge = g$, $\forall g \in G$

(3) An inverse operation $G \to G$, such that $gg^{-1} = g^{-1}g = e$

Definition 7.2 A *Lie group* is a smooth manifold G which is a group and for which the group operations of multiplication, $(g,h) \to gh$ for $g,h \in G$, and inversion, $g \to g^{-1}$ with $gg^{-1} = g^{-1}g = e$, are smooth.

Definition 7.3 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 $\mathbb{R}^{n\times n}$. 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 7.4 (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 matrices in $GL(n,\mathbb{R})$ is n^2.

Proposition 7.5 *Let $K \in GL(n,\mathbb{R})$ be a symmetric matrix, $K^T = K$. Then the subgroup S of $GL(n,\mathbb{R})$ defined by the mapping*

$$S = \{U \in GL(n,\mathbb{R}) | U^T K U = K\}$$

is a submanifold of $\mathbb{R}^{n\times n}$ of dimension $n(n-1)/2$.

Remark 7.6 The subgroup S leaves invariant a certain symmetric quadratic form under linear transformations, $S \times \mathbb{R}^n \to \mathbb{R}^n$ given by $\mathbf{x} \to U\mathbf{x}$, 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. Thus, S is the *isotropy subgroup* of the quadratic form associated with K.

Proof

- Is S a subgroup? We check the following three defining properties
  (1) Identity: $I \in S$ because $I^T K I = K$.
  (2) Inverse: $U \in S \Longrightarrow U^{-1} \in S$, because
  $$K = U^{-T}(U^T K U)U^{-1} = U^{-T}(K)U^{-1}.$$
  (3) Closed under multiplication: $U, V \in S \Longrightarrow UV \in S$, because
  $$(UV)^T KUV = V^T(U^T KU)V = V^T(K)V = K.$$
- Hence, S is a subgroup of $GL(n, \mathbb{R})$.
- Is S is a submanifold of $\mathbb{R}^{n\times n}$ of dimension $n(n-1)/2$?
  - Indeed, S is the zero locus of the mapping $UKU^T - K$. This makes it a submanifold, because it turns out to be a submersion.
  - 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 $n^2 - n(n+1)/2 = n(n-1)/2$. □

Exercise 7.7 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 linearization of this map.

Solution Here is why S is a submanifold of $R^{n\times n}$.

First, S is the zero locus of the mapping

$$U \to U^T KU - K, \qquad \text{(locus map)}$$

Let $U \in S$, and let δU be an arbitrary element of $R^{n\times n}$. Then linearize to find

$$(U + \delta U)^T K(U + \delta U) - K = U^T KU - K + \delta U^T KU + U^T K\delta U + O(\delta U)^2.$$

We may conclude that S is a submanifold of $R^{n\times n}$ if we can show that the linearization of the locus map, namely the linear mapping defined by

$$L \equiv \delta U \to \delta U^T KU + U^T K\delta U, \qquad R^{n\times n} \to R^{n\times n}$$

has constant rank for all $U \in S$.

Lemma 7.8 *The linearization map L is onto the space of $n \times n$ of symmetric matrices and hence the original map is a submersion.*

Proof that L is onto

- 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 S we can find δU such that

$$\delta U^T K U + U^T K \delta U = S.$$

Namely

$$\delta U = K^{-1} U^{-T} S/2.$$

- Thus, the linearization map L is onto the space of $n \times n$ of symmetric matrices and the original locus map $U \to U K U^T - K$ to the space of symmetric matrices is a submersion. □

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 $n^2 - n(n+1)/2 = n(n-1)/2$.

Corollary 7.9 (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 7.3, S is a matrix Lie group.*

Exercise 7.10 What is the tangent space to S at the identity, $T_I S$?

Exercise 7.11 Show that for any pair of matrices $A, B \in T_I S$, the matrix commutator $[A, B] \equiv AB - BA \in T_I S$.

Proposition 7.12 *The linear space of matrices A satisfying*

$$A^T K + KA = 0$$

defines $T_I S$, the tangent space at the identity of the matrix Lie group S defined in Proposition 7.5.

Proof Near the identity the defining condition for S expands to

$$(I + \epsilon A^T + O(\epsilon^2))K(I + \epsilon A + O(\epsilon^2)) = K, \quad \text{for} \quad \epsilon \ll 1.$$

At linear order $O(\epsilon)$ one finds,

$$A^T K + KA = 0.$$

This relation defines the linear space of matrices $A \in T_I S$. □

If $A, B \in T_I S$, does it follow that $[A, B] \in T_I S$?

Using $[A, B]^T = [B^T, A^T]$, we check closure by a direct computation,

$$\begin{aligned}[B^T, A^T]K + K[A, B] &= B^T A^T K - A^T B^T K + KAB - KBA \\ &= B^T A^T K - A^T B^T K - A^T KB + B^T KA = 0.\end{aligned}$$

Hence, the tangent space of S at the identity $T_I S$ is closed under the matrix commutator $[\cdot, \cdot]$.

Remark 7.13 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 $A^T K + KA = 0$ defines a matrix Lie algebra.

7.2 Defining matrix Lie algebras

We are ready to prove the following, in preparation for defining matrix Lie algebras.

Proposition 7.14 *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 7.15 *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 $R'(0) = B$. Define $S(t) = RR_B(t)R^{-1} \in T_I S$ 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 7.14 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 7.16 (Matrix commutator) For any pair of $n \times n$ matrices A, B, the *matrix commutator* is defined as $[A, B] = AB - BA$.

Proposition 7.17 (Properties of the matrix commutator) *The matrix commutator has the following two properties:*

(i) *Any two* $n \times n$ *matrices* A *and* B *satisfy*

$$[B, A] = -[A, B]$$

(This is the property of skew-symmetry.)

(ii) *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 7.18 (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 7.19 *For any matrix Lie group* S, *the tangent space at the identity* $T_I S$ *is a matrix Lie algebra.*

Proof This follows by Proposition 7.14 and because $T_I S$ is a vector space. □

7.3 Examples of matrix Lie groups

Example 7.20 (The Orthogonal Group $O(n)$) The mapping condition $U^T K U = K$ in Proposition 7.5 specializes for $K = I$ to $U^T U = I$, which defines the orthogonal group. Thus, in this case, S specializes to $O(n)$, the group of $n \times n$ orthogonal matrices. The orthogonal group is of special interest in mechanics.

Corollary 7.21 ($O(n)$ is a matrix Lie group) *By Proposition 7.5 the orthogonal group* $O(n)$ *is both a subgroup and a submanifold of the general linear group* $GL(n, \mathbb{R})$. *Thus, by Definition 7.3, the orthogonal group* $O(n)$ *is a matrix Lie group.*

Example 7.22 (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 7.23 (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 $U^T K U = K$ specializes to $U^T U = I$ with the extra condition $\det(U) = 1$.

Example 7.24 (The tangent space of $SO(n)$ at the identity $T_I SO(n)$) 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 a matrix Lie algebra.

Example 7.25 (The Symplectic Group) 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

$$Sp(l) = \{U \in GL(2l, \mathbb{R}) | U^T J U = J\}$$

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 $JA^T + AJ = 0$. This algebra is denoted as $sp(l)$.

Example 7.26 (The Special Euclidean Group) Consider the Lie group of 4×4 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 7.27 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?

Exercise 7.28 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 7.29 (Tripos question) When does a stone at the equator of the Earth weigh the most? Two hints: (1) Assume the Earth's orbit is a circle around the Sun and ignore the declination of the Earth's axis of rotation. (2) This is an exercise in using $SE(2)$.

Exercise 7.30 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.

7.4 Lie group actions

The action of a Lie group G on a manifold M is a group of transformations of M associated to elements of the group G, whose composition acting on M is corresponds to group multiplication in G.

Definition 7.31 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 $\Phi\colon G \times M \to M$ 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) above then simply reads $(gh)x = g(hx)$.

Similarly, one can define a *right action*, which is a map $\Psi\colon M \times G \to M$ 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 $\Psi(\Psi(x, g), h) = \Psi(x, gh)$ is then be expressed conveniently as $(xg)h = x(gh)$.

Example 7.32 (Properties of group actions) The action $\Phi\colon G \times M \to M$ of a group G on a manifold M is said to be

(1) *transitive*, if for every $x, y \in M$ there exists a $g \in G$, such that $gx = y$;

(2) *free*, if it has no fixed points, that is, $\Phi_g(x) = x$ implies $g = e$; and

(3) *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.

Orbits Given a group action of G on M, for a given point $x \in M$, the subset

$$\operatorname{Orb} x = \{gx \mid g \in G\} \subset M,$$

is called the *group orbit* through x. In finite dimensions, it can be shown that group orbits are always smooth (possibly immersed) manifolds. Group orbits generalize the notion of orbits of a dynamical system.

Exercise 7.33 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.

Theorem 7.34 *Orbits of proper group actions are embedded submanifolds.*

This theorem is stated by Marsden and Ratiu [44, Chapter 9], who refer to Abraham and Marsden [1] for the proof.

Example 7.35 (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 7.36 (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.

Left and right translations on the group are denoted L_g and R_g, respectively. For example, $L_g\colon G \to G$ is the map given by $h \to gh$, while $R_g\colon G \to G$ is the map given by $h \to hg$, for $g, h \in G$.

(a) *Left translation* $L_g\colon G \to G$; $h \to gh$ defines a transitive and free action of G on itself. Right multiplication $R_g\colon G \to G$; $h \to hg$ defines a right action, while $h \to hg^{-1}$ defines a left action of G on itself.

(b) G acts on G by conjugation, $g \to I_g = R_{g^{-1}} \circ L_g$. The map $I_g\colon G \to G$ given by $h \to ghg^{-1}$ is the *inner automorphism* associated with g. Orbits of this action are called *conjugacy classes*.

(c) Differentiating conjugation at e gives the *adjoint action* of G on $\mathfrak{g}$:

$$\mathrm{Ad}_g := T_e I_g \,:\, T_e G = \mathfrak{g} \to T_e G = \mathfrak{g}.$$

Explicitly, the *adjoint action* of G on $\mathfrak{g}$ is given by

$$\mathrm{Ad}\colon G \times \mathfrak{g} \to \mathfrak{g}, \quad \mathrm{Ad}_g(\xi) = T_e(R_{g^{-1}} \circ L_g)\xi$$

We have already seen an example of adjoint action for matrix Lie groups acting on matrix Lie algebras, when we defined $S(t) = R_A(t)BR_A(t)^{-1} \in T_I S$ as a key step in the proof of Proposition 7.14.

(d) The *coadjoint action* of G on $\mathfrak{g}^*$, the dual of the Lie algebra $\mathfrak{g}$ of G, is defined as follows. Let $\mathrm{Ad}^*_g\colon \mathfrak{g}^* \to \mathfrak{g}^*$ be the dual of Ad_g, defined by

$$\langle \mathrm{Ad}^*_g\, \alpha, \xi\rangle = \langle \alpha, \mathrm{Ad}_g\, \xi\rangle$$

for $\alpha \in \mathfrak{g}^*$, $\xi \in \mathfrak{g}$ and pairing $\langle\cdot,\cdot\rangle\colon \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$. Then the map

$$\Phi^*\colon G \times \mathfrak{g}^* \to \mathfrak{g}^* \quad \text{given by} \quad (g,\alpha) \mapsto \mathrm{Ad}^*_{g^{-1}}\, \alpha$$

is the coadjoint action of G on $\mathfrak{g}^*$.

7.5 Examples: $SO(3)$, $SE(3)$, etc

A basis for the matrix Lie algebra $so(3)$ and a map to $\mathbb{R}^3$

The Lie algebra of $SO(n)$ is called $so(n)$. A basis (e_1, e_2, e_3) for $so(3)$ when $n = 3$ is given by

$$\widehat{\mathbf{x}} = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} = xe_1 + ye_2 + ze_3$$

Exercise 7.37 Show that $[e_1, e_2] = e_3$ and cyclic permutations, while all other matrix commutators among the basis elements vanish.

Example 7.38 (The isomorphism between $so(3)$ and $\mathbb{R}^3$) The previous equation may be written equivalently by defining the hat-operation $\widehat{\cdot}$ as

$$\widehat{\mathbf{x}}_{ij} = \epsilon_{ijk} x^k, \quad \text{where} \quad (x^1, x^2, x^3) = (x, y, z).$$

Here $\epsilon_{123} = 1$ and $\epsilon_{213} = -1$, with cyclic permutations. The totally antisymmetric tensor $\epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{ikj}$ also defines the cross product of vectors in $\mathbb{R}^3$. Consequently, we may write,

$$(\mathbf{x} \times \mathbf{y})_i = \epsilon_{ijk} x^j y^k = \widehat{\mathbf{x}}_{ij} y^j, \quad \text{that is,} \quad \mathbf{x} \times \mathbf{y} = \widehat{\mathbf{x}}\mathbf{y}$$

Exercise 7.39 What is the analog of the hat map $so(3) \mapsto \mathbb{R}^3$ for the three dimensional Lie algebras $sp(2,\mathbb{R})$, $so(2,1)$, $su(1,1)$, or $sl(2,\mathbb{R})$?

Background reading for this lecture is Marsden and Ratiu [44, Chapter 9].

Compute the Adjoint and adjoint operations by differentiation

(1) Differentiate $I_g(h)$ with respect to h at $h = e$ to produce the *Adjoint operation*

$$\mathrm{Ad}\colon G \times \mathfrak{g} \to \mathfrak{g} : \quad \mathrm{Ad}_g\, \eta = T_e I_g \eta$$

(2) Differentiate $\mathrm{Ad}_g\, \eta$ with respect to g at $g = e$ in the direction ξ to get the *Lie bracket* $[\xi, \eta]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ and thereby to produce the *adjoint operation*

$$T_e(\mathrm{Ad}_g\, \eta)\xi = [\xi, \eta] = \mathrm{ad}_\xi\, \eta$$

Compute the co-Adjoint and coadjoint operations by taking duals

(1) $\mathrm{Ad}^*_g\colon \mathfrak{g}^* \to \mathfrak{g}^*$, the dual of Ad_g, is defined by

$$\langle \mathrm{Ad}^*_g\, \alpha, \xi \rangle = \langle \alpha, \mathrm{Ad}_g\, \xi \rangle$$

for $\alpha \in \mathfrak{g}^*$, $\xi \in \mathfrak{g}$ and pairing $\langle \cdot, \cdot \rangle\colon \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$. The map

$$\Phi^*\colon G \times \mathfrak{g}^* \to \mathfrak{g}^* \quad \text{given by} \quad (g, \alpha) \mapsto \mathrm{Ad}^*_{g^{-1}}\, \alpha$$

defines the *co-Adjoint action* of G on $\mathfrak{g}^*$.

(2) The pairing

$$\langle \mathrm{ad}^*_\xi\, \alpha, \eta \rangle = \langle \alpha, \mathrm{ad}_\xi\, \eta \rangle$$

defines the *coadjoint action* of $\mathfrak{g}$ on $\mathfrak{g}^*$, for $\alpha \in \mathfrak{g}^*$ and $\xi, \eta \in \mathfrak{g}$.

See [44, Chapter 9] for more discussion of the Ad and ad operations.

Example: the rotation group $SO(3)$

The Lie algebra $so(3)$ and its dual The special orthogonal group is defined by

$$SO(3) := \{A \mid A \text{ a } 3 \times 3 \text{ orthogonal matrix}, \det(A) = 1\}.$$

Its Lie algebra $so(3)$ is formed by 3×3 skew symmetric matrices, and its dual is denoted $so(3)^*$.

The isomorphism $\hat{\cdot}\colon (so(3), [\cdot, \cdot]) \to (\mathbb{R}^3, \times)$ The Lie algebra $(so(3), [\cdot, \cdot])$, where $[\cdot, \cdot]$ is the commutator bracket of matrices, is isomorphic to the Lie algebra $(\mathbb{R}^3, \times)$, where $\times$ denotes the vector product in $\mathbb{R}^3$, by the isomorphism

$$\mathbf{u} := (u^1, u^2, u^3) \in \mathbb{R}^3 \mapsto \hat{\mathbf{u}} := \begin{bmatrix} 0 & -u^3 & u^2 \\ u^3 & 0 & -u^1 \\ -u^2 & u^1 & 0 \end{bmatrix} \in so(3),$$

that is, $\widehat{\mathbf{u}}_{ij} := -\epsilon_{ijk} u^k$. Equivalently, this isomorphism is given by

$$\widehat{\mathbf{u}}\mathbf{v} = \mathbf{u} \times \mathbf{v} \quad \text{for all} \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^3.$$

The following formulas for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ may be easily verified:

$$\begin{aligned}(\mathbf{u} \times \mathbf{v})\widehat{\ } &= [\widehat{\mathbf{u}}, \widehat{\mathbf{v}}]\\ [\widehat{\mathbf{u}}, \widehat{\mathbf{v}}]\mathbf{w} &= (\mathbf{u} \times \mathbf{v}) \times \mathbf{w}\\ \mathbf{u} \cdot \mathbf{v} &= -\tfrac{1}{2}\operatorname{trace}(\widehat{\mathbf{u}}\widehat{\mathbf{v}}).\end{aligned}$$

The Ad **action of** $SO(3)$ **on** $so(3)$ The corresponding adjoint action of $SO(3)$ on $so(3)$ may be obtained as follows. For $SO(3)$ we have $I_A(B) = ABA^{-1}$. Differentiating $B(t)$ at $B(0) = \mathrm{Id}$ gives

$$\mathrm{Ad}_A \widehat{\mathbf{v}} = \frac{d}{dt}\bigg|_{t=0} AB(t)A^{-1} = A\widehat{\mathbf{v}}A^{-1}, \quad \text{with} \quad \widehat{\mathbf{v}} = B'(0).$$

One calculates the pairing with a vector $\mathbf{w} \in \mathbb{R}^3$ as

$$\mathrm{Ad}_A \widehat{\mathbf{v}}(\mathbf{w}) = A\widehat{\mathbf{v}}(A^{-1}\mathbf{w}) = A(\mathbf{v} \times A^{-1}\mathbf{w}) = A\mathbf{v} \times \mathbf{w} = (A\mathbf{v})\widehat{\ }\mathbf{w}$$

where we have used a 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,

$$\mathrm{Ad}_A \widehat{\mathbf{v}} = (A\mathbf{v})\widehat{\ }$$

Identifying $so(3) \simeq \mathbb{R}^3$ then gives

$$\mathrm{Ad}_A \mathbf{v} = A\mathbf{v}.$$

So (speaking prose all our lives) the adjoint action of $SO(3)$ on $so(3)$ may be identified with multiplication of a matrix in $SO(3)$ times a vector in $\mathbb{R}^3$.

The ad **action of** $so(2)$ **on** $so(3)$ Differentiating again gives the ad–action of the Lie algebra $so(3)$ on itself:

$$[\widehat{\mathbf{u}}, \widehat{\mathbf{v}}] = \mathrm{ad}_{\widehat{\mathbf{u}}} \widehat{\mathbf{v}} = \frac{d}{dt}\bigg|_{t=0} \left(e^{t\widehat{\mathbf{u}}}\mathbf{v}\right)\widehat{\ } = (\widehat{\mathbf{u}}\mathbf{v})\widehat{\ } = (\mathbf{u} \times \mathbf{v})\widehat{\ }.$$

So in this isomorphism the vector cross product is identified with the matrix commutator of skew symmetric matrices.

Infinitesimal generator Likewise, the *infinitesimal generator* corresponding to $\mathbf{u} \in \mathbb{R}^3$ has the expression

$$\mathbf{u}_{\mathbb{R}^3}(\mathbf{x}) := \left.\frac{d}{dt}\right|_{t=0} e^{t\hat{\mathbf{u}}}\mathbf{x} = \hat{\mathbf{u}}\,\mathbf{x} = \mathbf{u} \times \mathbf{x}.$$

Exercise 7.40 What is the analog of the hat map $so(3) \mapsto \mathbb{R}^3$ for the three dimensional Lie algebras $sp(2,\mathbb{R})$, $so(2,1)$, $su(1,1)$, or $sl(2,\mathbb{R})$?

The dual Lie algebra isomorphism $\tilde{}$: $so(3)^* \to \mathbb{R}^3$

Coadjoint actions The dual $so(3)^*$ is identified with $\mathbb{R}^3$ by the isomorphism

$$\in \mathbb{R}^3 \mapsto \tilde{} \in so(3)^* \colon \tilde{}(\hat{\mathbf{u}}) := \cdot\,\mathbf{u} \quad \text{for any} \quad \mathbf{u} \in \mathbb{R}^3.$$

In terms of this isomorphism, the co-Adjoint action of $SO(3)$ on $so(3)^*$ is given by

$$\mathrm{Ad}^*_{A^{-1}}\tilde{} = (A\mathbf{\Pi})^{\sim}$$

and the coadjoint action of $so(3)$ on $so(3)^*$ is given by

$$\mathrm{ad}^*_{\hat{\mathbf{u}}}\,\tilde{\mathbf{\Pi}} = (\mathbf{\Pi} \times \mathbf{u})^{\sim}. \tag{6}$$

Computing the co-Adjoint action of $SO(3)$ on $so(3)^*$ This is given by

$$\begin{aligned}\left(\mathrm{Ad}^*_{A^{-1}}\tilde{}\right)(\hat{\mathbf{u}}) &= \tilde{}\cdot \mathrm{Ad}_{A^{-1}}\,\hat{\mathbf{u}} = \tilde{}\cdot (A^{-1}\mathbf{u})\hat{} = \mathbf{\Pi}\cdot A^T\mathbf{u} \\ &= A\mathbf{\Pi}\cdot\mathbf{u} = (A\mathbf{\Pi})^{\sim}(\hat{\mathbf{u}}),\end{aligned}$$

that is, the co-Adjoint action of $SO(3)$ on $so(3)^*$ has the expression

$$\mathrm{Ad}^*_{A^{-1}}\tilde{} = (A\mathbf{\Pi})^{\sim},$$

Therefore, the co-Adjoint 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 2–sphere of radius $\|\mathbf{\Pi}\|$.

Computing the coadjoint action of $so(3)$ on $so(3)^*$ Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ and note that

$$\begin{aligned}\left\langle \mathrm{ad}^*_{\hat{\mathbf{u}}}\,\tilde{\mathbf{\Pi}}, \hat{\mathbf{v}}\right\rangle &= \left\langle \tilde{\mathbf{\Pi}}, [\hat{\mathbf{u}}, \hat{\mathbf{v}}]\right\rangle = \left\langle \tilde{\mathbf{\Pi}}, (\mathbf{u}\times\mathbf{v})\hat{}\right\rangle = \mathbf{\Pi}\cdot(\mathbf{u}\times\mathbf{v}) \\ &= (\mathbf{\Pi}\times\mathbf{u})\cdot\mathbf{v} = \left\langle(\mathbf{\Pi}\times\mathbf{u})^{\sim}, \hat{\mathbf{v}}\right\rangle,\end{aligned}$$

which shows that $\mathrm{ad}^*_{\hat{\mathbf{u}}}\,\tilde{\mathbf{\Pi}} = (\mathbf{\Pi}\times\mathbf{u})^{\sim}$, thereby proving (6). Therefore

$$T_{\mathbf{\Pi}}\mathcal{O} = \left\{\mathbf{\Pi}\times\mathbf{u} \;\mid\; \mathbf{u}\in\mathbb{R}^3\right\},$$

since the plane perpendicular to $\mathbf{\Pi}$, that is, the tangent space to the sphere centered at the origin of radius $\|\mathbf{\Pi}\|$, is given by $\left\{\mathbf{\Pi}\times\mathbf{u} \mid \mathbf{u}\in\mathbb{R}^3\right\}$.

8 Lifted actions

Definition 8.1 Let $\Phi\colon G \times M \to M$ be a left action, and write $\Phi_g(x) = \Phi(g,x)$ for $x \in M$. The *tangent lift action* of G on the tangent bundle TM is defined by $gv = T_x\Phi_g(v)$ for every $v \in T_xM$.

Remark 8.2 In standard calculus notation, the expression for tangent lift may be written as

$$T_x\Phi \cdot v = \frac{d}{dt}\Phi(c(t))\Big|_{t=0} = \frac{\partial \Phi}{\partial c}c'(t)\Big|_{t=0} =: D\Phi(x)\cdot v,$$

with $c(0) = x$ and $c'(0) = v$.

Definition 8.3 If X is a vector field on M and ϕ is a differentiable map from M to itself, then the *push-forward* of X by ϕ is the vector field ϕ_*X defined by $(\phi_*X)(\phi(x)) = T_x\phi\,(X(x))$. That is, the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "T\phi"] & TM \\ M \arrow[u, "X"] \arrow[r, "\phi"] & M \arrow[u, "\phi_* X"'] \end{tikzcd}$$

If ϕ is a diffeomorphism then the *pull-back* ϕ^*X is also defined: $(\phi^*X)(x) = T_{\phi(x)}\phi^{-1}\,(X\,(\phi(x)))$.

Definition 8.4 Let $\Phi\colon G \times M \to M$ be a left action, and write $\Phi_g(m) = \Phi(g,m)$. Then G has a left action on $X \in \mathfrak{X}(M)$ (the set of vector fields on M) by the push-forward: $gX = \left(\Phi_g\right)_* X$.

Definition 8.5 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 $x \in X$.

Definition 8.6 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 8.5, 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 $\mathfrak{X}^L(G)$.

Proposition 8.7 *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_ξ^L is the unique left-invariant vector field such that $X_\xi^L(e) = \xi$.*

Proof To show that X_ξ^L is left-invariant, we need to show that $g\big(X_\xi^L(h)\big) = X_\xi^L(gh)$ for every $g, h \in G$. This follows from the definition of X_ξ^L and the associativity property of group actions:

$$g\big(X_\xi^L(h)\big) = g(h\xi) = (gh)\xi = X_\xi^L(gh)$$

We repeat the last line in less compact notation:

$$T_h L_g\big(X_\xi^L(h)\big) = T_h L_g(h\xi) = T_e L_{gh}\xi = X_\xi^L(gh)$$

For uniqueness, suppose X is left-invariant and $X(e) = \xi$. Then for any $g \in G$, we have $X(g) = g(X(e)) = g\xi = X_\xi^L(g)$. □

Remark 8.8 Note that the map $\xi \mapsto X_\xi^L$ is an vector space isomorphism from $T_e G$ to $\mathfrak{X}^L(G)$.

All of the above definitions have analogues for right actions. The definitions of *right-invariant*, $\mathfrak{X}^R(G)$ and X_ξ^R use the right action of G on itself defined by $\Phi(g, h) = R_g(h) = hg$.

Exercise 8.9 There is a left action of G on itself defined by $\Phi_g(h) = hg^{-1}$.

We will use the map $\xi \mapsto X_\xi^L$ to relate the Lie bracket on $\mathfrak{g}$, defined as $[\xi, \eta] = \mathrm{ad}_\xi\, \eta$, with the Jacobi–Lie bracket on vector fields.

Definition 8.10 The *Jacobi–Lie bracket* on $\mathfrak{X}(M)$ is defined in local coordinates by

$$[X, Y]_{\text{J–L}} \equiv (DX)\cdot Y - (DY)\cdot X$$

which, in finite dimensions, is equivalent to

$$[X, Y]_{\text{J–L}} \equiv -(X\cdot\nabla)Y + (Y\cdot\nabla)X \equiv -[X, Y]$$

Theorem 8.11 (Properties of the Jacobi–Lie bracket)

(1) *The Jacobi–Lie bracket satisfies*

$$[X,Y]_{J\text{-}L} = \mathcal{L}_X Y \equiv \frac{d}{dt}\Big|_{t=0} \Phi_t^* Y,$$

where Φ is the flow of X. (This is coordinate-free, and can be used as an alternative definition.)

(2) *This bracket makes $\mathfrak{X}^L(M)$ a Lie algebra with $[X,Y]_{J\text{-}L} = -[X,Y]$, where $[X,Y]$ is the Lie algebra bracket on $\mathfrak{X}(M)$.*

(3) *$\phi_*[X,Y] = [\phi_* X, \phi_* Y]$ for any differentiable $\phi\colon M \to M$.*

Remark 8.12 The first property of the Jacobi–Lie bracket is proved for matrices in Section 9. The other two properties are proved below for the case that M is the Lie group G.

Theorem 8.13 *$\mathfrak{X}^L(G)$ 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

$$\left(L_g\right)_* [X,Y]_{\text{J–L}} = [\left(L_g\right)_* X, \left(L_g\right)_* Y]_{\text{J–L}}$$

for all $g \in G$. Hence $[X,Y]_{\text{J–L}} \in \mathfrak{X}^L(G)$. This is the second property in Theorem 8.11. □

Theorem 8.14 *Set $\left[X_\xi^L, X_\eta^L\right]_{J\text{-}L}(e) = [\xi,\eta]$ for every $\xi,\eta \in \mathfrak{g}$, where the bracket on the right is the Jacobi–Lie bracket. (We say: 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 $\mathrm{ad}_A B = AB - BA$. 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

$$\begin{aligned}[A,B] = \left[X_A^L, X_B^L\right]_{\text{J–L}}(I) &= DX_B^L(I)\cdot X_A^L(I) - DX_A^L(I)\cdot X_B^L(I)\\ &= DX_B^L(I)\cdot A - DX_A^L(I)\cdot B = AB - BA\end{aligned}$$

This is the third property of the Jacobi–Lie bracket listed in Theorem 8.11. For the general proof, see Marsden and Ratiu [44, Proposition 9.14]. □

Remark 8.15 This theorem, together with Item 2 in Theorem 8.11, proves that the Jacobi–Lie bracket makes $\mathfrak{g}$ into a Lie algebra.

Remark 8.16 By Theorem 8.13, the vector field $\left[X_\xi^L, X_\eta^L\right]$ is left-invariant. Since $\left[X_\xi^L, X_\eta^L\right]_{\text{J–L}}(e) = [\xi, \eta]$, it follows that

$$\left[X_\xi^L, X_\eta^L\right] = X_{[\xi,\eta]}^L.$$

Definition 8.17 Let $\Phi\colon G \times M \to M$ 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 ξ direction is the vector field ξ_M on M defined by

$$\xi_M(x) = \frac{d}{dt}\bigg|_{t=0} \Phi_{g(t)}(x)$$

Remark 8.18 Note: this definition does not depend on the choice of $g(t)$. For example, the choice in Marsden and Ratiu [44] is $\exp(t\xi)$, where exp denotes the exponentiation on Lie groups (not defined here).

Exercise 8.19 Consider the action of $SO(3)$ on the unit sphere S^2 around the origin, and let $\xi = (0,0,1)\hat{}$. Sketch the vector field ξ_M. (Hint: the vectors all point "Eastward.")

Theorem 8.20 *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 Marsden and Ratiu [44, Proposition 9.3.6].

Exercise 8.21 Express the statements and formulas of this lecture for the case of $SO(3)$ action on its Lie algebra $so(3)$. (Hint: look at the previous lecture.) Wherever possible, translate these formulas to $\mathbb{R}^3$ by using the $\hat{}$ map: $so(3) \to \mathbb{R}^3$.

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?

9 Handout: the Lie derivative and the Jacobi–Lie bracket

Let X and Y be two vector fields on the same manifold M.

Definition 9.1 The *Lie derivative* of Y with respect to X is $\mathcal{L}_X Y \equiv \frac{d}{dt}\Phi_t^* Y\big|_{t=0}$, where Φ is the flow of X.

The Lie derivative $\mathcal{L}_X Y$ is "the derivative of Y in the direction given by X." Its definition is coordinate-independent. By contrast, $DY \cdot X$ (also written as $X[Y]$) is also "the derivative of Y in the X direction", but the value of $DY \cdot X$ depends on the coordinate system, and in particular does not usually equal $\mathcal{L}_X Y$ in the chosen coordinate system.

Theorem 9.2 $\mathcal{L}_X Y = [X, Y]$, *where the bracket on the right is the* Jacobi–Lie bracket.

Proof In the following calculation, we assume that M is finite-dimensional, and we work in local coordinates. Thus we may consider everything as matrices, which allows us to use the product rule and the identities $\left(M^{-1}\right)' = -M^{-1} M' M^{-1}$ and $\frac{d}{dt}\big(D\Phi_t(x)\big) = D\big(\frac{d}{dt}\Phi_t\big)(x)$.

$$
\begin{aligned}
\mathcal{L}_X Y(x) &= \frac{d}{dt}\Phi_t^* Y(x)\Big|_{t=0} \\
&= \frac{d}{dt}\big(D\Phi_t(x)\big)^{-1} Y\big(\Phi_t(x)\big)\Big|_{t=0} \\
&= \left[\left(\frac{d}{dt}\left(D\Phi_t(x)\right)^{-1}\right) Y\big(\Phi_t(x)\big) + \big(D\Phi_t(x)\big)^{-1}\frac{d}{dt} Y\big(\Phi_t(x)\big)\right]_{t=0} \\
&= \Bigg[-\big(D\Phi_t(x)\big)^{-1}\left(\frac{d}{dt} D\Phi_t(x)\right)\big(D\Phi_t(x)\big)^{-1} Y\big(\Phi_t(x)\big) \\
&\qquad\qquad + \big(D\Phi_t(x)\big)^{-1}\frac{d}{dt} Y\big(\Phi_t(x)\big)\Bigg]_{t=0} \\
&= \left[-\left(\frac{d}{dt} D\Phi_t(x)\right) Y(x) + \frac{d}{dt} Y\big(\Phi_t(x)\big)\right]_{t=0} \\
&= -D\left(\frac{d}{dt}\Phi_t(x)\Big|_{t=0}\right) Y(x) + DY(x)\left(\frac{d}{dt}\Phi_t(x)\Big|_{t=0}\right) \\
&= -DX(x)\cdot Y(x) + DY(x)\cdot X(x) \\
&= [X, Y]_{\text{J–L}}(x)
\end{aligned}
$$

Therefore $\mathcal{L}_X Y = [X, Y]_{\text{J–L}}$. □

Vorticity dynamics

The same formula applies in infinite dimensions, although the proof is more elaborate. For example, the equation for the vorticity dynamics of an Euler fluid with velocity **u** (with $\operatorname{div}\mathbf{u}=0$) and vorticity $\omega=\operatorname{curl}\mathbf{u}$ may be written as

$$\begin{aligned}\partial_t\omega &= -\mathbf{u}\cdot\nabla\omega+\omega\cdot\nabla\mathbf{u}\\ &= -[u,\omega]\\ &= -\operatorname{ad}_u\omega\\ &= -\mathcal{L}_u\omega\end{aligned}$$

All of these equations express the invariance of the vorticity vector field ω under the flow of its corresponding divergenceless velocity vector field u. This is also encapsulated in the language of fluid dynamics in *characteristic form* as

$$\frac{d}{dt}\Big(\omega\cdot\frac{\partial}{\partial\mathbf{x}}\Big)=0,\quad \text{along}\quad \frac{d\mathbf{x}}{dt}=\mathbf{u}(\mathbf{x},t)=\operatorname{curl}^{-1}\omega.$$

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

$$\operatorname{curl}\operatorname{curl}\mathbf{u}=-\Delta\mathbf{u}+\nabla\operatorname{div}\mathbf{u},$$

and application of $\operatorname{div}\mathbf{u}=0$. Thus, in coordinates,

$$\frac{d\mathbf{x}}{dt}=\mathbf{u}(\mathbf{x},t)\quad\Longrightarrow\quad \mathbf{x}(t,\mathbf{x}_0)=\Phi_t\mathbf{x}_0$$

with $\Phi_0=\mathrm{Id}$, that is, $\mathbf{x}(0,\mathbf{x}_0)=\mathbf{x}_0$ at $t=0$, and

$$\omega^j(\mathbf{x}(t,\mathbf{x}_0),t)\frac{\partial}{\partial x^j(t,\mathbf{x}_0)}=\omega^A(\mathbf{x}_0)\frac{\partial}{\partial x_0^A}\circ\Phi_t^{-1}.$$

Consequently,

$$\Phi_{t*}\omega^j(\mathbf{x}(t,\mathbf{x}_0),t)=\omega^A(\mathbf{x}_0)\frac{\partial x^j(t,\mathbf{x}_0)}{\partial x_0^A}=:D\Phi_t\cdot\omega.$$

This is the Cauchy (1859) solution of *Euler's equation for vorticity*,

$$\frac{\partial\omega}{\partial t}=[\omega,\operatorname{curl}^{-1}\omega].$$

This type of equation will reappear several more times in the remaining lectures. In it, the vorticity ω evolves by the ad-action of the right-invariant vector field $\mathbf{u}=\operatorname{curl}^{-1}\omega$. That is,

$$\frac{\partial\omega}{\partial t}=-\operatorname{ad}_{\operatorname{curl}^{-1}\omega}\omega.$$

The Cauchy solution is the tangent lift of this flow, namely,

$$\Phi_{t*}\omega(\Phi_t(\mathbf{x}_0)) = T_{\mathbf{x}_0}\Phi_t(\omega(\mathbf{x}_0)).$$

10 Handout: summary of Euler's equations for incompressible flow

Euler's equation of incompressible fluid motion

$$\underbrace{\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u}}_{\substack{d\mathbf{u}/dt \text{ along } d\mathbf{x}/dt=\mathbf{u} \\ \text{(advective time derivative)}}} + \nabla p = 0$$

where $\mathbf{u}\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ satisfies $\operatorname{div}\mathbf{u} = 0$.

Geometric dynamics of vorticity

$$\begin{aligned}
\boldsymbol{\omega} &= \operatorname{curl}\mathbf{u} \\
\boldsymbol{\omega}_t &= -\mathbf{u}\cdot\nabla\boldsymbol{\omega} + \boldsymbol{\omega}\cdot\nabla\mathbf{u} \\
&= -[\mathbf{u},\boldsymbol{\omega}] \\
&= -\mathrm{ad}_{\mathbf{u}}\boldsymbol{\omega} \\
&= -\mathcal{L}_{\mathbf{u}}\boldsymbol{\omega}
\end{aligned}$$

In these equations, one denotes $\frac{d}{dt} = \frac{\partial}{\partial t} + \mathcal{L}_{\mathbf{u}}$ and, hence, may write Euler vorticity dynamics equivalently in any of the following three forms

$$\frac{d\boldsymbol{\omega}}{dt} = \boldsymbol{\omega}\cdot\nabla\mathbf{u},$$

as well as

$$\left(\frac{\partial}{\partial t} + \mathcal{L}_{\mathbf{u}}\right)\left(\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}}\right) = 0$$

or

$$\frac{d}{dt}(\boldsymbol{\omega}\cdot\nabla) = 0 \quad \text{along} \quad \frac{d\mathbf{x}}{dt} = \mathbf{u}$$

The last form is found using the chain rule as

$$\frac{d}{dt}(\boldsymbol{\omega}\cdot\nabla) = \frac{d\boldsymbol{\omega}}{dt}\cdot\nabla + \boldsymbol{\omega}\cdot\frac{d}{dt}\nabla = \left(\frac{d}{dt}\boldsymbol{\omega} - \boldsymbol{\omega}\cdot\nabla\mathbf{u}\right)\cdot\nabla = 0.$$

Ertel's theorem [18]

The operators d/dt and $\boldsymbol{\omega}\cdot\nabla$ commute on solutions of Euler's fluid equations. That is,

$$\left[\frac{d}{dt}, \boldsymbol{\omega}\cdot\nabla\right] = 0,$$

so that

$$\frac{d}{dt}(\boldsymbol{\omega}\cdot\nabla\mathbf{A}) = \boldsymbol{\omega}\cdot\nabla\frac{d}{dt}\mathbf{A}$$

for all differentiable $\mathbf{A}$ when $\boldsymbol{\omega} = \operatorname{curl}\mathbf{u}$ and $\mathbf{u}$ is a solution of Euler's equations for incompressible fluid flow. Consequently, one finds the following infinite set of *conservation laws*:

$$\text{If}\quad \frac{d\mathbf{A}}{dt} = 0, \quad \text{then} \quad \int \Phi(\boldsymbol{\omega}\cdot\nabla\mathbf{A})\,d^3\mathbf{x} = \text{const} \quad \text{for all differentiable } \Phi$$

Ohktani's formula [52]

$$\frac{d^2\boldsymbol{\omega}}{dt^2} = \frac{d}{dt}(\boldsymbol{\omega}\cdot\nabla\mathbf{u}) = \boldsymbol{\omega}\cdot\nabla\frac{d\mathbf{u}}{dt} = -\boldsymbol{\omega}\cdot\nabla\nabla p = -\mathbb{P}\boldsymbol{\omega}$$

where

$$\mathbb{P}_{ij} = \frac{\partial^2 p}{\partial x^i \partial x^j} \qquad \text{(“Hessian” of pressure).}$$

In addition, one has the relations

$$\begin{aligned} p &= -\Delta^{-1}\operatorname{tr}\left(\nabla\mathbf{u}^T\cdot\nabla\mathbf{u}\right) \\ S &= \tfrac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T) \qquad \text{(strain rate tensor)} \end{aligned}$$

so that, the following system of equations results,

$$\frac{d\boldsymbol{\omega}}{dt} = S\boldsymbol{\omega} \qquad\qquad \frac{d^2\boldsymbol{\omega}}{dt^2} = -\mathbb{P}\boldsymbol{\omega}$$

Kelvin (1890's) circulation theorem

$$\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}} = \omega^j\frac{\partial}{\partial x^j} \qquad\qquad \frac{d\mathbf{u}}{dt} + \nabla p = 0,$$

where $\operatorname{div}\mathbf{u} = 0$, or equivalently $u^j_{,j} = 0$ in index notation.

The motion equation may be rewritten equivalently as a 1–form relation,

$$\frac{du_i}{dt}\,dx^i = -dp = \nabla_i p\,dx^i \qquad \text{along} \quad \frac{d\mathbf{x}}{dt} = \mathbf{u}$$

$$\frac{d}{dt}(u_i\,dx^i) - \underbrace{u_i\frac{d}{dt}dx^i}_{=u_i\,du^i = d|\mathbf{u}|^2/2} = -dp$$

Consequently

$$\frac{d}{dt}(\mathbf{u}\cdot d\mathbf{x}) = -d\big(p - \tfrac{1}{2}|\mathbf{u}|^2\big)$$

which becomes

$$\frac{d}{dt}\oint_{C(\mathbf{u})}\mathbf{u}\cdot d\mathbf{x} = -\oint_{C(\mathbf{u})} d\big(p - \tfrac{1}{2}\mathbf{u}^2\big) = 0$$

upon integrating around a closed loop $C(\mathbf{u})$ moving with velocity $\mathbf{u}$ The 1–form relation above may be rewritten as

$$(\partial_t + \mathcal{L}_{\mathbf{u}})(\mathbf{u}\cdot d\mathbf{x}) = -d\big(p - \tfrac{1}{2}\mathbf{u}^2\big)$$

whose exterior derivative yields using $d^2 = 0$

$$(\partial_t + \mathcal{L}_{\mathbf{u}})(\boldsymbol{\omega}\cdot d\mathbf{S}) = 0$$

where $\boldsymbol{\omega}\cdot d\mathbf{S} = \operatorname{curl}\mathbf{u}\cdot d\mathbf{S} = d(\mathbf{u}\cdot d\mathbf{x})$.

For these geometric quantities, one sees that the characteristic, or advective derivative is equivalent to a Lie derivative. Namely,

$$\underbrace{\frac{d}{dt}\bigg|_{\text{advect}}}_{\text{fluids}} = \underbrace{\partial_t + \mathcal{L}_{\mathbf{u}}}_{\text{geometry}}$$

Stokes theorem

The classical theorem due to Stokes

$$\oint_{\partial S}\mathbf{u}\cdot d\mathbf{x} = \iint_S \operatorname{curl}\mathbf{u}\cdot d\mathbf{S}$$

shows that Kelvin's circulation theorem is equivalent to conservation of flux of vorticity

$$\frac{d}{dt}\iint_S \boldsymbol{\omega}\cdot d\mathbf{S} = 0,$$

with $\partial S = C(\mathbf{u})$ through any surface comoving with the flow.

Recall the definition,

$$\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}} \lrcorner\, d^3x = \boldsymbol{\omega}\cdot d\mathbf{S}$$

One may check this formula directly, by computing

$$\begin{aligned}\Big(\omega^1\frac{\partial}{\partial x^1}+\omega^2\frac{\partial}{\partial x^2}+\omega^3\frac{\partial}{\partial x^3}\Big) &\lrcorner\,\big(dx^1\wedge dx^2\wedge dx^3\big)\\ &= \omega^1 dx^2\wedge dx^3+\omega^2 dx^3\wedge dx^1\omega dx^1\wedge dx^2\\ &= \boldsymbol{\omega}\cdot d\mathbf{S}\end{aligned}$$

One may then use the vorticity equation in vector-field form,

$$(\partial_t+\mathcal{L}_{\mathbf{u}})\,\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}}=0$$

to prove that the flux of vorticity through any comoving surface is conserved, as follows.

$$\begin{aligned}&(\partial_t+\mathcal{L}_{\mathbf{u}})\Big(\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}}\lrcorner\, d^3x\Big)\\ &\qquad=\Big(\underbrace{(\partial_t+\mathcal{L}_{\mathbf{u}})\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}}}_{=0}\Big)\lrcorner\, d^3x+\boldsymbol{\omega}\cdot\frac{\partial}{\partial\mathbf{x}}\lrcorner\;\underbrace{(\partial_t+\mathcal{L}_{\mathbf{u}})d^3x}_{=\operatorname{div}\mathbf{u}\,d^3x=0}\;=0\end{aligned}$$

That is, as computed above using the exterior derivative

$$(\partial_t+\mathcal{L}_{\mathbf{u}})\,\boldsymbol{\omega}\cdot d\mathbf{S}=0.$$

Momentum conservation

From Euler's fluid equation $du_i/dt+\nabla_i p=0$ with $u^j_{\ ,j}=0$ one finds,

$$\begin{aligned}\int\big(\partial_t u_i+u^j\partial_j u_i+\partial_i p\big)d^3x&=0\\ &=\frac{d}{dt}\int u_i\,d^3x+\int\partial_j\big(u_iu^j+p\,\delta_i^j\big)d^3x\\ &=\underbrace{\frac{d}{dt}M_i}_{=0}+\underbrace{\oint\widehat{n}_j\big(u_iu^j+p\delta_i^j\big)dS}_{=0,\text{ if }\widehat{\mathbf{n}}\cdot\mathbf{u}=0}\end{aligned}$$

Local conservation of fluid momentum is expressed using differentiation by parts as

$$\partial_t u_i=-\partial_j T_i^{\ j}$$

where $T_i^{\ j}:=u_i\,u^j+p\,\delta_i^j$ is the *fluid stress tensor*.

Moreover, each component of the total momentum $M_i = \int u_i \, d^3x$ for $i = 1, 2, 3$, is conserved for an incompressible Euler flow, provided the flow is tangential to any fixed boundaries, that is, $\widehat{\mathbf{n}} \cdot \mathbf{u} = 0$.

Mass conservation

For mass density $D(\mathbf{x}, t)$ with total mass $\int D(\mathbf{x}, t) \, d^3x$, along $d\mathbf{x}/dt = \mathbf{u}(x, t)$ one finds,

$$\frac{d}{dt} D \, d^3x = (\partial_t + \mathcal{L}_{\mathbf{u}})(D d^3x) = \underbrace{(\partial_t D + \operatorname{div} D\mathbf{u})}_{\text{continuity equation}} \; d^3x = 0$$

The solution of this equation is written in Lagrangian form as

$$(D \, d^3x) \cdot g^{-1}(t) = D(\mathbf{x}_0) \, d^3x$$

For incompressible flow, this becomes

$$\frac{1}{D} = \det \frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} = \frac{d^3x}{d^3x_0} = 1$$

Likewise, in the Eulerian representation one finds the equivalent relations,

$$\left. \begin{array}{c} D = 1 \\ \partial_t D + \operatorname{div}(D\mathbf{u}) = 0 \end{array} \right\} \Longrightarrow \operatorname{div} \mathbf{u} = 0$$

Energy conservation

Euler's fluid equation for incompressible flow $\operatorname{div} \mathbf{u} = 0$

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = 0$$

conserves the total kinetic energy, defined by

$$KE = \int \tfrac{1}{2} |\mathbf{u}|^2 d^3x$$

The vector calculus identity

$$\mathbf{u} \cdot \nabla \mathbf{u} = -\mathbf{u} \times \operatorname{curl} \mathbf{u} + \tfrac{1}{2} \nabla |\mathbf{u}|^2$$

recasts Euler's equation as

$$\partial_t \mathbf{u} - \mathbf{u} \times \operatorname{curl} \mathbf{u} + \nabla \big(p + \tfrac{1}{2} \mathbf{u}^2 \big) = 0$$

So that

$$\frac{\partial}{\partial t} \tfrac{1}{2} |\mathbf{u}|^2 + \operatorname{div} \big(p + \tfrac{1}{2} |\mathbf{u}|^2 \big) \mathbf{u} = 0.$$

Consequently,

$$\frac{d}{dt}\int_{\Omega} \tfrac{1}{2}|\mathbf{u}|^2 d^3x = -\oint_{\partial\Omega} \left(p + \tfrac{1}{2}|\mathbf{u}|^2\right)\mathbf{u}\cdot d\mathbf{S} = 0$$

since $\mathbf{u}\cdot d\mathbf{S} = \mathbf{u}\cdot\widehat{\mathbf{n}}\,dS = 0$ on any fixed boundary and one finds

$$KE = \int \tfrac{1}{2}|\mathbf{u}|^2 d^3x = \text{const}$$

for Euler fluid motion.

11 Lie group action on its tangent bundle

Definition 11.1 A Lie group G acts on its tangent bundle TG by tangent lifts.[3] Given $X \in T_hG$ we can consider the action of G on X by either left or right translations, denoted as T_hL_gX or T_hR_gX, respectively. These expressions may be abbreviated as

$$T_hL_gX = L_g^*X = gX \qquad \text{and} \qquad T_hR_gX = R_g^*X = Xg.$$

Left action of a Lie group G on its tangent bundle TG is illustrated in the figure below.

$$\begin{tikzcd} TG \arrow[r, "TL_g"] & TG \\ G \arrow[u, "X"] \arrow[r, "L_g"] & G \arrow[u, "gX"'] \end{tikzcd}$$

For matrix Lie groups, this action is just multiplication on the left or right, respectively.

Left- and right-invariant vector fields A vector field X on G is called left-invariant, if for every $g \in G$ one has $L_g^*X = X$, that is, if

$$(T_hL_g)X(h) = X(gh)$$

for every $h \in G$. The commutative diagram for a left-invariant vector field is illustrated in the figure below.

[3] Recall Definition 6.3 of tangent lifts of a differentiable manifold.

$$\begin{tikzcd} TG \arrow[r, "TL_g"] & TG \\ G \arrow[u, "X"] \arrow[r, "L_g"] & G \arrow[u, "X"'] \end{tikzcd}$$

Proposition 11.2 *The set $\mathfrak{X}_L(G)$ of left invariant vector fields on the Lie group G is a subalgebra of $\mathfrak{X}(G)$ the set of all vector fields on G.*

Proof If $X, Y \in \mathfrak{X}_L(G)$ and $g \in G$, then

$$L_g^*[X,Y] = \left[L_g^* X, L_g^* Y\right] = [X,Y].$$

Consequently, the Lie bracket $[X,Y] \in \mathfrak{X}_L(G)$. Therefore, $\mathfrak{X}_L(G)$ is a subalgebra of $\mathfrak{X}(G)$, the set of all vector fields on G. □

Proposition 11.3 *The linear maps $\mathfrak{X}_L(G)$ and T_eG are isomorphic as vector spaces.*

Demonstration of proposition. For each $\xi \in T_eG$, define a vector field X_ξ on G by letting $X_\xi(g) = T_eL_g(\xi)$. Then

$$\begin{aligned} X_\xi(gh) &= T_eL_{gh}(\xi) = T_e(L_g \circ L_h)(\xi) \\ &= T_hL_g(T_eL_h(\xi)) = T_hL_g(X_\xi(h)), \end{aligned}$$

which shows that X_ξ is left invariant. (This proposition is stated by Marsden and Ratiu [44, Chapter 9], who refer to Abraham and Marsden [1] for the full proof.)

Definition 11.4 (*Jacobi–Lie bracket of vector fields*) Let $g(t)$ and $h(s)$ be curves in G with $g(0) = e$, $h(0) = e$ and define vector fields at the identity of G by the tangent vectors $g'(0) = \xi$, $h'(0) = \eta$. Compute the linearization of the Adjoint action of G on T_eG as

$$[\xi, \eta] := \frac{d}{dt}\frac{d}{ds} g(t)h(s)g(t)^{-1}\Big|_{s=0,t=0} = \frac{d}{dt} g(t)\eta g(t)^{-1}\Big|_{t=0} = \xi\eta - \eta\xi.$$

This is the *Jacobi–Lie bracket* of the vector fields ξ and η.

Definition 11.5 The *Lie bracket* in T_eG is defined by

$$[\xi, \eta] := [X_\xi, X_\eta](e),$$

for $\xi, \eta \in T_eG$ and for $[X_\xi, X_\eta]$ the Jacobi–Lie bracket of vector fields. This makes T_eG into a Lie algebra. Note that

$$[X_\xi, X_\eta] = X_{[\xi,\eta]},$$

for all $\xi, \eta \in T_eG$.

Definition 11.6 The vector space T_eG with this Lie algebra structure is called the *Lie algebra of* G and is denoted by $\mathfrak{g}$.

If we let $\xi_L(g) = T_eL_g\xi$, then the Jacobi–Lie bracket of two such left-invariant vector fields in fact gives the Lie algebra bracket:

$$[\xi_L, \eta_L](g) = [\xi, \eta]_L(g)$$

For the right-invariant case, the right hand side obtains a minus sign,

$$[\xi_R, \eta_R](g) = -[\xi, \eta]_R(g).$$

The relative minus sign arises because of the difference in action $(xh^{-1})g^{-1} = x(gh)^{-1}$ on the right versus $(gh)x = g(hx)$ on the left.

Infinitesimal generators In mechanics, group actions often appear as symmetry transformations, which arise through their infinitesimal generators, defined as follows.

Definition 11.7 Suppose $\Phi\colon G\times M \to M$ is an action. For $\xi \in \mathfrak{g}$, $\Phi^\xi(t,x)\colon \mathbb{R}\times M \to M$ defined by $\Phi^\xi(x) = \Phi(\exp t\xi, x) = \Phi_{\exp t\xi}(x)$ is an $\mathbb{R}$–action on M. In other words, $\Phi_{\exp t\xi} \to M$ is a flow on M. The vector field on M defined by[4]

$$\xi_M(x) = \frac{d}{dt}\bigg|_{t=0} \Phi_{\exp t\xi}(x)$$

is called the *infinitesimal generator* of the action corresponding to ξ.

The Jacobi–Lie bracket of infinitesimal generators is related to the Lie algebra bracket as follows:

$$[\xi_M, \eta_M] = -[\xi, \eta]_M$$

See, for example, Marsden and Ratiu [44, Chapter 9] for the proof.

[4] Recall Definition 5.22 of vector fields.

12 Lie algebras as vector fields

Definition 12.1 (The ad–operation) For $A \in \mathfrak{g}$ we define the operator ad_A to be the operator ad: $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that maps $B \in \mathfrak{g}$ to $[A, B]$. We write $\mathrm{ad}_A B = [A, B]$.

Definition 12.2 A *representation* of a Lie algebra $\mathfrak{g}$ on a vector space V is a mapping ρ from $\mathfrak{g}$ to the linear transformations of V such that for $A, B \in \mathfrak{g}$ and any constant scalar c,

(i) $\rho(A + cB) = \rho(A) + c\rho(B)$,

(ii) $\rho([A, B]) = \rho(A)\rho(B) - \rho(B)\rho(A)$.

If the map ρ is one-to-one, the representation is said to *faithful*.

Exercise 12.3 For a Lie algebra $\mathfrak{g}$, show that the map $A \to \mathrm{ad}\, A$ is a representation of the Lie algebra $\mathfrak{g}$, with $\mathfrak{g}$ itself the vector space of the representation. This is called the *adjoint representation*.

Example 12.4 (Vector field representations of Lie algebras) The Jacobi–Lie bracket of the vector fields ξ and η in Theorem 9.2 may be represented in coordinate charts as

$$\eta = \frac{dx}{ds}\Big|_{s=0} = v(x), \quad \text{and} \quad \xi = \frac{dx}{dt}\Big|_{t=0} = u(x).$$

The Jacobi–Lie bracket of these two vector fields yields a third vector field,

$$\begin{aligned}\xi\eta - \eta\xi &= \frac{d\eta}{dt}\Big|_{t=0} - \frac{d\xi}{ds}\Big|_{s=0} \\ &= \frac{dv}{dx}\frac{dx}{dt}\Big|_{t=0} - \frac{du}{dx}\frac{dx}{ds}\Big|_{s=0} = \frac{dv}{dx}\cdot u - \frac{du}{dx}\cdot v = u\cdot\nabla v - v\cdot\nabla u.\end{aligned}$$

Thus, the Jacobi–Lie bracket of vector fields at the tangent space of the identity T_eG is closed and may be represented in coordinate charts by the Lie bracket (commutator of vector fields)

$$[\xi, \eta] := \xi\eta - \eta\xi = u\cdot\nabla v - v\cdot\nabla u =: [u, v].$$

This example also proves the following

Proposition 12.5 *Let $\mathfrak{X}(\mathbb{R}^n)$ be the set of vector fields defined on $\mathbb{R}^n$. A Lie algebra $\mathfrak{g}$ may be represented on coordinate charts by vector fields $X_\xi = X^i_\xi \frac{\partial}{\partial x^i} \in \mathfrak{X}(\mathbb{R}^n)$ for each element $\xi \in \mathfrak{g}$. This vector field representation satisfies*

$$X_{[\xi,\eta]} = [X_\xi, X_\eta]$$

where $[\xi, \eta] \in \mathfrak{g}$ is the Lie algebra product and $[X_\xi, X_\eta]$ is the vector field commutator.

13 Lagrangian and Hamiltonian formulations

13.1 Newton's equations for particle motion in Euclidean space

Newton's equations

$$m_i \ddot{\mathbf{q}}_i = \mathbf{F}_i, \quad i = 1, \dots, N, \quad \text{(no sum on } i\text{)} \tag{7}$$

describe the *accelerations* $\ddot{\mathbf{q}}_i$ of N particles with

$$\begin{aligned} \text{Masses} \quad & m_i, \qquad i = 1, \dots, N, \\ \text{Euclidean positions} \quad & \mathbf{q} := (\mathbf{q}_1, \dots, \mathbf{q}_N) \in \mathbb{R}^{3N}, \end{aligned}$$

in response to *prescribed forces*,

$$\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_N),$$

acting on these particles. Suppose the forces arise from a *potential*. That is, let

$$\mathbf{F}_i(\mathbf{q}) = -\frac{\partial V(\{\mathbf{q}\})}{\partial \mathbf{q}_i}, \qquad V \colon \mathbb{R}^{3N} \to \mathbb{R}, \tag{8}$$

where $\partial V / \partial \mathbf{q}_i$ denotes the gradient of the potential with respect to the variable $\mathbf{q}_i$. Then Newton's equations (7) become

$$m_i \ddot{\mathbf{q}}_i = -\frac{\partial V}{\partial \mathbf{q}_i}, \qquad i = 1, \dots, N. \tag{9}$$

Remark 13.1 Newton (1620) introduced the gravitational potential for celestial mechanics, now called the *Newtonian potential*,

$$V(\{\mathbf{q}\}) = \sum_{i,j=1}^{N} \frac{-G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}. \tag{10}$$

13.2 Equivalence Theorem

Theorem 13.2 (Lagrangian and Hamiltonian formulations) *Newton's equations in potential form,*

$$m_i \ddot{\mathbf{q}}_i = -\frac{\partial V}{\partial \mathbf{q}_i}, \qquad i = 1, \dots, N, \tag{11}$$

for particle motion in Euclidean space $\mathbb{R}^{3N}$ are equivalent *to the following four statements:*

(i) *The Euler–Lagrange equations*

(12) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i}\right) - \frac{\partial L}{\partial \mathbf{q}_i} = 0, \qquad i = 1, \dots, N,$$

hold for the Lagrangian $L\colon \mathbb{R}^{6N} = \{(\mathbf{q}, \dot{\mathbf{q}}) \mid \mathbf{q}, \dot{\mathbf{q}} \in \mathbb{R}^{3N}\} \to \mathbb{R}$, *defined by*

(13) $$L(\mathbf{q}, \dot{\mathbf{q}}) := \sum_{i=1}^{N} \frac{m_i}{2} \|\dot{\mathbf{q}}_i\|^2 - V(\mathbf{q}),$$

with $\|\dot{\mathbf{q}}_i\|^2 = \dot{\mathbf{q}}_i \cdot \dot{\mathbf{q}}_i = \dot{q}_i^j \dot{q}_i^k \delta_{jk}$ *(no sum on* i *).*

(ii) Hamilton's principle of stationary action, $\delta\mathcal{S} = 0$, *holds for the* action functional *(dropping* i *s)*

(14) $$\mathcal{S}[\mathbf{q}(\cdot)] := \int_a^b L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt.$$

(iii) Hamilton's equations of motion,

(15) $$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}},$$

hold for the Hamiltonian *resulting from the* Legendre transform,

(16) $$H(\mathbf{q}, \mathbf{p}) := \mathbf{p} \cdot \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}) - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})),$$

where $\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})$ *solves for* $\dot{\mathbf{q}}$ *from the definition* $\mathbf{p} := \partial L(\mathbf{q}, \dot{\mathbf{q}})/\partial \dot{\mathbf{q}}$.
In the case of Newton's equations in potential form (11), the Lagrangian in equation (13) yields $\mathbf{p}_i = m_i \dot{\mathbf{q}}_i$ *and the resulting Hamiltonian is (restoring* i *s)*

$$H = \underbrace{\sum_{i=1}^{N} \frac{1}{2m_i} \|\mathbf{p}_i\|^2}_{\text{Kinetic energy}} + \underbrace{V(\mathbf{q})}_{\text{Potential}}$$

(iv) *Hamilton's equations in their Poisson bracket formulation,*

(17) $$\dot{F} = \{F, H\} \quad \text{for all } F \in \mathcal{F}(P),$$

hold with Poisson bracket *defined by*

(18) $$\{F, G\} := \sum_{i=1}^{N} \left(\frac{\partial F}{\partial \mathbf{q}_i} \cdot \frac{\partial G}{\partial \mathbf{p}_i} - \frac{\partial F}{\partial \mathbf{p}_i} \cdot \frac{\partial G}{\partial \mathbf{q}_i} \right) \quad \text{for all} \quad F, G \in \mathcal{F}(P).$$

We will prove this theorem by proving a chain of linked equivalence relations: (11) $\Leftrightarrow$ (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) as propositions. (The symbol $\Leftrightarrow$ means "equivalent to".)

Step I: Proof that Newton's equations (11) $\Leftrightarrow$ (i) Check by direct verification. $\square$

Step II: Proof that (i) $\Leftrightarrow$ (ii) The *Euler–Lagrange equations (12) are equivalent to Hamilton's principle of stationary action.*

To simplify notation, we momentarily suppress the particle index i.

We need to prove the solutions of (12) are critical points $\delta\mathcal{S} = 0$ of the *action functional*

$$\mathcal{S}[\mathbf{q}(\cdot)] := \int_a^b L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt, \tag{19}$$

(where $\dot{\mathbf{q}} = d\mathbf{q}(t)/dt$) with respect to variations on $C^\infty([a,b], \mathbb{R}^{3N})$, the space of smooth trajectories $\mathbf{q}\colon [a,b] \to \mathbb{R}^{3N}$ with fixed endpoints $\mathbf{q}_a$, $\mathbf{q}_b$.

In $C^\infty([a,b], \mathbb{R}^{3N})$ consider a *deformation* $\mathbf{q}(t,s)$, $s \in (-\epsilon, \epsilon)$, $\epsilon > 0$, with fixed endpoints $\mathbf{q}_a$, $\mathbf{q}_b$, of a curve $\mathbf{q}_0(t)$, that is, $\mathbf{q}(t,0) = \mathbf{q}_0(t)$ for all $t \in [a,b]$ and $\mathbf{q}(a,s) = \mathbf{q}_0(a) = \mathbf{q}_a$, $\mathbf{q}(b,s) = \mathbf{q}_0(b) = \mathbf{q}_b$ for all $s \in (-\epsilon, \epsilon)$.

Define a *variation* of the curve $\mathbf{q}_0(\cdot)$ in $C^\infty([a,b], \mathbb{R}^{3N})$ by

$$\delta\mathbf{q}(\cdot) := \frac{d}{ds}\bigg|_{s=0} \mathbf{q}(\cdot, s) \in T_{\mathbf{q}_0(\cdot)} C^\infty([a,b], \mathbb{R}^{3N}),$$

and define the *first variation* of $\mathcal{S}$ at $\mathbf{q}_0(t)$ to be the derivative

$$\delta\mathcal{S} := \mathbf{D}\mathcal{S}[\mathbf{q}_0(\cdot)](\delta\mathbf{q}(\cdot)) := \frac{d}{ds}\bigg|_{s=0} \mathcal{S}[\mathbf{q}(\cdot, s)]. \tag{20}$$

Note that $\delta\mathbf{q}(a) = \delta\mathbf{q}(b) = \mathbf{0}$. With these notations, *Hamilton's principle of stationary action* states that the curve $\mathbf{q}_0(t)$ satisfies the Euler–Lagrange equations (12) if and only if $\mathbf{q}_0(\cdot)$ is a critical point of the action functional, that is, $\mathbf{D}\mathcal{S}[\mathbf{q}_0(\cdot)] = 0$. Indeed, using the equality of mixed partials, integrating by parts, and taking into account that $\delta\mathbf{q}(a) = \delta\mathbf{q}(b) = 0$, one finds

$$\begin{aligned}
\delta\mathcal{S} := \mathbf{D}\mathcal{S}[\mathbf{q}_0(\cdot)](\delta\mathbf{q}(\cdot)) &= \frac{d}{ds}\bigg|_{s=0} \mathcal{S}[\mathbf{q}(\cdot, s)] = \frac{d}{ds}\bigg|_{s=0} \int_a^b L(\mathbf{q}(t,s), \dot{\mathbf{q}}(t,s))\, dt \\
&= \sum_{i=1}^N \int_a^b \left[\frac{\partial L}{\partial \mathbf{q}_i} \cdot \delta\mathbf{q}_i(t,s) + \frac{\partial L}{\partial \dot{\mathbf{q}}_i} \cdot \delta\dot{\mathbf{q}}_i \right] dt \\
&= -\sum_{i=1}^N \int_a^b \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i} \right) - \frac{\partial L}{\partial \mathbf{q}_i} \right] \cdot \delta\mathbf{q}_i \; dt = 0
\end{aligned}$$

for all smooth $\delta\mathbf{q}_i(t)$ satisfying $\delta\mathbf{q}_i(a) = \delta\mathbf{q}_i(b) = 0$. This proves the equivalence of (i) and (ii), upon restoring particle index i in the last two lines. $\square$

Definition 13.3 The conjugate momenta for the Lagrangian in (13) are defined as

$$\mathbf{p}_i := \frac{\partial L}{\partial \dot{\mathbf{q}}_i} = m_i \dot{\mathbf{q}}_i \in \mathbb{R}^3, \quad i = 1, \dots, N, \quad \text{(no sum on } i\text{)} \tag{21}$$

Definition 13.4 The *Hamiltonian* is defined via the change of variables $(\mathbf{q}, \dot{\mathbf{q}}) \mapsto (\mathbf{q}, \mathbf{p})$, called the *Legendre transform*,

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &:= \mathbf{p} \cdot \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}) - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})) \\ &= \sum_{i=1}^{N} \frac{m_i}{2} \|\dot{\mathbf{q}}_i\|^2 + V(\mathbf{q}) \\ &= \underbrace{\sum_{i=1}^{N} \frac{1}{2m_i} \|\mathbf{p}_i\|^2}_{\text{Kinetic energy}} + \underbrace{V(\mathbf{q})}_{\text{Potential}} \end{aligned} \tag{22}$$

Remark 13.5 The value of the Hamiltonian coincides with the total energy of the system. This value will be shown momentarily to remain constant under the evolution of Euler–Lagrange equations (12).

Remark 13.6 The Hamiltonian H may be obtained from the Legendre transformation as a function of the variables $(\mathbf{q}, \mathbf{p})$, provided one may solve for $\dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})$, which requires the Lagrangian to be *regular*, for example,

$$\det \frac{\partial^2 L}{\partial \dot{\mathbf{q}}_i \partial \dot{\mathbf{q}}_i} \neq 0 \quad \text{(no sum on } i\text{)}.$$

Step III: Proof that (ii) $\Leftrightarrow$ (iii) (Hamilton's principle of stationary action is equivalent to Hamilton's canonical equations.) Lagrangian (13) is regular and the derivatives of the Hamiltonian may be shown to satisfy,

$$\frac{\partial H}{\partial \mathbf{p}_i} = \frac{1}{m_i} \mathbf{p}_i = \dot{\mathbf{q}}_i = \frac{d\mathbf{q}_i}{dt} \quad \text{and} \quad \frac{\partial H}{\partial \mathbf{q}_i} = \frac{\partial V}{\partial \mathbf{q}_i} = -\frac{\partial L}{\partial \mathbf{q}_i}.$$

Consequently, the Euler–Lagrange equations (12) imply

$$\dot{\mathbf{p}}_i = \frac{d\mathbf{p}_i}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i}\right) = \frac{\partial L}{\partial \mathbf{q}_i} = -\frac{\partial H}{\partial \mathbf{q}_i}.$$

These calculations show that the Euler–Lagrange equations (12) are equivalent to Hamilton's canonical equations:

$$\dot{\mathbf{q}}_i = \frac{\partial H}{\partial \mathbf{p}_i}, \qquad \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{q}_i}, \tag{23}$$

where $\partial H/\partial \mathbf{q}_i, \partial H/\partial \mathbf{p}_i \in \mathbb{R}^3$ are the gradients of H with respect to $\mathbf{q}_i, \mathbf{p}_i \in \mathbb{R}^3$, respectively. This proves the equivalence of (ii) and (iii). □

Remark 13.7 The Euler–Lagrange equations are second order and they determine *curves in configuration space* $\mathbf{q}_i \in C^\infty([a,b], \mathbb{R}^{3N})$. In contrast, Hamilton's equations are first order and they determine *curves in phase space* $(\mathbf{q}_i, \mathbf{p}_i) \in C^\infty([a,b], \mathbb{R}^{6N})$, a space whose dimension is twice the dimension of the configuration space.

Step IV: Proof that (iii) ⇔ (iv) (Hamilton's canonical equations may be written using a Poisson bracket.)

By the chain rule and (23) any $F \in \mathcal{F}(P)$ satisfies

$$\begin{aligned}\frac{dF}{dt} &= \sum_{i=1}^{N}\left(\frac{\partial F}{\partial \mathbf{q}_i}\cdot\dot{\mathbf{q}}_i + \frac{\partial F}{\partial \mathbf{p}_i}\cdot\dot{\mathbf{p}}_i\right)\\ &= \sum_{i=1}^{N}\left(\frac{\partial F}{\partial \mathbf{q}_i}\cdot\frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial F}{\partial \mathbf{p}_i}\cdot\frac{\partial H}{\partial \mathbf{q}_i}\right) = \{F, H\}.\end{aligned}$$

This finishes the proof of the theorem, by proving the equivalence of (iii) and (iv). □

Remark 13.8 (Energy conservation) Since the Poisson bracket is skew symmetric, $\{H, F\} = -\{F, H\}$, one finds that $\dot{H} = \{H, H\} = 0$. Consequently, the value of the Hamiltonian is preserved by the evolution. Thus, the Hamiltonian is said to be a *constant of the motion*.

Exercise 13.9 Show that the Poisson bracket is bilinear, skew symmetric, satisfies the Jacobi identity and acts as derivation on products of functions in phase space.

Exercise 13.10 Given two constants of motion, what does the Jacobi identity imply about additional constants of motion?

Exercise 13.11 Compute the Poisson brackets among

$$J_i = \epsilon_{ijk} p_j q_k$$

in Euclidean space. What Lie algebra do these Poisson brackets recall to you?

Exercise 13.12 Verify that Hamilton's equations determined by the function

$$\langle J(z), \xi\rangle = \xi\cdot(\mathbf{q}\times\mathbf{p})$$

give infinitesimal rotations about the ξ–axis.

14 Hamilton's principle on manifolds

Theorem 14.1 (Hamilton's Principle of Stationary Action) *Let the smooth function $L\colon TQ \to \mathbb{R}$ be a Lagrangian on TQ. A C^2 curve $c\colon [a,b] \to Q$ joining $q_a = c(a)$ to $q_b = c(b)$ satisfies the Euler–Lagrange equations if and only if*

$$\delta \int_a^b L(c(t), \dot{c}(t))dt = 0.$$

Proof The meaning of the variational derivative in the statement is the following. Consider a family of C^2 curves $c(t,s)$ for $|s| < \varepsilon$ satisfying $c_0(t) = c(t)$, $c(a,s) = q_a$, and $c(b,s) = q_b$ for all $s \in (-\varepsilon, \varepsilon)$. Then

$$\delta \int_a^b L(c(t), \dot{c}(t))dt := \frac{d}{ds}\bigg|_{s=0} \int_a^b L(c(t,s), \dot{c}(t,s))dt.$$

Differentiating under the integral sign, working in local coordinates (covering the curve $c(t)$ by a finite number of coordinate charts), integrating by parts, denoting

$$v(t) := \frac{d}{ds}\bigg|_{s=0} c(t,s),$$

and taking into account that $v(a) = v(b) = 0$, yields

$$\int_a^b \left(\frac{\partial L}{\partial q^i} v^i + \frac{\partial L}{\partial \dot{q}^i} \dot{v}^i \right) dt = \int_a^b \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \right) v^i dt.$$

This vanishes for any C^1 function $v(t)$ if and only if the Euler–Lagrange equations hold. □

Remark 14.2 The integral appearing in this theorem

$$\mathcal{S}(c(\cdot)) := \int_a^b L(c(t), \dot{c}(t))dt$$

is called the *action integral*. It is defined on C^2 curves $c\colon [a,b] \to Q$ with fixed endpoints, $c(a) = q_a$ and $c(b) = q_b$.

Remark 14.3 (Variational derivatives of functionals vs Lie derivatives of functions) The variational derivative of a functional $S[u]$ is defined as the linearization

$$\lim_{\epsilon \to 0} \frac{S[u + \epsilon v] - S[u]}{\epsilon} = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} S[u + \epsilon v] = \left\langle \frac{\delta S}{\delta v}, v \right\rangle.$$

Compare this to the expression for the Lie derivative of a function. If f is a real valued function on a manifold M and X is a vector field on M, the Lie derivative of f along X is defined as the directional derivative

$$\mathcal{L}_X f = X(f) := \mathbf{d} f \cdot X.$$

If M is finite-dimensional, this is

$$\mathcal{L}_X f = X[f] := \mathbf{d} f \cdot X = \frac{\partial f}{\partial x^i} X^i = \lim_{\epsilon \to 0} \frac{f(x + \epsilon X) - f(x)}{\epsilon}.$$

The similarity is suggestive: Namely, the Lie derivative of a function and the variational derivative of a functional are both defined as linearizations of smooth maps in certain directions.

The next theorem emphasizes the role of Lagrangian one-forms and two-forms in the variational principle. The following is a direct corollary of the previous theorem.

Theorem 14.4 *Given a C^k Lagrangian $L\colon TQ \to \mathbb{R}$ for $k \geq 2$, there exists a unique C^{k-2} map $\mathcal{E}L(L)\colon \ddot{Q} \to T^*Q$, where*

$$\ddot{Q} := \left\{ \frac{d^2 q}{dt^2}\bigg|_{t=0} \in T(TQ) \;\bigg|\; q(t) \text{ is a } C^2 \text{ curve in } Q \right\}$$

is a submanifold of $T(TQ)$, and a unique C^{k-1} one-form $\Theta_L \in \Lambda^1(TQ)$, such that for all C^2 variations $q(t,s)$ (defined on a fixed t–interval) of $q(t,0) = q_0(t) := q(t)$, we have

$$\begin{aligned}(24)\quad \delta\mathcal{S} := \frac{d}{ds}\bigg|_{s=0} \mathcal{S}[c(\cdot,s)] &= \mathbf{D}\mathcal{S}[q(\cdot)] \cdot \delta q(\cdot) \\ &= \int_a^b \mathcal{E}L(L)\,(q,\dot{q},\ddot{q}) \cdot \delta q \, dt + \underbrace{\Theta_L\big(q,\dot{q}\big) \cdot \delta q\Big|_a^b}_{\text{cf. Noether's Theorem}}\end{aligned}$$

where

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

15 Summary handout for differential forms

Vector fields and 1–forms

Let M be a manifold. In what follows, all maps may be assumed to be C^∞, although that's not necessary.

A *vector field* on M is a map $X\colon M \to TM$ such that $X(x) \in T_xM$ for every $x \in M$. The set of all smooth vector fields on M is written $\mathfrak{X}(M)$. ("Smooth" means differentiable or C^r for some $r \leq \infty$, depending on context.)

A *(differential)* 1*–form* on M is a map $\theta\colon M \to T^*M$ such that $\theta(x) \in T_x^*M$ for every $x \in M$.

More generally, if $\pi\colon E \to M$ is a bundle, then a *section* of the bundle is a map $\varphi\colon M \to E$ such that $\pi \circ \varphi(x) = x$ for all $x \in M$. So a vector field is a section of the tangent bundle, while a 1–form is section of the cotangent bundle.

Vector fields can added and also multiplied by scalar functions $k\colon M \to \mathbb{R}$, as follows: $(X_1 + X_2)(x) = X_1(x) + X_2(x)$, $(kX)(x) = k(x)X(x)$.

Differential forms can added and also multiplied by scalar functions $k\colon M \to \mathbb{R}$, as follows: $(\alpha + \beta)(x) = \alpha(x) + \beta(x)$, $(k\theta)(x) = k(x)\theta(x)$.

We have already defined the push-forward and pull-back of a vector field. The *pull-back* of a 1–form θ on N by a map $\varphi\colon M \to N$ is the 1–form $\varphi^*\theta$ on M defined by

$$\left(\varphi^*\theta\right)(x) \cdot v = \theta\left(\varphi(x)\right) \cdot T\varphi(v)$$

The *push-forward* of a 1–form α on M by a diffeomorphism $\psi\colon M \to N$ is the pull-back of α by ψ^{-1}.

A vector field can be *contracted* with a differential form, using the pairing between tangent and cotangent vectors: $(X \lrcorner\, \theta)(x) = \theta(x) \cdot X(x)$. Note that $X \lrcorner\, \theta$ is a map from M to $\mathbb{R}$. Many books write $i_X\theta$ in place of $X \lrcorner\, \theta$, and the contraction operation is often called *interior product*.

The *differential* of $f\colon M \to \mathbb{R}$ is a 1–form df on M defined by

$$df(x) \cdot v = \frac{d}{dt} f(c(t)\Big|_{t=0}$$

for any $x \in M$, any $v \in T_xM$ and any path $c(t)$ in M such that $c(0) = 0$ and $c'(0) = v$. The left hand side, $df(x) \cdot v$, means the pairing between cotangent and tangent vectors, which could also be written $df(x)(v)$ or $\langle df(x), v\rangle$.

Note:

$$X \lrcorner\, df = \pounds_X f = X[f]$$

Remark 15.1 df is very similar to Tf, but Tf is defined for all differentiable $f\colon M \to N$, whereas df is only defined when $N = \mathbb{R}$ (in this course, anyway). In this case, Tf is a map from TM to $T\mathbb{R}$, and $Tf(v) = df(x) \cdot v \in T_{f(x)}\mathbb{R}$ for every $v \in T_xM$ (we have identified $T_f(x)\mathbb{R}$ with $\mathbb{R}$.)

In coordinates...

Let M be n–dimensional, and let $x^1, \dots, x^n$ be differentiable local coordinates for M. This means that there's an open subset U of M and an open subset V of $\mathbb{R}^n$ such that the map $\varphi\colon U \to V$ defined by $\varphi(x) = \big(x^1(x), \dots, x^n(x)\big)$ is a diffeomorphism. In particular, each x^i is a map from M to $\mathbb{R}$, so the differential dx^i is defined. There is also a vector field $\frac{\partial}{\partial x^i}$ for every i, which is defined by $\frac{\partial}{\partial x^i}(x) = \frac{d}{dt}\varphi^{-1}\big(\varphi(x) + t\mathbf{e}_i\big)\big|_{t=0}$, where $\mathbf{e}_i$ is the i^{th} standard basis vector.

Exercise 15.2 Verify that

$$\frac{\partial}{\partial x^i} \lrcorner\, dx^j \equiv \delta^i_j$$

(where $\equiv$ means the left hand side is a constant function with value δ^i_j).

Remark 15.3 Of course, given a coordinate system $\varphi = \big(x^1, \dots, x^n\big)$, it is usual to write $x = \big(x^1, \dots, x^n\big)$, which means x is identified with $\big(x^1(x), \dots, x^n(x)\big) = \varphi(x)$.

For every $x \in M$, the vectors $\frac{\partial}{\partial x^i}(x)$ form a basis for $T_x M$, so every $v \in T^x M$ can be uniquely expressed as $v = v^i \frac{\partial}{\partial x^i}(x)$. This expression defines the *tangent-lifted coordinates* $x^1, \dots, x^n, v^1, \dots v^n$ on TM (they are local coordinates, defined on $TU \subset TM$).

For every $x \in M$, the covectors $dx^i(x)$ form a basis for $T^*_x M$, so every $\alpha \in T^x M$ can be uniquely expressed as $\alpha = \alpha_i dx^i(x)$. This expression defines the *cotangent-lifted coordinates* $x^1, \dots, x^n, \alpha_1, \dots, \alpha_n$ on T^*M (they are local coordinates, defined on $T^*U \subset T^*M$).

Note that the basis $\big(\frac{\partial}{\partial x^i}\big)$ is dual to the basis $\big(dx^1, \dots, dx^n\big)$, by the previous exercise. It follows that

$$\big(\alpha_i dx^i\big) \cdot \left(v^i \frac{\partial}{\partial x^i}\right) = \alpha_i v^i$$

(we have used the summation convention).

In mechanics, the configuration space is often called Q, and the lifted coordinates are written: $q^1, \dots, q^n, \dot q^1, \dots, \dot q^n$ (on TQ) and $q^1, \dots, q^n, p_1, \dots, p_n$ (on T^*Q).

Why the distinction between subscripts and superscripts? This is to keep track of how quantities vary if coordinates are changed (see next exercise). One benefit is that using the summation convention gives coordinate-independent answers.

Exercise 15.4 Consider two sets of local coordinates q^i and s^i on Q, related by $\left(s^1,\ldots,s^n\right)=\psi\left(q^1,\ldots,q^n\right)$. Verify that the corresponding tangent lifted coordinates $\dot{q}^i$ and $\dot{s}^i$ are related by

$$\dot{s}^i=\frac{\partial\psi^i}{\partial q^j}\dot{q}^j.$$

Note that the last equation can be written as $\dot{\mathbf{s}}=D\psi(q)\dot{\mathbf{q}}$, where $\dot{\mathbf{s}}$ is the column vector $(\dot{s}^1,\ldots\dot{s}^n)$, and similarly for $\dot{\mathbf{q}}$.

Do the corresponding calculation on the cotangent bundle side. See Definition 16.11.

The next level: TTQ, T^*T^*Q, et cetera

Since TQ is a manifold, we can consider vector fields on it, which are sections of $T(TQ)$. In coordinates, every vector field on TTQ has the form $X=a^i\frac{\partial}{\partial q^i}+b^i\frac{\partial}{\partial\dot{q}^i}$, where the a^i and b^i are functions of q and $\dot{q}$. Note that the same symbol q^i has two interpretations: as a coordinate on TQ and as a coordinate on Q, so $\frac{\partial}{\partial q^i}$ can mean a vector field TQ (as above) or on Q.

The tangent lift of the bundle projection $\tau\colon TQ\to Q$ is a map $T\tau\colon TTQ\to TQ$. If X is written in coordinates as above, then $T\tau\circ X=a^i\frac{\partial}{\partial q^i}$. A vector field X on TTQ is *second order* if $T\tau\circ X(v)=v$; in coordinates, $a^i=\dot{q}^i$. The name comes from the process of reducing of second order equations to first order ones by introducing new variables $\dot{q}^i=\frac{dq^i}{dt}$.

One may also consider T^*TQ, TT^*Q and T^*T^*Q. However, the subscript/superscript distinction is problematic here.

1**–forms** The 1–forms on T^*Q are sections of T^*T^*Q. Given cotangent-lifted local coordinates

$$\left(q^1,\ldots,q^n,p_1,\ldots,p_n\right)$$

on T^*Q, the general 1–form on T^*Q has the form $a_idq^i+b_idp_i$, where a_i and b_i are functions of (q,p). The *canonical* 1*–form* on T^*Q is

$$\theta=p_idq^i,$$

also written in the short form $p\,dq$. Pairing $\theta(q,p)$ with an arbitrary tangent vector $v=a^i\frac{\partial}{\partial q^i}+b^i\frac{\partial}{\partial p^i}\in T_{(q,p)}T^*Q$ gives

$$\langle\theta(q,p),v\rangle=\left\langle p_idq^i,a^i\frac{\partial}{\partial q^i}+b^i\frac{\partial}{\partial p^i}\right\rangle=p_ia^i=\left\langle p_idq^i,a^i\frac{\partial}{\partial q^i}\right\rangle=\langle p,T\tau^*(v)\rangle,$$

where $\tau^*\colon T^*Q \to Q$ is projection. In the last line we have interpreted q^i as a coordinate on Q, which implies that $p_i dq^i = p$, by definition of the coordinates p_i. Note that the last line is coordinate-free.

2**–forms** Recall that a 1–form on M, evaluated at a point $x \in M$, is a linear map from $T_x M$ to $\mathbb{R}$.

A 2–form on M, evaluated at a point $x \in M$, is a skew-symmetric bilinear form on $T_x M$; and the bilinear form has to vary smoothly as x changes. (Confusingly, bilinear forms can be skew-symmetric, symmetric or neither; *differential* forms are assumed to be skew-symmetric.)

The *pull-back* of a 2–form ω on N by a map $\varphi\colon M \to N$ is the 2–form $\varphi^*\omega$ on M defined by

$$\left(\varphi^*\omega\right)(x)(v,w) = \theta\left(\varphi(x)\right)\left(T\varphi(v), T\varphi(w)\right)$$

The *push-forward* of a 2–form ω on M by a diffeomorphism $\psi\colon M \to N$ is the pull-back of ω by ψ^{-1}.

A vector field X can be *contracted* with a 2–form ω to get a 1–form $X \rfloor \omega$ defined by

$$(X \rfloor \omega)(x)(v) = \omega(x)\left(X(x), v\right)$$

for any $v \in T_x M$. A shorthand for this is $(X \rfloor \omega)(v) = \omega(X, v)$, or just $X \rfloor \omega = \omega(X, \cdot)$.

The *tensor product* of two 1–forms α and β is the 2–form $\alpha \otimes \beta$ defined by

$$(\alpha \otimes \beta)(v,w) = \alpha(v)\beta(w)$$

for all $v, w \in T_x^* M$.

The *wedge product* of two 1–forms α and β is the skew-symmetric 2–form $\alpha \wedge \beta$ defined by

$$(\alpha \wedge \beta)(v,w) = \alpha(v)\beta(w) - \alpha(w)\beta(v)\,.$$

Exterior derivative The differential df of a real-valued function is also called the exterior derivative of f. In this context, real-valued functions can be called 0–forms. The exterior derivative is a linear operation from 0–forms to 1–forms that satisfies the Leibniz identity, a.k.a. the product rule,

$$d(fg) = f\ dg + g\ df$$

The exterior derivative of a 1–form is an alternating 2–form, defined as follows:

$$d\left(a_i dx^i\right) = \frac{\partial a_i}{\partial x^j} dx^j \wedge dx^i.$$

Exterior derivative is a linear operation from 1–forms to 2 forms. The following identity is easily checked:

$$d(df) = 0$$

for all scalar functions f.

***n*–forms** See Marsden and Ratiu [44], Lee [38], or Abraham and Marsden [1]. Unless otherwise specified, n–forms are assumed to be alternating. Wedge products and contractions generalise.

It is a fact that all n–forms are linear combinations of wedge products of 1–forms. Thus we can define exterior derivative recursively by the properties

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta,$$

for all k–forms α and all forms β, and

$$d \circ d = 0$$

In local coordinates, if $\alpha = \alpha_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ (sum over all $i_1 < \cdots < i_k$), then

$$d\alpha = \frac{\partial \alpha_{i_1 \cdots i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i^k}$$

The *Lie derivative* of an n–form θ in the direction of the vector field X is defined as

$$\pounds_X \theta = \frac{d}{dt} \varphi_t^* \theta \bigg|_{t=0},$$

where φ is the flow of X.

Pull-back commutes with the operations d, $\lrcorner$, $\wedge$ and Lie derivative.

Cartan's magic formula

$$\pounds_X \alpha = d\,(X \lrcorner \alpha) + X \lrcorner d\alpha$$

This looks even more magic when written using the notation $i_X \alpha = X \lrcorner \alpha$:

$$\pounds_X = d i_X + i_X d$$

An n–form α is *closed* if $d\alpha = 0$, and *exact* if $\alpha = d\beta$ for some β. All exact forms are closed (since $d \circ d = 0$), but the converse is false. It is true that all closed forms are *locally* exact; this is the *Poincaré Lemma*.

Remark 15.5 For a survey of the basic definitions, properties, and operations on differential forms, as well as useful of tables of relations between differential calculus and vector calculus, see, for example, Bloch [5, Chapter 2].

16 Euler–Lagrange equations of manifolds

In Theorem 14.4,

$$\delta\mathcal{S} := \frac{d}{ds}\bigg|_{s=0} \mathcal{S}[c(\cdot,s)] = \mathbf{D}\mathcal{S}[q(\cdot)]\cdot\delta q(\cdot) \tag{25}$$
$$= \int_a^b \mathcal{E}L(L)\big(q,\dot q,\ddot q\big)\cdot\delta q\,dt + \underbrace{\Theta_L\,(q,\dot q)\cdot\delta q\Big|_a^b}_{\text{cf. Noether's Theorem}}$$

where

$$\delta q(t) = \frac{d}{ds}\bigg|_{s=0} q(t,s),$$

the map $\mathcal{EL}\colon \ddot{Q}\to T^*Q$ is called the *Euler–Lagrange operator* and its expression in local coordinates is

$$\mathcal{EL}(q,\dot q,\ddot q)_i = \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot q^i}.$$

One understands that the formal time derivative is taken in the second summand and everything is expressed as a function of $(q,\dot q,\ddot q)$.

Theorem 16.1 (Noether (1918) Symmetries and Conservation Laws) *If the action variation in (25) vanishes $\delta\mathcal{S}=0$ because of a symmetry transformation which does* not *preserve the end points and the Euler–Lagrange equations hold, then the term marked* cf. Noether's Theorem *must also vanish. However, vanishing of this term now is interpreted as a constant of motion. Namely, the term,*

$$A(v,w) := \langle \mathbb{F}L(v), w\rangle, \quad \text{or, in coordinates} \quad A(q,\dot q,\delta q) = \frac{\partial L}{\partial \dot q^i}\,\delta q^i,$$

is constant for solutions of the Euler–Lagrange equations. This result first appeared in Noether [36]. In fact, the result in [36] is more general than this. In particular, in the PDE (Partial Differential Equation) setting one must also include the transformation of the volume element in the action principle. See, for example, Olver [53] for good discussions of the history, framework and applications of Noether's theorem.

Exercise 16.2 Show that conservation of energy results from Noether's Theorem if, in Hamilton's principle, the variations are chosen as

$$\delta q(t) = \frac{d}{ds}\bigg|_{s=0} q(t,s),$$

corresponding to symmetry of the Lagrangian under reparametrizations of time along the given curve $q(t)\to q(\tau(t,s))$.

The canonical Lagrangian one-form and two-form.The one-form Θ_L, whose existence and uniqueness is guaranteed by Theorem 14.4, appears as the boundary term of the derivative of the action integral, when the endpoints of the curves on the configuration manifold are free. In finite dimensions, its local expression is

$$\Theta_L(q,\dot q) := \frac{\partial L}{\partial \dot q^i}\mathbf{d}q^i \qquad \big(= p_i(q,\dot q)\,\mathbf{d}q^i\big).$$

The corresponding closed two-form $\Omega_L = \mathbf{d}\Theta_L$ obtained by taking its exterior derivative may be expressed as

$$\Omega_L := -\mathbf{d}\Theta_L = \frac{\partial^2 L}{\partial \dot q^i \partial q^j}\mathbf{d}q^i \wedge \mathbf{d}q^j + \frac{\partial^2 L}{\partial \dot q^i \partial \dot q^j}\mathbf{d}q^i \wedge \mathbf{d}\dot q^j \qquad \big(= \mathbf{d}p_i(q,\dot q) \wedge \mathbf{d}q^i\big).$$

These coefficients may be written as the $2n \times 2n$ skew-symmetric matrix

$$\Omega_L = \begin{pmatrix} \mathcal{A} & \frac{\partial^2 L}{\partial \dot q^i \partial \dot q^j} \\ -\frac{\partial^2 L}{\partial \dot q^i \partial \dot q^j} & 0 \end{pmatrix}, \tag{26}$$

where $\mathcal{A}$ is the skew-symmetric $n{\times}n$ matrix $\big(\frac{\partial^2 L}{\partial \dot q^i \partial q^j}\big) - \big(\frac{\partial^2 L}{\partial \dot q^i \partial q^j}\big)^T$.

Non-degeneracy of Ω_L is equivalent to the invertibility of the matrix $\big(\frac{\partial^2 L}{\partial \dot q^i \partial \dot q^j}\big)$.

Definition 16.3 The *Legendre transformation* $\mathbb{F}L\colon TQ \to T^*Q$ is the smooth map near the identity defined by

$$\langle \mathbb{F}L(v_q), w_q\rangle := \frac{d}{ds}\bigg|_{s=0} L(v_q + s w_q).$$

In the finite dimensional case, the local expression of $\mathbb{F}L$ is

$$\mathbb{F}L(q^i, \dot q^i) = \left(q^i, \frac{\partial L}{\partial \dot q^i}\right) = (q^i, p_i(q,\dot q)).$$

If the skew-symmetric matrix (26) is invertible, the Lagrangian L is said to be *regular*. In this case, by the implicit function theorem, $\mathbb{F}L$ is locally invertible. If $\mathbb{F}L$ is a diffeomorphism, L is called *hyperregular*.

Definition 16.4 Given a Lagrangian L, the *action* of L is the map $A\colon TQ \to \mathbb{R}$ given by

$$A(v) := \langle \mathbb{F}L(v), v\rangle, \quad \text{or, in coordinates} \quad A(q,\dot q) = \frac{\partial L}{\partial \dot q^i}\dot q^i, \tag{27}$$

and the *energy* of L is

$$E(v) := A(v) - L(v), \quad \text{or, in coordinates} \quad E(q,\dot q) = \frac{\partial L}{\partial \dot q^i}\dot q^i - L(q,\dot q). \tag{28}$$

16.1 Lagrangian vector fields and conservation laws

Definition 16.5 A vector field Z on TQ is called a *Lagrangian vector field* if

$$\Omega_L(v)(Z(v), w) = \langle \mathbf{d}E(v), w\rangle,$$

for all $v \in T_qQ$, $w \in T_v(TQ)$.

Proposition 16.6 *The* energy is conserved *along the flow of a Lagrangian vector field* Z.

Proof Let $v(t) \in TQ$ be an integral curve of Z. Skew-symmetry of Ω_L implies

$$\begin{aligned}\frac{d}{dt}E(v(t)) = \langle \mathbf{d}E(v(t)), \dot v(t)\rangle = \langle \mathbf{d}E(v(t)), Z(v(t))\rangle \\ = \Omega_L(v(t))\,(Z(v(t)), Z(v(t))) = 0.\end{aligned}$$

Thus, $E(v(t))$ is constant in t. □

16.2 Equivalence of dynamics for hyperregular Lagrangians and Hamiltonians

Recall that a Lagrangian L is said to be *hyperregular* if its Legendre transformation $\mathbb{F}L\colon TQ \to T^*Q$ is a diffeomorphism.

The equivalence between the Lagrangian and Hamiltonian formulations for hyperregular Lagrangians and Hamiltonians is summarized below, following Marsden and Ratiu [44].

(a) Let L be a hyperregular Lagrangian on TQ and $H = E \circ (\mathbb{F}L)^{-1}$, where E is the energy of L and $(\mathbb{F}L)^{-1}\colon T^*Q \to TQ$ is the inverse of the Legendre transformation. Then the Lagrangian vector field Z on TQ and the Hamiltonian vector field X_H on T^*Q are related by the identity

$$(\mathbb{F}L)^* X_H = Z.$$

Furthermore, if $c(t)$ is an integral curve of Z and $d(t)$ an integral curve of X_H with $\mathbb{F}L(c(0)) = d(0)$, then $\mathbb{F}L(c(t)) = d(t)$ and their integral curves coincide on the manifold Q. That is, $\tau_Q(c(t)) = \pi_Q(d(t)) = \gamma(t)$, where $\tau_Q\colon TQ \to Q$ and $\pi_Q\colon T^*Q \to Q$ are the canonical bundle projections.

In particular, the pull back of the inverse Legendre transformation $\mathbb{F}L^{-1}$ induces a one-form Θ and a closed two-form Ω on T^*Q by

$$\Theta = (\mathbb{F}L^{-1})^*\Theta_L, \qquad \Omega = -\mathbf{d}\Theta = (\mathbb{F}L^{-1})^*\Omega_L.$$

In coordinates, these are the canonical presymplectic and symplectic forms, respectively,

$$\Theta = p_i\,\mathbf{d}q^i, \qquad \Omega = -\mathbf{d}\Theta = \mathbf{d}p_i \wedge \mathbf{d}q^i.$$

(b) A Hamiltonian $H\colon T^*Q \to \mathbb{R}$ is said to be *hyperregular* if the smooth map $\mathbb{F}H\colon T^*Q \to TQ$, defined by

$$\langle \mathbb{F}H(\alpha_q), \beta_q \rangle := \frac{d}{ds}\bigg|_{s=0} H(\alpha_q + s\beta_q), \qquad \alpha_q, \beta_q \in T_q^*Q,$$

is a diffeomorphism. Define the *action* of H by $G := \langle \Theta, X_H \rangle$. If H is a hyperregular Hamiltonian then the energies of L and H and the actions of L and H are related by

$$E = H \circ (\mathbb{F}H)^{-1}, \qquad A = G \circ (\mathbb{F}H)^{-1}.$$

Also, the Lagrangian $L = A - E$ is hyperregular and $\mathbb{F}L = \mathbb{F}H^{-1}$.

(c) These constructions define a bijective correspondence between hyperregular Lagrangians and Hamiltonians.

Remark 16.7 For thorough discussions of many additional results arising from the Hamilton's principle for hyperregular Lagrangians see, for example, Marsden and Ratiu [44, Chapters 7 and 8].

Exercise 16.8 (Spherical pendulum) A particle rolling on the interior of a spherical surface under gravity is called a spherical pendulum. Write down the Lagrangian and the equations of motion for a spherical pendulum with S^2 as its configuration space. Show explicitly that the Lagrangian is hyperregular. Use the Legendre transformation to convert the equations to Hamiltonian form. Find the conservation law corresponding to angular momentum about the axis of gravity by "bare hands" methods.

Exercise 16.9 (Differentially rotating frames) The Lagrangian for a free particle of unit mass relative to a moving frame is obtained by setting

$$L(\dot{\mathbf{q}}, \mathbf{q}, t) = \tfrac{1}{2}\|\dot{\mathbf{q}}\|^2 + \dot{\mathbf{q}} \cdot \mathbf{R}(\mathbf{q}, t)$$

for a function $\mathbf{R}(\mathbf{q}, t)$ which prescribes the space and time dependence of the moving frame velocity. For example, a frame rotating with time-dependent frequency $\Omega(t)$

about the vertical axis $\hat{\mathbf{z}}$ is obtained by choosing $\mathbf{R}(\mathbf{q},t) = \mathbf{q} \times \Omega(t)\hat{\mathbf{z}}$. Calculate $\Theta_L(q,\dot{q})$, $\Omega_L(q,\dot{q})$, the Euler–Lagrange operator $\mathcal{E}L(L)(q,\dot{q},\ddot{q})$, the Hamiltonian and its corresponding canonical equations.

Exercise 16.10 Calculate the action and the energy for the Lagrangian in Exercise 16.9.

Definition 16.11 (Cotangent lift) Given two manifolds Q and S related by a diffeomorphism $f\colon Q \mapsto S$, the *cotangent lift* $T^*f\colon T^*S \mapsto T^*Q$ of f is defined by

$$\langle T^* f(\alpha), v\rangle = \langle \alpha, Tf(v)\rangle \tag{29}$$

where

$$\alpha \in T_s^* S, \quad v \in T_q Q, \quad \text{and} \quad s = f(q).$$

As explained by Marsden and Ratiu [44, Chapter 6], cotangent lifts preserve the *action* of the Lagrangian L, which we write as

$$\langle \mathbf{p}, \dot{\mathbf{q}}\rangle = \langle \alpha, \dot{\mathbf{s}}\rangle, \tag{30}$$

where $\mathbf{p} = T^* f(\alpha)$ is the cotangent lift of α under the diffeomorphism f and $\dot{\mathbf{s}} = Tf(\dot{\mathbf{q}})$ is the tangent lift of $\dot{\mathbf{q}}$ under the function f, which is written in Euclidean coordinate components as $q^i \to s^i = f^i(\mathbf{q})$. Preservation of the action in (30) yields the coordinate relations,

$$\text{(Tangent lift in coordinates)} \quad \dot{s}^j = \frac{\partial f^j}{\partial q^i}\dot{q}^i \qquad \Longrightarrow$$

$$p_i = \alpha_k \frac{\partial f^k}{\partial q^i} \quad \text{(Cotangent lift in coordinates)}$$

Thus, in coordinates, the cotangent lift is the inverse transpose of the tangent lift.

Remark 16.12 The cotangent lift of a function preserves the induced action one-form,

$$\langle \mathbf{p}, \mathbf{dq}\rangle = \langle \alpha, \mathbf{ds}\rangle,$$

so it is a source of (pre-)symplectic transformations.

16.3 The classic Euler–Lagrange example: geodesic flow

An important example of a Lagrangian vector field is the geodesic spray of a Riemannian metric. A *Riemannian manifold* is a smooth manifold Q endowed with a symmetric nondegenerate covariant tensor g, which is positive definite. Thus, on each tangent space $T_q Q$ there is a nondegenerate definite inner product defined by pairing with $g(q)$.

If (Q, g) is a Riemannian manifold, there is a natural Lagrangian on it given by the *kinetic energy* K of the metric g, namely,

$$K(v) := \tfrac{1}{2} g(q)(v_q, v_q),$$

for $q \in Q$ and $v_q \in T_q Q$. In finite dimensions, in a local chart,

$$K(q, \dot{q}) = \tfrac{1}{2} g_{ij}(q) \dot{q}^i \dot{q}^j.$$

The Legendre transformation is in this case $\mathbb{F}K(v_q) = g(q)(v_q, \cdot)$, for $v_q \in T_q Q$. In coordinates, this is

$$\mathbb{F}K(q, \dot{q}) = \left(q^i, \frac{\partial K}{\partial \dot{q}^i} \right) = (q^i, g_{ij}(q)\dot{q}^j) =: (q^i, p_i).$$

The Euler–Lagrange equations become the *geodesic equations* for the metric g, given (for finite dimensional Q in a local chart) by

$$\ddot{q}^i + \Gamma^i_{jk} \dot{q}^j \dot{q}^k = 0, \quad i = 1, \dots n,$$

where the three-index quantities

$$\Gamma^h_{jk} = \tfrac{1}{2} g^{hl} \left(\frac{\partial g_{jl}}{\partial q^k} + \frac{\partial g_{kl}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l} \right), \quad \text{with} \quad g_{ih} g^{hl} = \delta^l_i,$$

are the *Christoffel symbols* of the Levi-Civita connection on (Q, g).

Exercise 16.13 Explicitly compute the geodesic equation as an Euler–Lagrange equation for the kinetic energy Lagrangian $K(q, \dot{q}) = \frac{1}{2} g_{ij}(q) \dot{q}^i \dot{q}^j$.

Exercise 16.14 For the kinetic energy Lagrangian $K(q, \dot{q}) = \frac{1}{2} g_{ij}(q) \dot{q}^i \dot{q}^j$ with $i, j = 1, 2, \dots, N$:

- Compute the momentum p_i canonical to q^i for geodesic motion.
- Perform the Legendre transformation to obtain the Hamiltonian for geodesic motion.
- Write out the geodesic equations in terms of q^i and its canonical momentum p_i.
- Check directly that Hamilton's equations are satisfied.

Remark 16.15 A classic problem is to determine the metric tensors $g_{ij}(q)$ for which these geodesic equations admit enough additional conservation laws to be integrable.

Exercise 16.16 Consider the Lagrangian

$$L_\epsilon(\mathbf{q},\dot{\mathbf{q}}) = \tfrac{1}{2}\|\dot{\mathbf{q}}\|^2 - \tfrac{1}{2\epsilon}(1-\|\mathbf{q}\|^2)^2$$

for a particle in $\mathbb{R}^3$. Let $\gamma_\epsilon(t)$ be the curve in $\mathbb{R}^3$ obtained by solving the Euler–Lagrange equations for L_ϵ with the initial conditions $\mathbf{q}_0 = \gamma_\epsilon(0), \dot{\mathbf{q}}_0 = \dot{\gamma}_\epsilon(0)$. Show that

$$\lim_{\epsilon\to 0} \gamma_\epsilon(t)$$

is a great circle on the two-sphere S^2, provided that $\mathbf{q}_0$ has unit length and the initial conditions satisfy $\mathbf{q}_0 \cdot \dot{\mathbf{q}}_0 = 0$.

Remark 16.17 The Lagrangian vector field associated to K is called the *geodesic spray*. Since the Legendre transformation is a diffeomorphism (in finite dimensions or in infinite dimensions if the metric is assumed to be strong), the geodesic spray is always a second order equation.

16.4 Covariant derivative

The variational approach to geodesics recovers the classical formulation using covariant derivatives, as follows. Let $\mathfrak{X}(Q)$ denote the set of vector fields on the manifold Q. The *covariant derivative*

$$\nabla\colon \mathfrak{X}(Q)\times\mathfrak{X}(Q)\to\mathfrak{X}(Q) \qquad (X,Y)\mapsto \nabla_X(Y),$$

of the Levi-Civita connection on (Q,g) is given in local charts by

$$\nabla_X(Y) = \Gamma^k_{ij}X^iY^j\frac{\partial}{\partial q^k} + X^i\frac{\partial Y^k}{\partial q^i}\frac{\partial}{\partial q^k}.$$

If $c(t)$ is a curve on Q and $Y \in \mathfrak{X}(Q)$, the covariant derivative of Y along $c(t)$ is defined by

$$\frac{DY}{Dt} := \nabla_{\dot c}Y,$$

or locally,

$$\left(\frac{DY}{Dt}\right)^k = \Gamma^k_{ij}(c(t))\dot c^i(t)Y^j(c(t)) + \frac{d}{dt}Y^k(c(t)).$$

A vector field is said to be *parallel transported* along $c(t)$ if

$$\frac{DY}{Dt} = 0.$$

Thus $\dot c(t)$ is parallel transported along $c(t)$ if and only if

$$\ddot c^{\,i} + \Gamma^i_{jk}\dot c^j\,\dot c^k = 0.$$

In classical differential geometry a *geodesic* is defined to be a curve $c(t)$ in Q whose tangent vector $\dot{c}(t)$ is parallel transported along $c(t)$. As the expression above shows, geodesics are integral curves of the Lagrangian vector field defined by the kinetic energy of g.

Definition 16.18 A *classical mechanical system* is given by a Lagrangian of the form $L(v_q) = K(v_q) - V(q)$, for $v_q \in T_q Q$. The smooth function $V\colon Q \to \mathbb{R}$ is called the *potential energy*. The total energy of this system is given by $E = K + V$ and the Euler–Lagrange equations (which are always second order for a hyperregular Lagrangian) are

$$\ddot{q}^{\,i} + \Gamma^i_{jk}\dot{q}^j\dot{q}^k + g^{il}\frac{\partial V}{\partial q^l} = 0, \quad i = 1, \dots n,$$

where g^{ij} are the entries of the inverse matrix of (g_{ij}).

Definition 16.19 If $Q = \mathbb{R}^3$ and the metric is given by $g_{ij} = \delta_{ij}$, these equations are Newton's equations of motion (9) of a particle in a potential field which launched our discussion in Lecture 13.

Exercise 16.20 (Gauge invariance) Show that the Euler–Lagrange equations are unchanged under

$$L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \to L' = L + \frac{d}{dt}\gamma(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \tag{31}$$

for any function $\gamma\colon \mathbb{R}^{6N} = \{(\mathbf{q}, \dot{\mathbf{q}}) \mid \mathbf{q}, \dot{\mathbf{q}} \in \mathbb{R}^{3N}\} \to \mathbb{R}$.

Exercise 16.21 (Generalized coordinate theorem) Show that the Euler–Lagrange equations are *unchanged in form* under any smooth invertible mapping $f\colon \{\mathbf{q} \mapsto \mathbf{s}\}$. That is, with

$$L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = \widetilde{L}(\mathbf{s}(t), \dot{\mathbf{s}}(t)), \tag{32}$$

show that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}} = 0 \quad \Longleftrightarrow \quad \frac{d}{dt}\left(\frac{\partial \widetilde{L}}{\partial \dot{\mathbf{s}}}\right) - \frac{\partial \widetilde{L}}{\partial \mathbf{s}} = 0. \tag{33}$$

Exercise 16.22 How do the Euler–Lagrange equations transform under $\mathbf{q}(t) = \mathbf{r}(t) + \mathbf{s}(t)$?

Exercise 16.23 (Other example Lagrangians) Write the Euler–Lagrange equations, then apply the Legendre transformation to determine the Hamiltonian and Hamilton's canonical equations for the following Lagrangians. Determine which of them are hyperregular.

- $L(q,\dot q) = \Big(g_{ij}(q)\dot q^i \dot q^j\Big)^{1/2}$ (Is it possible to assume that $L(q,\dot q) = 1$? Why?)
- $L(q,\dot q) = -\Big(1 - \dot{\mathbf{q}}\cdot\dot{\mathbf{q}}\Big)^{1/2}$
- $L(q,\dot q) = \frac{m}{2}\dot{\mathbf{q}}\cdot\dot{\mathbf{q}} + \frac{e}{c}\dot{\mathbf{q}}\cdot\mathbf{A}(\mathbf{q})$, for constants m, c and prescribed function $\mathbf{A}(\mathbf{q})$. How do the Euler–Lagrange equations for this Lagrangian differ from free motion in a moving frame with velocity $\frac{e}{mc}\mathbf{A}(\mathbf{q})$?

Example: charged particle in a magnetic field Consider a particle of charge e and mass m moving in a magnetic field $\mathbf{B}$, where $\mathbf{B} = \nabla\times\mathbf{A}$ is a given magnetic field on $\mathbb{R}^3$. The Lagrangian for the motion is given by the "minimal coupling" prescription (jay-dot-ay)

$$L(\mathbf{q},\dot{\mathbf{q}}) = \frac{m}{2}\|\dot{\mathbf{q}}\|^2 + \frac{e}{c}\mathbf{A}(\mathbf{q})\cdot\dot{\mathbf{q}},$$

in which the constant c is the speed of light. The derivatives of this Lagrangian are

$$\frac{\partial L}{\partial\dot{\mathbf{q}}} = m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A} =: \mathbf{p} \quad\text{and}\quad \frac{\partial L}{\partial\mathbf{q}} = \frac{e}{c}\nabla\mathbf{A}^T\cdot\dot{\mathbf{q}}$$

Hence, the Euler–Lagrange equations for this system are

$$m\,\ddot{\mathbf{q}} = \frac{e}{c}(\nabla\mathbf{A}^T\cdot\dot{\mathbf{q}} - \nabla\mathbf{A}\cdot\dot{\mathbf{q}}) = \frac{e}{c}\,\dot{\mathbf{q}}\times\mathbf{B}$$

(Newton's equations for the Lorentz force). The Lagrangian L is hyperregular, because

$$\mathbf{p} = \mathbb{F}L(\mathbf{q},\dot{\mathbf{q}}) = m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A}(\mathbf{q})$$

has the inverse

$$\dot{\mathbf{q}} = \mathbb{F}H(\mathbf{q},\mathbf{p}) = \frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{q})\right).$$

The corresponding Hamiltonian is given by the invertible change of variables,

$$H(\mathbf{q},\mathbf{p}) = \mathbf{p}\cdot\dot{\mathbf{q}} - L(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2m}\left\|\mathbf{p} - \frac{e}{c}\mathbf{A}\right\|^2. \tag{34}$$

The Hamiltonian H is hyperregular since

$$\dot{\mathbf{q}} = \mathbb{F}H(\mathbf{q},\mathbf{p}) = \frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) \quad\text{has the inverse}\quad \mathbf{p} = \mathbb{F}L(\mathbf{q},\dot{\mathbf{q}}) = m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A}.$$

The canonical equations for this Hamiltonian recover Newton's equations for the Lorentz force law.

Example: charged particle in a magnetic field by the Kaluza–Klein construction Although the minimal-coupling Lagrangian is not expressed as the kinetic energy of a metric, Newton's equations for the Lorentz force law may still be obtained as geodesic equations. This is accomplished by suspending them in a higher dimensional space via the *Kaluza–Klein construction*, which proceeds as follows.

Let Q_{KK} be the manifold $\mathbb{R}^3 \times S^1$ with variables $(\mathbf{q}, \theta)$. On Q_{KK} introduce the one-form $A + \mathbf{d}\theta$ (which defines a connection one-form on the trivial circle bundle $\mathbb{R}^3 \times S^1 \to \mathbb{R}^3$) and introduce the *Kaluza–Klein Lagrangian* L_{KK}: $TQ_{\mathrm{KK}} \simeq T\mathbb{R}^3 \times TS^1 \mapsto \mathbb{R}$ as

$$\begin{aligned} L_{\mathrm{KK}}(\mathbf{q}, \theta, \dot{\mathbf{q}}, \dot{\theta}) &= \tfrac{1}{2} m \|\dot{\mathbf{q}}\|^2 + \tfrac{1}{2} \big\| \big\langle A + \mathbf{d}\theta, (\mathbf{q}, \dot{\mathbf{q}}, \theta, \dot{\theta}) \big\rangle \big\|^2 \\ &= \tfrac{1}{2} m \|\dot{\mathbf{q}}\|^2 + \tfrac{1}{2} \big(\mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta} \big)^2. \end{aligned}$$

The Lagrangian L_{KK} is positive definite in $(\dot{\mathbf{q}}, \dot{\theta})$; so it may be regarded as the kinetic energy of a metric, the *Kaluza–Klein metric* on TQ_{KK}. (This construction fits the idea of $U(1)$ gauge symmetry for electromagnetic fields in $\mathbb{R}^3$. It can be generalized to a principal bundle with compact structure group endowed with a connection. The Kaluza–Klein Lagrangian in this generalization leads to Wong's equations for a color-charged particle moving in a classical Yang–Mills field.) The Legendre transformation for L_{KK} gives the momenta

$$\mathbf{p} = m\dot{\mathbf{q}} + (\mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta})\mathbf{A} \qquad \text{and} \qquad \pi = \mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta}. \tag{35}$$

Since L_{KK} does not depend on θ, the Euler–Lagrange equation

$$\frac{d}{dt} \frac{\partial L_{\mathrm{KK}}}{\partial \dot{\theta}} = \frac{\partial L_{\mathrm{KK}}}{\partial \theta} = 0,$$

shows that $\pi = \partial L_{\mathrm{KK}} / \partial \dot{\theta}$ is conserved. The *charge* is now defined by $e := c\pi$. The Hamiltonian H_{KK} associated to L_{KK} by the Legendre transformation (35) is

$$\begin{aligned} H_{\mathrm{KK}}(\mathbf{q}, \theta, \mathbf{p}, \pi) &= \mathbf{p} \cdot \dot{\mathbf{q}} + \pi \dot{\theta} - L_{\mathrm{KK}}(\mathbf{q}, \dot{\mathbf{q}}, \theta, \dot{\theta}) \\ &= \mathbf{p} \cdot \tfrac{1}{m} (\mathbf{p} - \pi \mathbf{A}) + \pi (\pi - \mathbf{A} \cdot \dot{\mathbf{q}}) - \tfrac{1}{2} m \|\dot{\mathbf{q}}\|^2 - \tfrac{1}{2} \pi^2 \\ &= \mathbf{p} \cdot \tfrac{1}{m} (\mathbf{p} - \pi \mathbf{A}) + \tfrac{1}{2} \pi^2 - \pi \mathbf{A} \cdot \tfrac{1}{m} (\mathbf{p} - \pi \mathbf{A}) - \tfrac{1}{2m} \|\mathbf{p} - \pi \mathbf{A}\|^2 \\ &= \tfrac{1}{2m} \|\mathbf{p} - \pi \mathbf{A}\|^2 + \tfrac{1}{2} \pi^2. \end{aligned} \tag{36}$$

On the constant level set $\pi = e/c$, the Kaluza–Klein Hamiltonian H_{KK} is a function of only the variables $(\mathbf{q}, \mathbf{p})$ and is equal to the Hamiltonian (34) for charged particle motion under the Lorentz force up to an additive constant. This example provides an easy but fundamental illustration of the geometry of (Lagrangian) reduction by symmetry. The canonical equations for the Kaluza–Klein Hamiltonian H_{KK} now reproduce Newton's equations for the Lorentz force law.

17 The rigid body in three dimensions

In the absence of external torques, Euler's equations for rigid body motion are:

$$\begin{aligned} I_1\dot{\Omega}_1 &= (I_2 - I_3)\Omega_2\Omega_3, \\ I_2\dot{\Omega}_2 &= (I_3 - I_1)\Omega_3\Omega_1, \\ I_3\dot{\Omega}_3 &= (I_1 - I_2)\Omega_1\Omega_2, \end{aligned} \tag{37}$$

or, equivalently,

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega},$$

where $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is the body angular velocity vector and I_1, I_2, I_3 are the moments of inertia of the rigid body.

Question 17.1 Can these equations – as they are written – be cast into Lagrangian or Hamiltonian form in any sense? (Since there are an odd number of equations, they cannot be put into canonical Hamiltonian form.)

We could reformulate them as:

- Euler–Lagrange equations on $T\mathrm{SO}(3)$ or
- Canonical Hamiltonian equations on $T^*\mathrm{SO}(3)$,

by using Euler angles and their velocities, or their conjugate momenta. However, these reformulations on $T\mathrm{SO}(3)$ or $T^*\mathrm{SO}(3)$ would answer a different question for a *six* dimensional system. We are interested in these structures for the equations as given above.

Answer 17.2 (Lagrangian formulation) The Lagrangian answer is this: These equations may be expressed in Euler–Poincaré form on the Lie algebra $\mathbb{R}^3$ using the Lagrangian

$$l(\mathbf{\Omega}) = \tfrac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2) = \tfrac{1}{2}\mathbf{\Omega}^T \cdot \mathbb{I}\mathbf{\Omega}, \tag{38}$$

which is the (rotational) kinetic energy of the rigid body.

The Hamiltonian answer to this question will be discussed later.

Proposition 17.3 *The Euler rigid body equations are equivalent to the* rigid body action principle *for a* reduced action

(39) $$\delta S_{\mathrm{red}} = \delta \int_a^b l(\mathbf{\Omega})\, dt = 0,$$

where variations of $\mathbf{\Omega}$ *are restricted to be of the form*

(40) $$\delta \mathbf{\Omega} = \dot{\mathbf{\Sigma}} + \mathbf{\Omega} \times \mathbf{\Sigma},$$

in which $\mathbf{\Sigma}(t)$ *is a curve in* $\mathbb{R}^3$ *that vanishes at the endpoints in time.*

Proof Since $l(\mathbf{\Omega}) = \frac{1}{2}\langle \mathbb{I}\mathbf{\Omega}, \mathbf{\Omega}\rangle$, and $\mathbb{I}$ is symmetric, we obtain

$$\begin{aligned}
\delta \int_a^b l(\mathbf{\Omega})\, dt &= \int_a^b \langle \mathbb{I}\mathbf{\Omega}, \delta\mathbf{\Omega}\rangle\, dt \\
&= \int_a^b \langle \mathbb{I}\mathbf{\Omega}, \dot{\mathbf{\Sigma}} + \mathbf{\Omega} \times \mathbf{\Sigma}\rangle\, dt \\
&= \int_a^b \left[\left\langle -\frac{d}{dt}\mathbb{I}\mathbf{\Omega}, \mathbf{\Sigma} \right\rangle + \langle \mathbb{I}\mathbf{\Omega}, \mathbf{\Omega} \times \mathbf{\Sigma}\rangle \right] dt \\
&= \int_a^b \left\langle -\frac{d}{dt}\mathbb{I}\mathbf{\Omega} + \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega}, \mathbf{\Sigma} \right\rangle dt,
\end{aligned}$$

upon integrating by parts and using the endpoint conditions, $\mathbf{\Sigma}(b) = \mathbf{\Sigma}(a) = 0$. Since $\mathbf{\Sigma}$ is otherwise arbitrary, (102) is equivalent to

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

which are Euler's equations (37). □

Let's derive this variational principle from the *standard* Hamilton's principle.

17.1 Hamilton's principle for rigid body motion on $T\,\mathrm{SO}(3)$

An element $\mathbf{R} \in \mathrm{SO}(3)$ gives the configuration of the body as a map of a *reference configuration* $\mathcal{B} \subset \mathbb{R}^3$ to the current configuration $\mathbf{R}(\mathcal{B})$; the map $\mathbf{R}$ takes a reference or label point $X \in \mathcal{B}$ to a current point $x = \mathbf{R}(X) \in \mathbf{R}(\mathcal{B})$.

When the rigid body is in motion, the matrix $\mathbf{R}$ is time-dependent. Thus,

$$x(t) = \mathbf{R}(t)X$$

with $\mathbf{R}(t)$ a curve parametrized by time in $\mathrm{SO}(3)$. The velocity of a point of the body is

$$\dot{x}(t) = \dot{\mathbf{R}}(t)X = \dot{\mathbf{R}}\mathbf{R}^{-1}(t)x(t).$$

Since $\mathbf{R}$ is an orthogonal matrix, $\mathbf{R}^{-1}\dot{\mathbf{R}}$ and $\dot{\mathbf{R}}\mathbf{R}^{-1}$ are skew matrices. Consequently, we can write (recall the hat map)

$$\dot{x} = \dot{\mathbf{R}}\mathbf{R}^{-1}x = \boldsymbol{\omega} \times x. \tag{41}$$

This formula defines the *spatial angular velocity vector* $\boldsymbol{\omega}$. Thus, $\boldsymbol{\omega}$ is essentially given by *right* translation of $\dot{\mathbf{R}}$ to the identity. That is, the vector

$$\boldsymbol{\omega} = \big(\dot{\mathbf{R}}\mathbf{R}^{-1}\big)\widehat{\ }.$$

The corresponding *body angular velocity* is defined by

$$\boldsymbol{\Omega} = \mathbf{R}^{-1}\boldsymbol{\omega}, \tag{42}$$

so that $\boldsymbol{\Omega}$ is the angular velocity relative to a body fixed frame. Notice that

$$\begin{aligned} \mathbf{R}^{-1}\dot{\mathbf{R}}X &= \mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}x = \mathbf{R}^{-1}(\boldsymbol{\omega} \times x) \\ &= \mathbf{R}^{-1}\boldsymbol{\omega} \times \mathbf{R}^{-1}x = \boldsymbol{\Omega} \times X, \end{aligned} \tag{43}$$

so that $\boldsymbol{\Omega}$ is given by *left* translation of $\dot{\mathbf{R}}$ to the identity. That is, the vector

$$\boldsymbol{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\widehat{\ }.$$

The *kinetic energy* is obtained by summing up $m|\dot{x}|^2/2$ (where $|\cdot|$ denotes the Euclidean norm) over the body. This yields

$$K = \frac{1}{2}\int_{\mathcal{B}} \rho(X)|\dot{\mathbf{R}}X|^2\, d^3X, \tag{44}$$

in which ρ is a given mass density in the reference configuration. Since

$$|\dot{\mathbf{R}}X| = |\boldsymbol{\omega} \times x| = |\mathbf{R}^{-1}(\boldsymbol{\omega} \times x)| = |\boldsymbol{\Omega} \times X|,$$

K is a quadratic function of $\boldsymbol{\Omega}$. Writing

$$K = \tfrac{1}{2}\boldsymbol{\Omega}^T \cdot \mathbb{I}\boldsymbol{\Omega} \tag{45}$$

defines the *moment of inertia tensor* $\mathbb{I}$, which, provided the body does not degenerate to a line, is a positive-definite (3×3) matrix, or better, a quadratic form. This quadratic form can be diagonalized by a change of basis; thereby defining the principal axes and moments of inertia. In this basis, we write $\mathbb{I} = \mathrm{diag}(I_1, I_2, I_3)$.

The function K is taken to be the Lagrangian of the system on $T\mathrm{SO}(3)$ (and by means of the Legendre transformation we obtain the corresponding Hamiltonian description on $T^*\mathrm{SO}(3)$). Notice that K in equation (44) is *left* (not right) invariant on $T\mathrm{SO}(3)$, since

$$\mathbf{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\hat{}.$$

It follows that the corresponding Hamiltonian will also be *left* invariant.

In the framework of Hamilton's principle, the relation between motion in $\mathbf{R}$ space and motion in body angular velocity (or $\mathbf{\Omega}$) space is as follows.

Proposition 17.4 *The curve $\mathbf{R}(t) \in \mathrm{SO}(3)$ satisfies the Euler–Lagrange equations for*

(46) $$L(\mathbf{R}, \dot{\mathbf{R}}) = \frac{1}{2}\int_{\mathcal{B}} \rho(X)|\dot{\mathbf{R}}X|^2\, d^3X,$$

if and only if $\mathbf{\Omega}(t)$ defined by $\mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{v} = \mathbf{\Omega}\times\mathbf{v}$ for all $\mathbf{v}\in\mathbb{R}^3$ satisfies Euler's equations

(47) $$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega}\times\mathbf{\Omega}.$$

The proof of this relation will illustrate how to reduce variational principles using their symmetry groups. By Hamilton's principle, $\mathbf{R}(t)$ satisfies the Euler–Lagrange equations, if and only if

$$\delta\int L(\mathbf{R}, \dot{\mathbf{R}})\, dt = 0.$$

Let $l(\mathbf{\Omega}) = \frac{1}{2}(\mathbb{I}\mathbf{\Omega})\cdot\mathbf{\Omega}$, so that $l(\mathbf{\Omega}) = L(\mathbf{R}, \dot{\mathbf{R}})$ where the matrix $\mathbf{R}$ and the vector $\mathbf{\Omega}$ are related by the hat map, $\mathbf{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\hat{}$. Thus, the Lagrangian L is left SO(3)-invariant. That is,

$$l(\mathbf{\Omega}) = L(\mathbf{R}, \dot{\mathbf{R}}) = L(\mathbf{e}, \mathbf{R}^{-1}\dot{\mathbf{R}}).$$

To see how we should use this left-invariance to transform Hamilton's principle, define the skew matrix $\widehat{\mathbf{\Omega}}$ by $\widehat{\mathbf{\Omega}}\mathbf{v} = \mathbf{\Omega}\times\mathbf{v}$ for any $\mathbf{v}\in\mathbb{R}^3$.

We differentiate the relation $\mathbf{R}^{-1}\dot{\mathbf{R}} = \widehat{\mathbf{\Omega}}$ with respect to $\mathbf{R}$ to get

(48) $$-\mathbf{R}^{-1}(\delta\mathbf{R})\mathbf{R}^{-1}\dot{\mathbf{R}} + \mathbf{R}^{-1}(\delta\dot{\mathbf{R}}) = \widehat{\delta\mathbf{\Omega}}.$$

Let the skew matrix $\widehat{\mathbf{\Sigma}}$ be defined by

(49) $$\widehat{\mathbf{\Sigma}} = \mathbf{R}^{-1}\delta\mathbf{R},$$

and define the vector $\mathbf{\Sigma}$ by

(50) $$\widehat{\mathbf{\Sigma}}\mathbf{v} = \mathbf{\Sigma}\times\mathbf{v}.$$

Note that

$$\dot{\widehat{\mathbf{\Sigma}}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}\delta\mathbf{R} + \mathbf{R}^{-1}\delta\dot{\mathbf{R}},$$

so

$$\mathbf{R}^{-1}\delta\dot{\mathbf{R}} = \dot{\widehat{\boldsymbol{\Sigma}}} + \mathbf{R}^{-1}\dot{\mathbf{R}}\widehat{\boldsymbol{\Sigma}}. \tag{51}$$

Substituting (51) and (49) into (48) gives

$$-\widehat{\boldsymbol{\Sigma}}\,\widehat{\boldsymbol{\Omega}} + \dot{\widehat{\boldsymbol{\Sigma}}} + \widehat{\boldsymbol{\Omega}}\,\widehat{\boldsymbol{\Sigma}} = \widehat{\delta\boldsymbol{\Omega}}\,,$$

that is,

$$\widehat{\delta\boldsymbol{\Omega}} = \dot{\widehat{\boldsymbol{\Sigma}}} + [\widehat{\boldsymbol{\Omega}}, \widehat{\boldsymbol{\Sigma}}]. \tag{52}$$

The identity $[\widehat{\boldsymbol{\Omega}}, \widehat{\boldsymbol{\Sigma}}] = (\boldsymbol{\Omega}\times\boldsymbol{\Sigma})\widehat{\ }$ holds by Jacobi's identity for the cross product and so

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega}\times\boldsymbol{\Sigma}. \tag{53}$$

These calculations prove the following:

Theorem 17.5 *For a Lagrangian which is left-invariant under SO(3), Hamilton's variational principle*

$$\delta S = \delta\int_a^b L(\mathbf{R}, \dot{\mathbf{R}})\,dt = 0 \tag{54}$$

on $T\mathrm{SO}(3)$ *is equivalent to the* reduced variational principle

$$\delta S_{\mathrm{red}} = \delta\int_a^b l(\boldsymbol{\Omega})\,dt = 0 \tag{55}$$

with $\boldsymbol{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\widehat{\ }$ *on* $\mathbb{R}^3$ *where the variations* $\delta\boldsymbol{\Omega}$ *are of the form*

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega}\times\boldsymbol{\Sigma},$$

with $\boldsymbol{\Sigma}(a) = \boldsymbol{\Sigma}(b) = 0$.

Reconstruction of $\mathbf{R}(t)\in SO(3)$ In Theorem 17.5, Euler's equations for the rigid body

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

follow from the reduced variational principle (55) for the Lagrangian

$$l(\boldsymbol{\Omega}) = \tfrac{1}{2}(\mathbb{I}\boldsymbol{\Omega})\cdot\boldsymbol{\Omega}, \tag{56}$$

which is expressed in terms of the left-invariant time-dependent angular velocity in the body, $\boldsymbol{\Omega}\in\mathfrak{so}(3)$. The body angular velocity $\boldsymbol{\Omega}(t)$ yields the tangent vector

$\dot{\mathbf{R}}(t) \in T_{\mathbf{R}(t)}SO(3)$ along the integral curve in the rotation group $\mathbf{R}(t) \in SO(3)$ by the relation,

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

This relation provides the *reconstruction formula*. It's solution as a linear differential equation with time-dependent coefficients yields the integral curve $\mathbf{R}(t) \in SO(3)$ for the orientation of the rigid body, once the time dependence of $\mathbf{\Omega}(t)$ is determined from the Euler equations.

17.2 Hamiltonian form of rigid body motion

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

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

in terms of a Poisson bracket operation,

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

which is bilinear, skew-symmetric and satisfies the Jacobi identity and (usually) the Leibniz rule.

As we shall explain, reduced equations arising from group-invariant Hamilton's principles on Lie groups are naturally Hamiltonian. If we *Legendre transform* our reduced Lagrangian for the $SO(3)$ left invariant variational principle (55) for rigid body dynamics, then its simple, beautiful and well-known Hamiltonian formulation emerges.

Definition 17.6 The Legendre transformation $\mathbb{F}l\colon \mathfrak{so}(3) \to \mathfrak{so}(3)^*$ is defined by

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

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 (56), the $\mathbb{R}^3$ components of the body angular momentum are

$$\Pi_i = I_i \Omega_i = \frac{\partial l}{\partial \Omega_i}, \quad i = 1, 2, 3. \tag{57}$$

17.3 Lie–Poisson Hamiltonian formulation of rigid body dynamics

Let

$$h(\Pi) := \langle \Pi, \Omega \rangle - l(\Omega),$$

where the pairing $\langle \cdot, \cdot \rangle \colon \mathfrak{so}(3)^* \times \mathfrak{so}(3) \to \mathbb{R}$ is understood in components as the vector dot product on $\mathbb{R}^3$

$$\langle \Pi, \Omega \rangle := \mathbf{\Pi} \cdot \mathbf{\Omega}.$$

Hence, one finds the expected expression for the rigid-body Hamiltonian

$$h = \tfrac{1}{2} \mathbf{\Pi} \cdot \mathbb{I}^{-1} \mathbf{\Pi} := \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3}. \tag{58}$$

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

$$\frac{\partial h}{\partial \Pi} = \Omega + \left\langle \Pi, \frac{\partial \Omega}{\partial \Pi} \right\rangle - \left\langle \frac{\partial l}{\partial \Omega}, \frac{\partial \Omega}{\partial \Pi} \right\rangle = \Omega.$$

In $\mathbb{R}^3$ coordinates, this relation expresses the body angular velocity as the derivative of the reduced Hamiltonian with respect to the body angular momentum, namely (introducing grad-notation),

$$\nabla_\Pi h := \frac{\partial h}{\partial \mathbf{\Pi}} = \mathbf{\Omega}.$$

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:

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

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

$$\{f, h\}(\mathbf{\Pi}) := -\mathbf{\Pi} \cdot (\nabla_\Pi f \times \nabla_\Pi h). \tag{59}$$

For the Hamiltonian (58), one checks that the Euler equations in terms of the rigid-body angular momenta,

$$\dot{\Pi}_1 = \frac{I_2 - I_3}{I_2 I_3} \Pi_2 \Pi_3, \quad \dot{\Pi}_2 = \frac{I_3 - I_1}{I_3 I_1} \Pi_3 \Pi_1, \quad \dot{\Pi}_3 = \frac{I_1 - I_2}{I_1 I_2} \Pi_1 \Pi_2, \tag{60}$$

that is,

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega}. \tag{61}$$

are equivalent to

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

The Poisson bracket proposed in (59) is an example of a *Lie Poisson bracket*, which we will show separately satisfies the defining relations to be a Poisson bracket.

17.4 $\mathbb{R}^3$ Poisson bracket

The rigid body Poisson bracket (59) is a special case of the Poisson bracket for functions on $\mathbb{R}^3$,

$$\{f,h\} = -\nabla c \cdot \nabla f \times \nabla h \tag{62}$$

This bracket generates the motion

$$\dot{\mathbf{x}} = \{\mathbf{x}, h\} = \nabla c \times \nabla h \tag{63}$$

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 Marsden and Ratiu [2003].)

Exercise 17.7 Consider the $\mathbb{R}^3$ Poisson bracket

$$\{f,h\} = -\nabla c \cdot \nabla f \times \nabla h \tag{64}$$

Let $c = \mathbf{x}^T \cdot \mathbb{C}\mathbf{x}$ be a quadratic form on $\mathbb{R}^3$, and let $\mathbb{C}$ be the associated symmetric 3×3 matrix. Determine the conditions on the quadratic function $c(\mathbf{x})$ so that this Poisson bracket will satisfy the Jacobi identity.

Exercise 17.8 Find the general conditions on the function $\mathbf{c}(\mathbf{x})$ 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?

Exercise 17.9 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 17.10 (Casimirs of the $\mathbb{R}^3$ bracket) The Casimirs (or distinguished functions, as Lie called them) of a Poisson bracket satisfy

$$\{c,h\}(\mathbf{x}) = 0, \quad \forall h(\mathbf{x}).$$

Suppose the function $\mathbf{c}(\mathbf{x})$ is chosen so that the $\mathbb{R}^3$ bracket (62) satisfies the defining properties of a Poisson bracket. What are the Casimirs for the $\mathbb{R}^3$ bracket (62)? Why?

Exercise 17.11 Show that the motion equation

$$\dot{\mathbf{x}} = \{\mathbf{x}, h\}$$

for the $\mathbb{R}^3$ bracket (62) is invariant under a certain linear combination of the functions c and h. Interpret this invariance geometrically.

18 Momentum maps

The main idea Symmetries are often associated with conserved quantities. For example, the flow of any $SO(3)$–invariant Hamiltonian vector field on $T^*\mathbb{R}^3$ conserves angular momentum, $\mathbf{q}\times\mathbf{p}$. More generally, given a Hamiltonian H on a phase space P, and a group action of G on P that conserves H, there is often an associated "momentum map" $J\colon P\to\mathfrak{g}^*$ that is conserved by the flow of the Hamiltonian vector field.

Note: all group actions in this section will be left actions until otherwise specified.

18.1 Hamiltonian systems on Poisson manifolds

Definition 18.1 A *Poisson bracket* on a manifold P is a skew-symmetric bilinear operation on

$$\mathcal{F}(P) := \mathcal{C}^\infty(P,\mathbb{R})$$

satisfying the Jacobi identity and the Leibniz identity,

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

The pair $(P,\{\cdot,\cdot\})$ is called a *Poisson manifold.*

Remark 18.2 The Leibniz identity is sometimes not included in the definition. Note that bilinearity, skew-symmetry and the Jacobi identity are the axioms of a Lie algebra. In what follows, a Poisson bracket is a binary operation that makes $\mathcal{F}(P)$ into a Lie algebra and also satisfies the Leibniz identity.

Exercise 18.3 Show that the *classical Poisson bracket*, defined in cotangent-lifted coordinates

$$\left(q^1,\dots,q^N,p_1,\dots,p_N\right)$$

on an $2N$–dimensional cotangent bundle T^*Q by

$$\{F,G\} = \sum_{i=1}^N\left(\frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i}\right),$$

satisfies the axioms of a Poisson bracket. Show also that the definition of this bracket is independent of the choice of local coordinates $\left(q^1, \ldots, q^N\right)$.

Definition 18.4 A *Poisson map* between two Poisson manifolds is a map

$$\varphi\colon \left(P_1, \{\cdot,\cdot\}_1\right) \to \left(P_2, \{\cdot,\cdot\}_2\right)$$

that preserves the brackets, meaning

$$\{F \circ \varphi, G \circ \varphi\}_1 = \{F, G\}_2 \circ \varphi, \qquad \text{for all } F, G \in \mathcal{F}(P_2).$$

Definition 18.5 An action Φ of G on a Poisson manifold $(P, \{,\})$ is *canonical* if Φ_g is a Poisson map for every g, that is,

$$\left\{F \circ \Phi_g, K \circ \Phi_g\right\} = \{F, K\} \circ \Phi_g$$

for every $F, K \in \mathcal{F}(P)$.

Definition 18.6 Let $(P, \{\cdot,\cdot\})$ be a Poisson manifold, and let $H\colon P \to \mathbb{R}$ be differentiable. The *Hamiltonian vector field* for H is the vector field X_H defined by

$$X_H(F) = \{F, H\}, \qquad \text{for any } F \in \mathcal{F}(P)$$

Remark 18.7 X_H is well-defined because of the Leibniz identity and the correspondence between vector fields and derivations (see Lee [38]).

Remark 18.8 $X_H(F) = \pounds_{X_H} F = \dot{F}$, the Lie derivative of F along the flow of X_H. The equations

$$\dot{F} = \{F, H\},$$

called "Hamilton's equations", have already appeared in Theorem 13.2, and are an equivalent definition of X_H.

Exercise 18.9 Show that Hamilton's equations for the classical Poisson bracket are the canonical Hamilton's equations,

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}.$$

18.2 Infinitesimal invariance under Hamiltonian vector fields

Let G act smoothly on P, and let $\xi \in \mathfrak{g}$. Recall (from Lecture 11) that the infinitesimal generator ξ_P is the vector field on P defined by

$$\xi_P(x) = \frac{d}{dt} g(t)x \bigg|_{t=0},$$

for some path $g(t)$ in G such that $g(0) = e$ and $g'(0) = \xi$.

Remark 18.10 For matrix groups, we can take $g(t) = \exp(t\xi)$. This works in general for the exponential map of an arbitrary Lie group. For matrix groups,

$$\xi_P(\mathbf{x}) = \frac{d}{dt} \exp(t\xi)\mathbf{x} \bigg|_{t=0} = \xi\mathbf{x} \qquad \text{(matrix multiplication)}.$$

Exercise 18.11 If $H\colon P \to \mathbb{R}$ is G–invariant, meaning that $H(gx) = H(x)$ for all $g \in G$ and $x \in P$, then $\pounds_{\xi_P} H = 0$ for all $\xi \in \mathfrak{g}$. This property is called *infinitesimal invariance*.

Example 18.12 (The momentum map for the rotation group) Consider the cotangent bundle of ordinary Euclidean space $\mathbb{R}^3$. This is the Poisson (symplectic) manifold with coordinates $(\mathbf{q}, \mathbf{p}) \in T^*\mathbb{R}^3 \simeq \mathbb{R}^6$, equipped with the canonical Poisson bracket. An element g of the rotation group $SO(3)$ acts on $T^*\mathbb{R}^3$ according to

$$g(\mathbf{q}, \mathbf{p}) = (g\mathbf{q}, g\mathbf{p})$$

Set $g(t) = \exp(tA)$, so that $\frac{d}{dt}\big|_{t=0} g(t) = A$ and the corresponding Hamiltonian vector field is

$$X_A = (\dot{\mathbf{q}}, \dot{\mathbf{p}}) = (A\mathbf{q}, A\mathbf{p})$$

where $A \in so(3)$ is a skew-symmetric matrix. The corresponding Hamiltonian equations read

$$\dot{\mathbf{q}} = A\mathbf{q} = \frac{\partial J_A}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = A\mathbf{p} = -\frac{\partial J_A}{\partial \mathbf{q}}.$$

Hence,

$$J_A(\mathbf{q}, \mathbf{p}) = -A\mathbf{p} \cdot \mathbf{q} = a_i \epsilon_{ijk} p_k q_j = \mathbf{a} \cdot \mathbf{q} \times \mathbf{p}.$$

for a vector $\mathbf{a} \in \mathbb{R}^3$ with components a_i, $i = 1, 2, 3$. So the momentum map for the rotation group is the angular momentum $J = \mathbf{q} \times \mathbf{p}$.

Example 18.13 Consider angular momentum $J = \mathbf{q} \times \mathbf{p}$, defined on $P = T^*\mathbb{R}^3$. For every $\xi \in \mathbb{R}^3$, define

$$J_\xi(\mathbf{q}, \mathbf{p}) := \xi \cdot (\mathbf{q} \times \mathbf{p}) = \mathbf{p} \cdot (\xi \times \mathbf{q})$$

Using Exercise 18.9 and Remark 18.10,

$$X_{J_\xi}(\mathbf{q},\mathbf{p}) = \left(\frac{\partial J_\xi}{\partial \mathbf{p}}, -\frac{\partial J_\xi}{\partial \mathbf{q}}\right) = (\xi \times \mathbf{q}, \xi \times \mathbf{p}) = \widehat{\xi}_P(\mathbf{q},\mathbf{p}),$$

where the last line is the infinitesimal generator corresponding to $\widehat{\xi} \in so(3)$. Now suppose $H\colon P \to \mathbb{R}$ is $SO(3)$–invariant. From Exercise 18.11, we have $\pounds_{\widehat{\xi}} H = 0$. It follows that

$$\pounds_{X_H} J_\xi = \left\{J_\xi, H\right\} = -\left\{H, J_\xi\right\} = -\pounds_{X_{J_\xi}} H = -\pounds_{\xi_P} H = 0.$$

Since this holds for all ξ, we have shown that J is conserved by the Hamiltonian flow.

18.3 Defining momentum maps

In order to generalise this example, we recast it using the hat map $\widehat{\ }\colon \mathbb{R}^3 \to so(3)$ and the associated map $\widetilde{\ }\colon \left(\mathbb{R}^3\right)^* \to so(3)^*$, and the standard identification $\left(\mathbb{R}^3\right)^* \cong \mathbb{R}^3$ via the Euclidean dot product. We consider J as a function from P to $so(3)^*$ given by $J\big(\mathbf{q},\mathbf{p}\big) = (\mathbf{q}\times\mathbf{p})\widetilde{\ }$. For any $\xi = \widehat{\mathbf{v}}$, we define $J_\xi(\mathbf{q},\mathbf{p}) = \big\langle (\mathbf{q}\times\mathbf{p})\widetilde{\ }, \widehat{\mathbf{v}}\big\rangle = (\mathbf{q}\times\mathbf{p})\cdot\mathbf{v}$. As before, we find that $X_{J_\xi} = \xi_P$ for every ξ, and J is conserved by the Hamiltonian flow. We take the first property, $X_{J_\xi} = \xi_P$, as the general definition of a momentum map. The conservation of J follows by the same Poisson bracket calculation as in the example; the result is Noether's Theorem.

Definition 18.14 A *momentum map* for a canonical action of G on P is a map $J\colon P \to \mathfrak{g}^*$ such that, for every $\xi \in \mathfrak{g}$, the map $J_\xi\colon P \to \mathbb{R}$ defined by $J_\xi(p) = \langle J(p), \xi\rangle$ satisfies

$$X_{J_\xi} = \xi_P$$

Theorem 18.15 (Noether's Theorem) *Let G act canonically on $(P, \{\cdot,\cdot\})$ with momentum map J. If H is G–invariant, then J is conserved by the flow of X_H.*

Proof For every $\xi \in \mathfrak{g}$,

$$\pounds_{X_H} J_\xi = \left\{J_\xi, H\right\} = -\left\{H, J_\xi\right\} = -\pounds_{X_{J_\xi}} H = -\pounds_{\xi_P} H = 0. \qquad \square$$

Exercise 18.16 Momentum maps are unique up to a choice of a constant element of $\mathfrak{g}^*$ on every connected component of M.

Exercise 18.17 Show that the S^1 action on the torus $T^2 := S^1 \times S^1$ given by $\alpha(\theta,\phi) = (\alpha+\theta,\phi)$ is canonical with respect to the classical bracket (with θ,ϕ in place of q, p), but doesn't have a momentum map.

Exercise 18.18 Show that the Petzval invariant for Fermat's principle in axisymmetric, translation-invariant media is a momentum map, $T^*\mathbb{R}^2 \mapsto sp(2,\mathbb{R})^*$ taking $(\mathbf{q},\mathbf{p}) \mapsto (X,Y,Z)$. What is its symmetry? What is its Hamiltonian vector field?

Theorem 18.19 (also due to Noether) *Let G act on Q, and by cotangent lifts on T^*Q. Then $J\colon T^*Q \to \mathfrak{g}^*$ defined by, for every $\xi \in \mathfrak{g}$,*

$$J_\xi\left(\alpha_q\right) = \left\langle \alpha_q, \xi_Q\left(q\right)\right\rangle, \text{ for every } \alpha_q \in T_q^*Q,$$

is a momentum map (the "standard one") for the G action with respect to the classical Poisson bracket.

(A proof using symplectic forms is given in Marsden and Ratiu [45].)

Proof We need to show that $X_{J_\xi} = \xi_{T^*Q}$, for every $\xi \in \mathfrak{g}$. From the definition of Hamiltonian vector fields, this is equivalent to showing that $\xi_{T^*Q}[F] = \{F, J_\xi\}$ for every $F \in \mathcal{F}(T^*Q)$. We verify this for finite-dimensional Q by using cotangent-lifted local coordinates.

$$\begin{aligned}
\frac{\partial J_\xi}{\partial p}(q,p) &= \xi_Q(q) \\
\frac{\partial J_\xi}{\partial q^i}(q,p) &= \left\langle p, \frac{\partial}{\partial q^i}\left(\xi_Q(q)\right)\right\rangle \\
&= \left\langle p, \frac{\partial}{\partial q^i}\left(\frac{\partial}{\partial t}\Phi_{(\exp(t\xi))}(q)\Big|_{t=0}\right)\right\rangle \quad = \left\langle p, \frac{\partial}{\partial t}\left(\frac{\partial}{\partial q^i}\Phi_{(\exp(t\xi))}(q)\right)\Big|_{t=0}\right\rangle \\
&= \frac{\partial}{\partial t}\left\langle p, T\Phi_{(\exp(t\xi))}\frac{\partial}{\partial q^i}(q)\right\rangle\Big|_{t=0} \quad = \frac{\partial}{\partial t}\left\langle T^*\Phi_{(\exp(t\xi))}\,p, \frac{\partial}{\partial q^i}(q)\right\rangle\Big|_{t=0} \\
&= \left\langle -\xi_{T^*Q}(q,p), \frac{\partial}{\partial q^i}(q)\right\rangle \\
\frac{\partial J_\xi}{\partial q}(q,p) &= -\xi_{T^*Q}(q,p)
\end{aligned}$$

So for every $F \in \mathcal{F}(T^*Q)$,

$$\begin{aligned}
\xi_{T^*Q}[F] &= \frac{\partial}{\partial t}F\left(\exp(t\xi)q, \exp(t\xi)p\right)\Big|_{t=0} \\
&= \frac{\partial F}{\partial q}\xi_Q(q) + \frac{\partial F}{\partial p}\xi_{T^*Q}(q,p) = \frac{\partial F}{\partial q}\frac{\partial J_\xi}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial J_\xi}{\partial q} = \{F, J_\xi\}
\end{aligned}$$

which completes the proof. □

Example 18.20 Let $G \subset M_n(\mathbb{R})$ be a matrix group, with cotangent-lifted action on $(q,p) \in T^*\mathbb{R}^n$. For every $\mathrm{g} \subset M_n(\mathbb{R})$, $q \mapsto gq$. The cotangent-lifted action is $(q,p) \mapsto (gq, g^{-T}p)$. Thus, writing $g = \exp(t\xi)$, the linearization of this group action yields the vector field

$$X_\xi = (\xi q, -\xi^T p)$$

The corresponding Hamiltonian equations read

$$\xi q = \frac{\partial J_\xi}{\partial p}, \qquad -\xi^T p = -\frac{\partial J_\xi}{\partial q}$$

This yields the momentum map $J(q,p)$ given by

$$J_\xi(q,p) = \langle J(q,p), \xi \rangle = p^T \xi_Q(q) = p^T \xi q.$$

In coordinates, $p^T \xi q = p_i \xi^i_j q^j$, so $J(q,p) = q^i p_j$.

Exercise 18.21 Calculate the momentum map of the cotangent lifted action of the group of translations of $\mathbb{R}^3$.

Solution The element $\mathbf{x} \in \mathbb{R}^3$ acts on $\mathbf{q} \in \mathbb{R}^3$ by addition of vectors,

$$\mathbf{x} \cdot (\mathbf{q}) = \mathbf{q} + \mathbf{x}.$$

The infinitesimal generator is $\lim_{\mathbf{x} \to 0} \frac{d}{d\mathbf{x}}(\mathbf{q} + \mathbf{x}) = \mathrm{Id}$. Thus, $\xi_{\mathbf{q}} = \mathrm{Id}$ and

$$\langle J_k, \xi \rangle = \langle (\mathbf{q}, \mathbf{p}), \xi_{\mathbf{q}} \rangle = \langle \mathbf{p}, \mathrm{Id} \rangle = p_i \delta^i_k = p_k$$

This is also Hamiltonian with $J_\xi = \mathbf{p}$, so that $\{\mathbf{p}, J_\xi\} = 0$ and $\{\mathbf{q}, J_\xi\} = \mathrm{Id}$.

Example 18.22 Let G act on itself by left multiplication, and by cotangent lifts on T^*G. We first note that the infinitesimal action on G is

$$\xi_G(g) = \frac{d}{dt} \exp(t\xi) g \bigg|_{t=0} = TR_g \xi.$$

Let J_L be the momentum map for this action. Let $\alpha_g \in T^*_g G$. For every $\xi \in \mathfrak{g}$, we have

$$\left\langle J_L(\alpha_g), \xi \right\rangle = \left\langle \alpha_g, \xi_G(g) \right\rangle = \left\langle \alpha_g, TR_g \xi \right\rangle = \left\langle TR_g^* \alpha_g, \xi \right\rangle$$

so $J_L(\alpha_g) = TR_g^* \alpha_g$. Alternatively, writing $\alpha_g = T^* L_{g^{-1}} \mu$ for some $\mu \in \mathfrak{g}^*$ we have

$$J_L\left(T^* L_{g^{-1}} \mu\right) = TR_g^* T^* L_{g^{-1}} \mu = Ad^*_{g^{-1}} \mu.$$

Exercise 18.23 Show that the momentum map for the right multiplication action $R_g(h) = hg$ is $J_R(\alpha_g) = TL_g^* \alpha_g$.

For matrix groups, the tangent lift of the left (or right) multiplication action is again matrix multiplication. Indeed, to compute $TR_G(A)$ for any $A \in T_Q SO(3)$, let $B(t)$ be a path in $SO(3)$ such that $B(0) = Q$ and $B'(0) = A$. Then

$$TR_G(A) = \frac{d}{dt} B(t)G \bigg|_{t=0} = AG.$$

Similarly, $TL_G(A) = GA$. To compute the cotangent lift similarly, we need to be able to consider elements of T^*G as matrices. This can be done using any nondegenerate bilinear form on each tangent space $T_Q G$. We will use the pairing defined by

$$\langle\langle A, B \rangle\rangle := -\tfrac{1}{2} \operatorname{tr}\left(A^T B\right) = -\tfrac{1}{2} \operatorname{tr}\left(A B^T\right).$$

(The equivalence of the two formulas follows from the properties $\operatorname{tr}(CD) = \operatorname{tr}(DC)$ and $\operatorname{tr}(C^T) = \operatorname{tr}(C)$).

Exercise 18.24 Check that this pairing, restricted to $so(3)$, corresponds to the Euclidean inner product via the hat map.

Example 18.25 Consider the previous example for a *matrix* group G. For any $Q \in G$, the pairing given above allows use to consider any element $P \in T_Q^* G$ as a matrix. The natural pairing of $T_Q^* G$ with $T_Q G$ now has the formula,

$$\langle P, A \rangle = -\tfrac{1}{2} \operatorname{tr}\left(P^T A\right), \quad \text{for all } A \in T_Q G.$$

We compute the cotangent-lifts of the left and right multiplication actions:

$$\begin{aligned}
\left\langle T^* L_Q(P), A \right\rangle &= \left\langle P, TL_Q(A) \right\rangle = \left\langle P, QA \right\rangle \\
&= -\tfrac{1}{2} \operatorname{tr}\left(P^T QA\right) = -\tfrac{1}{2} \operatorname{tr}\left((Q^T P)^T A\right) = \left\langle Q^T P, A \right\rangle \\
\left\langle T^* R_Q(P), A \right\rangle &= \left\langle P, TR_Q(A) \right\rangle = \langle P, AQ \rangle \\
&= -\tfrac{1}{2} \operatorname{tr}\left(P(AQ)^T\right) = -\tfrac{1}{2} \operatorname{tr}\left(PQ^T A^T\right) = \left\langle PQ^T, A \right\rangle
\end{aligned}$$

In summary,

$$T^* L_Q(P) = Q^T P \quad \text{and} \quad T^* R_Q(P) = PQ^T$$

We thus compute the momentum maps as

$$\begin{aligned}
J_L(Q, P) &= T^* R_Q P = PQ^T \\
J_R(Q, P) &= T^* L_Q P = Q^T P
\end{aligned}$$

In the special case of $G = SO(3)$, these matrices PQ^T and $Q^T P$ are skew-symmetric, since they are elements of $so(3)$. Therefore,

$$J_L(Q, P) = T^* R_Q P = \tfrac{1}{2}\big(PQ^T - QP^T\big)$$
$$J_R(Q, P) = T^* L_Q P = \tfrac{1}{2}\big(Q^T P - P^T Q\big)$$

Exercise 18.26 Show that the cotangent lifted action on $SO(n)$ is expressed as

$$Q \cdot P = Q^T P$$

as matrix multiplication.

Definition 18.27 A momentum map is said to be *equivariant* when it is equivariant with respect to the given action on P and the coadjoint action on $\mathfrak{g}^*$. That is,

$$J(g \cdot p) = \mathrm{Ad}^*_{g^{-1}} J(p)$$

for every $g \in G$, $p \in P$, where $g \cdot p$ denotes the action of g on the point p and where Ad denotes the adjoint action.

Exercise 18.28 Show that the momentum map derived from the cotangent lift in Theorem 18.19 is equivariant.

Example 18.29 (Momentum map for symplectic representations) Let (V, Ω) be a symplectic vector space and let G be a Lie group acting linearly and symplectically on V. This action admits an equivariant momentum map $\mathbf{J}\colon V \to \mathfrak{g}$ given by

$$J^\xi(v) = \langle \mathbf{J}(v), \xi \rangle = \tfrac{1}{2}\Omega(\xi \cdot v, v),$$

where $\xi \cdot v$ denotes the Lie algebra representation of the element $\xi \in \mathfrak{g}$ on the vector $v \in V$. To verify this, note that the infinitesimal generator $\xi_V(v) = \xi \cdot v$, by the definition of the Lie algebra representation induced by the given Lie group representation, and that $\Omega(\xi \cdot u, v) = -\Omega(u, \xi \cdot v)$ for all $u, v \in V$. Therefore

$$\mathbf{d}J^\xi(u)(v) = \tfrac{1}{2}\Omega(\xi \cdot u, v) + \tfrac{1}{2}\Omega(\xi \cdot v, u) = \Omega(\xi \cdot u, v).$$

Equivariance of $\mathbf{J}$ follows from the obvious relation $g^{-1} \cdot \xi \cdot g \cdot v = (\mathrm{Ad}_{g^{-1}} \xi) \cdot v$ for any $g \in G$, $\xi \in \mathfrak{g}$, and $v \in V$.

Example 18.30 (Cayley–Klein parameters and the Hopf fibration) Consider the natural action of $SU(2)$ on $\mathbb{C}^2$. Since this action is by isometries of the Hermitian metric, it is automatically symplectic and therefore has a momentum map $\mathbf{J}\colon \mathbb{C}^2 \to \mathfrak{su}(2)^*$ given in Example 18.29, that is,

$$\langle \mathbf{J}(z,w), \xi\rangle = \tfrac{1}{2}\Omega(\xi\cdot(z,w),(z,w)),$$

where $z,w\in\mathbb{C}$ and $\xi\in\mathfrak{su}(2)$. Now the symplectic form on $\mathbb{C}^2$ is given by minus the imaginary part of the Hermitian inner product. That is, $\mathbb{C}^n$ has Hermitian inner product given by $\mathbf{z}\cdot\mathbf{w} := \sum_{j=1}^n z_j\overline{w}_j$, where $\mathbf{z}=(z_1,\dots,z_n), \mathbf{w}=(w_1,\dots,w_n)\in\mathbb{C}^n$. The symplectic form is thus given by $\Omega(\mathbf{z},\mathbf{w}) := -\operatorname{Im}(\mathbf{z}\cdot\mathbf{w})$ and it is identical to the one given before on $\mathbb{R}^{2n}$ by identifying $\mathbf{z}=\mathbf{u}+i\mathbf{v}\in\mathbb{C}^n$ with $(\mathbf{u},\mathbf{v})\in\mathbb{R}^{2n}$ and $\mathbf{w}=\mathbf{u}'+i\mathbf{v}'\in\mathbb{C}^n$ with $(\mathbf{u}',\mathbf{v}')\in\mathbb{R}^{2n}$.

The Lie algebra $\mathfrak{su}(2)$ of $SU(2)$ consists of 2×2 skew Hermitian matrices of trace zero. This Lie algebra is isomorphic to $\mathfrak{so}(3)$ and therefore to $(\mathbb{R}^3,\times)$ by the isomorphism given by

$$\mathbf{x}=(x^1,x^2,x^3)\in\mathbb{R}^3 \mapsto \widetilde{\mathbf{x}} := \frac{1}{2}\begin{bmatrix} -ix^3 & -ix^1-x^2 \\ -ix^1+x^2 & ix^3\end{bmatrix}\in\mathfrak{su}(2).$$

Thus we have $[\widetilde{\mathbf{x}},\widetilde{\mathbf{y}}]=(\mathbf{x}\times\mathbf{y})^{\sim}$ for any $\mathbf{x},\mathbf{y}\in\mathbb{R}^3$. Other useful relations are $\det(2\widetilde{\mathbf{x}})=\|\mathbf{x}\|^2$ and $\operatorname{trace}(\widetilde{\mathbf{x}}\widetilde{\mathbf{y}}) = -\frac{1}{2}\mathbf{x}\cdot\mathbf{y}$. Identify $\mathfrak{su}(2)^*$ with $\mathbb{R}^3$ by the map $\mu\in\mathfrak{su}(2)^*\mapsto\check{\mu}\in\mathbb{R}^3$ defined by

$$\check{\mu}\cdot\mathbf{x} := -2\langle\mu,\widetilde{\mathbf{x}}\rangle$$

for any $\mathbf{x}\in\mathbb{R}^3$. With these notations, the momentum map $\check{\mathbf{J}}\colon\mathbb{C}^2\to\mathbb{R}^3$ can be explicitly computed in coordinates: for any $\mathbf{x}\in\mathbb{R}^3$ we have

$$\begin{aligned}\check{\mathbf{J}}(z,w)\cdot\mathbf{x} &= -2\langle\mathbf{J}(z,w),\widetilde{\mathbf{x}}\rangle\\ &= \frac{1}{2}\operatorname{Im}\left(\begin{bmatrix} -ix^3 & -ix^1-x^2 \\ -ix^1+x^2 & ix^3\end{bmatrix}\begin{bmatrix} z\\ w\end{bmatrix}\cdot\begin{bmatrix} z\\ w\end{bmatrix}\right)\\ &= -\frac{1}{2}(2\operatorname{Re}(w\overline{z}), 2\operatorname{Im}(w\overline{z}), |z|^2-|w|^2)\cdot\mathbf{x}.\end{aligned}$$

Therefore

$$\check{\mathbf{J}}(z,w) = -\frac{1}{2}(2w\overline{z}, |z|^2-|w|^2)\in\mathbb{R}^3.$$

Thus, $\check{\mathbf{J}}$ is a Poisson map from $\mathbb{C}^2$, endowed with the canonical symplectic structure, to $\mathbb{R}^3$, endowed with the $+$ Lie–Poisson structure. Therefore, $-\check{\mathbf{J}}\colon\mathbb{C}^2\to\mathbb{R}^3$ is a canonical map, if $\mathbb{R}^3$ has the $-$ Lie–Poisson bracket relative to which the free rigid body equations are Hamiltonian. Pulling back the Hamiltonian $H(\boldsymbol{\Pi})=\boldsymbol{\Pi}\cdot\mathbb{I}^{-1}\boldsymbol{\Pi}/2$ to $\mathbb{C}^2$ gives a Hamiltonian function (called collective) on $\mathbb{C}^2$. The classical Hamilton equations for this function are therefore projected by $-\check{\mathbf{J}}$ to the rigid body equations $\dot{\boldsymbol{\Pi}}=\boldsymbol{\Pi}\times\mathbb{I}^{-1}\boldsymbol{\Pi}$. In this context, the variables (z,w) are called the *Cayley–Klein parameters*.

Exercise 18.31 Show that $-\check{\mathbf{J}}|_{S^3}\colon S^3 \to S^2$ is the *Hopf fibration*. In other words, *the momentum map of the $SU(2)$–action on $\mathbb{C}^2$, the Cayley–Klein parameters and the family of Hopf fibrations on concentric three-spheres in $\mathbb{C}^2$ are all the same map.*

Exercise 18.32 (Optical traveling wave pulses) The equation for the evolution of the complex amplitude of a polarized optical traveling wave pulse in a material medium is given as

$$\dot{z}_i = \frac{1}{\sqrt{-1}} \frac{\partial H}{\partial z_i^*}$$

with Hamiltonian $H\colon \mathbb{C}^2 \to \mathbb{R}$ defined by

$$H = z_i^* \chi_{ij}^{(1)} z_j + 3 z_i^* z_j^* \chi_{ijkl}^{(3)} z_k z_l$$

and the constant complex tensor coefficients $\chi_{ij}^{(1)}$ and $\chi_{ijkl}^{(1)}$ have the proper Hermitian and permutation symmetries for H to be real. Define the Stokes vectors by the isomorphism,

$$\mathbf{u} = (u^1, u^2, u^3) \in \mathbb{R}^3 \mapsto \widetilde{\mathbf{u}} := \frac{1}{2} \begin{bmatrix} -iu^3 & -iu^1 - U^2 \\ -iu^1 + u^2 & iu^3 \end{bmatrix} \in \mathfrak{su}(2).$$

(1) Prove that this isomorphism is an equivariant momentum map.

(2) Deduce the equations of motion for the Stokes vectors of this optical traveling wave and write it as a Lie–Poisson Hamiltonian system.

(3) Determine how this system is related to the equations for an $SO(3)$ rigid body.

Exercise 18.33 The formula determining the momentum map for the cotangent-lifted action of a Lie group G on a smooth manifold Q may be expressed in terms of the pairing $\langle\cdot,\cdot\rangle\colon \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$ as

$$\langle J, \xi \rangle = \langle p, \pounds_\xi q \rangle,$$

where $(q, p) \in T_q^* Q$ and $\pounds_\xi q$ is the infinitesimal generator of the action of the Lie algebra element ξ on the coordinate q.

Define appropriate pairings and determine the momentum maps explicitly for the following actions:

(a) $\pounds_\xi q = \xi \times q$ for $\mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$

(b) $\pounds_\xi q = \mathrm{ad}_\xi\, q$ for ad–action ad: $\mathfrak{g} \times \mathfrak{g} \mapsto \mathfrak{g}$ in a Lie algebra $\mathfrak{g}$

(c) AqA^{-1} for $A \in GL(3, R)$ acting on $q \in GL(3, R)$ by matrix conjugation

(d) Aq for left action of $A \in SO(3)$ on $q \in SO(3)$

(e) AqA^T for $A \in GL(3,R)$ acting on $q \in \mathrm{Sym}(3)$, that is $q = q^T$.

Answer 18.34

(a) $p \cdot \xi \times q = q \times p \cdot \xi \Rightarrow J = q \times p$. (The pairing is scalar product of vectors.)

(b) $\langle p, \mathrm{ad}_\xi\, q\rangle = -\langle \mathrm{ad}_q^*\, p, \xi\rangle \Rightarrow J = \mathrm{ad}_q^*\, p$ for the pairing $\langle\cdot,\cdot\rangle\colon \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$

(c) Compute $T_e(AqA^{-1}) = \xi q - q\xi = [\xi, q]$ for $\xi = A'(0) \in gl(3,R)$ acting on $q \in GL(3,R)$ by matrix Lie bracket $[\cdot,\cdot]$. For the matrix pairing $\langle A, B\rangle = \mathrm{trace}(A^T B)$, we have

$$\mathrm{trace}\left(p^T[\xi,q]\right) = \mathrm{trace}\left((pq^T - q^T p)^T \xi\right) \Rightarrow J = pq^T - q^T p.$$

(d) Compute $T_e(Aq) = \xi q$ for $\xi = A'(0) \in so(3)$ acting on $q \in SO(3)$ by left matrix multiplication. For the matrix pairing $\langle A, B\rangle = \mathrm{trace}(A^T B)$, we have

$$\mathrm{trace}\left(p^T \xi q\right) = \mathrm{trace}\left((pq^T)^T \xi\right) \Rightarrow J = \tfrac{1}{2}(pq^T - q^T p),$$

where we have used antisymmetry of the matrix $\xi \in so(3)$.

(e) Compute $T_e\big(AqA^T\big) = \xi q + q\xi^T$ for $\xi = A'(0) \in gl(3,R)$ acting on $q \in \mathrm{Sym}(3)$. For the matrix pairing $\langle A, B\rangle = \mathrm{trace}(A^T B)$, we have

$$\mathrm{trace}\left(p^T\left(\xi q + q\xi^T\right)\right) = \mathrm{trace}\left(q\left(p^T + p\right)\xi\right) = \mathrm{trace}\left((2qp)^T \xi\right) \Rightarrow J = 2qp,$$

where we have used 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 to have the correct number of components!)

Equivariance

Definition 18.35 A momentum map is *Ad^*–equivariant* iff

$$J(g \cdot x) = Ad^*_{g^{-1}} J(x)$$

for all $g \in G, x \in P$.

Proposition 18.36 *All cotangent-lifted actions are Ad^*–equivariant.*

Proposition 18.37 *Every Ad^*–equivariant momentum map $J\colon P \to \mathfrak{g}^*$ is a Poisson map, with respect to the '$+$' Lie–Poisson bracket on $\mathfrak{g}^*$.*

19 Quick summary for momentum maps

Let G be a Lie group, $\mathfrak{g}$ its Lie algebra, and let $\mathfrak{g}^*$ be its dual. Suppose that G acts symplectically on a symplectic manifold P with symplectic form denoted by Ω. Denote the infinitesimal generator associated with the Lie algebra element ξ by ξ_P and let the Hamiltonian vector field associated to a function $f\colon P \to \mathbb{R}$ be denoted X_f, so that $df = X_f \lrcorner \Omega$.

19.1 Definition, history and overview

A *momentum map* $J\colon P \to \mathfrak{g}^*$ is defined by the condition relating the infinitesimal generator ξ_P of a symmetry to the vector field of its corresponding conservation law, $\langle J, \xi\rangle$,

$$\xi_P = X_{\langle J,\xi\,\rangle}$$

for all $\xi \in \mathfrak{g}$. Here $\langle J, \xi\,\rangle\colon P \to \mathbb{R}$ is defined by the natural pointwise pairing.

A momentum map is said to be *equivariant* when it is equivariant with respect to the given action on P and the coadjoint action on $\mathfrak{g}^*$. That is,

$$J(g\cdot p) = \mathrm{Ad}^*_{g^{-1}}\, J(p)$$

for every $g \in G$, $p \in P$, where $g \cdot p$ denotes the action of g on the point p and where Ad denotes the adjoint action.

According to Weinstein [62], Lie [40] already knew that

(1) An action of a Lie group G with Lie algebra $\mathfrak{g}$ on a symplectic manifold P should be accompanied by such an equivariant momentum map $J\colon P \to \mathfrak{g}^*$ and

(2) The orbits of this action are themselves symplectic manifolds.

The links with mechanics were developed in the work of Lagrange, Poisson, Jacobi and, later, Noether. In particular, Noether showed that a momentum map for the action of a group G that is a symmetry of the Hamiltonian for a given system is a *conservation law* for that system.

In modern form, the momentum map and its equivariance were rediscovered by Kostant [37] and Souriau [61] in the general symplectic case, and by Smale [59; 60] for the case of the lifted action from a manifold Q to its cotangent bundle $P = T^*Q$. In this case, the equivariant momentum map is given explicitly by

$$\langle J(\alpha_q), \xi\,\rangle = \langle \alpha_q, \xi_Q(q)\rangle,$$

where $\alpha_q \in T^*Q$, $\xi \in \mathfrak{g}$, and where the angular brackets denote the natural pairing on the appropriate spaces. See Marsden and Ratiu [44] and Ortega and Ratiu [54] for additional history and description of the momentum map and its properties.

20 Rigid body equations on SO(n)

Recall from Manakov [41] and Ratiu [56] that the left invariant generalized rigid body equations on $\mathrm{SO}(n)$ may be written as

$$\begin{aligned} \dot{Q} &= Q\Omega, \\ \dot{M} &= M\Omega - \Omega M =: [M, \Omega], \end{aligned} \tag{RBn}$$

where $Q \in \mathrm{SO}(n)$ denotes the configuration space variable (the attitude of the body), $\Omega = Q^{-1}\dot{Q} \in so(n)$ is the body angular velocity, and

$$M := J(\Omega) = D^2\Omega + \Omega D^2 \in so^*(n),$$

is the body angular momentum. Here $J\colon so(n) \to so(n)^*$ is the symmetric (with respect to the above inner product) positive definite operator defined by

$$J(\Omega) = D^2\Omega + \Omega D^2,$$

where D^2 is the square of the constant diagonal matrix $D = \mathrm{diag}\,\{d_1, d_2, d_3\}$ satisfying $d_i^2 + d_j^2 > 0$ for all $i \neq j$. For $n = 3$ the elements of d_i^2 are related to the standard diagonal moment of inertia tensor I by

$$I = \mathrm{diag}\{I_1, I_2, I_3\}, \qquad I_1 = d_2^2 + d_3^2, \qquad I_2 = d_3^2 + d_1^2, \qquad I_3 = d_1^2 + d_2^2.$$

The Euler equations for the $SO(n)$ rigid body $\dot{M} = [M, \Omega]$ are readily checked to be the Euler–Lagrange equations on $so(n)$ for the Lagrangian

$$L(Q, \dot{Q}) = l(\Omega) = \tfrac{1}{2}\left\langle \Omega, J(\Omega) \right\rangle, \quad \text{with} \quad \Omega = Q^T\dot{Q}.$$

The momentum is found via the Legendre transformation to be

$$\frac{\partial l}{\partial \Omega} = J(\Omega) = M,$$

and the corresponding Hamiltonian is

$$H(M) = \frac{\partial l}{\partial \Omega} \cdot \Omega - l(\Omega) = \tfrac{1}{2}\big\langle M, J^{-1}(M) \big\rangle.$$

The quantity M is the angular momentum in the body frame. The corresponding angular momentum in space,

$$m = QMQ^T, \quad \text{is conserved} \quad \dot{m} = 0.$$

Indeed, conservation of spatial angular momentum m implies Euler's equations for the body angular momentum $M = Q^T m Q = \mathrm{Ad}^*_Q\, m$.

20.1 Implications of left invariance

This Hamiltonian $H(M)$ is invariant under the action of $SO(n)$ from the left. The corresponding conserved momentum map under this symmetry is known from the previous lecture as

$$J_L\colon T^*SO(n) \mapsto so(n)^* \quad \text{is} \quad J_L(Q,P) = PQ^T$$

On the other hand, we know (from Lectures 18 and 19) that the momentum map for right action is

$$J_R\colon T^*SO(n) \mapsto so(n)^*, \quad J_R(Q,P) = Q^T P$$

Hence $M = Q^T P = J_R$. Therefore, one computes

$$\begin{aligned} H(Q,P) &= H(Q, Q\cdot M) = H(\mathrm{Id}, M) \quad \text{(by left invariance)} \\ &= H(M) = \tfrac{1}{2}\langle M, J^{-1}(M)\rangle \\ &= \tfrac{1}{2}\langle Q^T P, J^{-1}(Q^T P)\rangle \end{aligned}$$

Hence, we may write the $SO(n)$ rigid body Hamiltonian as

$$H(Q,P) = \tfrac{1}{2}\big\langle Q^T P, \Omega(Q,P)\big\rangle$$

Consequently, the variational derivatives of $H(Q,P) = \frac{1}{2}\langle Q^T P, \Omega(Q,P)\rangle$ are

$$\begin{aligned} \delta H &= \Big\langle Q^T \delta P + \delta Q^T P, \Omega(Q,P)\Big\rangle \\ &= \mathrm{tr}(\delta P^T Q\Omega) + \mathrm{tr}(P^T \delta Q \Omega) \\ &= \mathrm{tr}(\delta P^T Q\Omega) + \mathrm{tr}(\delta Q \Omega P^T) \\ &= \mathrm{tr}(\delta P^T Q\Omega) + \mathrm{tr}(\delta Q^T P \Omega^T) \\ &= \langle \delta P, Q\Omega\rangle - \langle \delta Q, P\Omega\rangle \end{aligned}$$

where skew symmetry of Ω is used in the last step, that is, $\Omega^T = -\Omega$. Thus, Hamilton's canonical equations take the form,

$$\begin{aligned} \dot{Q} &= \frac{\delta H}{\delta P} = Q\Omega, \\ \dot{P} &= -\frac{\delta H}{\delta Q} = P\Omega. \end{aligned} \tag{65}$$

Equations (65) are the *symmetric generalized rigid body equations*, derived earlier by Bloch, Brockett and Crouch [7; 6] from the viewpoint of optimal control. Combining them yields,

$$Q^{-1}\dot{Q} = \Omega = P^{-1}\dot{P} \iff (PQ^T)^{\cdot} = 0,$$

in agreement with conservation of the momentum map $J_L(Q,P) = PQ^T$ corresponding to symmetry of the Hamiltonian under left action of $SO(n)$. This momentum map is the angular momentum in space, which is related to the angular momentum in the body by $PQ^T = m = QMQ^T$. Thus, we recognize the canonical momentum as $P = QM$ (see Example 18.22), and the momentum maps for left and right actions as

$$J_L = m = PQ^T \qquad \text{(spatial angular momentum)}$$
$$J_R = M = Q^T P \qquad \text{(body angular momentum)}$$

Thus, momentum maps $TG^* \mapsto \mathfrak{g}^*$ corresponding to symmetries of the Hamiltonian produce conservation laws; while momentum maps $TG^* \mapsto \mathfrak{g}^*$ which do *not* correspond to symmetries may be used to re-express the equations on $\mathfrak{g}^*$, in terms of variables on TG^*.

21 Manakov's formulation of the $SO(4)$ rigid body

The Euler equations on $SO(4)$ are

$$\text{(RBn)} \qquad \frac{dM}{dt} = M\Omega - \Omega M = [M, \Omega],$$

where Ω and M are skew symmetric 4×4 matrices. The angular frequency Ω is a linear function of the angular momentum, M. Manakov [41] "deformed" these equations into

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

where A, B are also skew symmetric 4×4 matrices and λ is a scalar constant parameter. For these equations to hold for any value of λ, the coefficent of each power must vanish.

- The coefficent of λ^2 is

  $$0 = [A, B]$$

  So A and B must commute. So, let them be constant and diagonal:

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

- The coefficent of λ is

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

  Therefore, by antisymmetry of M and Ω,

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

- Finally, the coefficent of λ^0 is the Euler equation,

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

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

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

which turns out to possess only 5 free parameters.

With these conditions, Manakov's deformation of the $SO(4)$ rigid body implies for every power n that

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

Since the commutator is antisymmetric, its trace vanishes and one has

$$\frac{d}{dt}\operatorname{trace}(M + \lambda A)^n = 0$$

after commuting the trace operation with time derivative. Consequently,

$$\operatorname{trace}(M + \lambda A)^n = \text{ constant}$$

for each power of λ. That is, all the coefficients of each power of λ are constant in time for the $SO(4)$ rigid body. Manakov [41] proved that these constants of motion are sufficient to completely determine the solution.

Remark 21.1 This result generalizes considerably. First, it holds for $SO(n)$. Indeed, as as proven using the theory of algebraic varieties by Haine [22], Manakov's method captures all the algebraically integrable rigid bodies on $SO(n)$ and 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.) Moreover, Miščenko and Fomenko [51] prove that every compact Lie group admits a family of left-invariant metrics with completely integrable geodesic flows.

Exercise 21.2 Try computing the constants of motion $\operatorname{trace}(M + \lambda A)^n$ for the values $n = 2, 3, 4$. How many additional constants of motion are needed for integrability for these cases? How many for general n? 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 21.3 The traces of the powers $\mathrm{trace}(M+\lambda A)^n$ are given by

$$n{=}2: \mathrm{tr}\, M^2 + 2\lambda\, \mathrm{tr}(AM) + \lambda^2\, \mathrm{tr}\, A^2$$
$$n{=}3: \mathrm{tr}\, M^3 + 3\lambda\, \mathrm{tr}(AM^2) + 3\lambda^2\, \mathrm{tr}\, A^2M + \lambda^3\, \mathrm{tr}\, A^3$$
$$n{=}4: \mathrm{tr}\, M^4 + 4\lambda\, \mathrm{tr}(AM^3) + \lambda^2(2\, \mathrm{tr}\, A^2M^2 + 4\, \mathrm{tr}\, AMAM) + \lambda^3\, \mathrm{tr}\, A^3M + \lambda^4\, \mathrm{tr}\, A^4$$

The number of conserved quantities for $n=2,3,4$ are, respectively, one ($C_1 = \mathrm{tr}\, M^2$), one ($I_1 = \mathrm{tr}\, AM^2$) and two ($C_2 = \mathrm{tr}\, M^4$ and $I_2 = 2\, \mathrm{tr}\, A^2M^2 + 4\, \mathrm{tr}\, AMAM$). The quantities C_1 and C_2 are Casimirs for the Lie–Poisson bracket for the rigid body. Thus, $\{C_1, H\} = 0 = \{C_2, H\}$ for any Hamiltonian $H(M)$; so of course C_1 and C_2 are conserved. However, each Casimir only reduces the dimension of the system by one. The dimension of the original phase space is $\dim T^*SO(n) = n(n-1)$. This is reduced in half by left invariance of the Hamiltonian to the dimension of the dual Lie algebra $\dim so(n)^* = n(n-1)/2$. For $n=4$, $\dim so(4)^* = 6$. One then subtracts the number of Casimirs (two) by passing to their level surfaces, which leaves four dimensions remaining in this case. The other two constants of motion I_1 and I_2 turn out to be sufficient for integrability, because they are in involution $\{I_1, I_2\} = 0$ and because the level surfaces of the Casimirs are symplectic manifolds, by the Marsden–Weinstein reduction theorem [48]. For more details, see Ratiu [56].

Exercise 21.4 How do the Euler equations look on $so(4)^*$ as a matrix equation? Is there an analog of the hat map for $so(3)^*$? Hint: the Lie algebra $so(4)$ is locally isomorphic to $so(3) \times so(3)$.

Exercise 21.5 Write Manakov's deformation of the rigid body equations in the symmetric form (65).

22 Free ellipsoidal motion on GL(n)

Riemann [58] considered the deformation of a body in $\mathbb{R}^n$ given by

$$x(t, x_0) = Q(t)\, x_0, \tag{66}$$

with $x, x_0 \in \mathbb{R}^n$, $Q(t) \in GL_+(n, \mathbb{R})$ and $x(t_0, x_0) = x_0$, so that $Q(t_0) = \mathrm{Id}$. (The subscript $+$ in $GL_+(n,\mathbb{R})$ means $n \times n$ matrices with positive determinant.) Thus, $x(t, x_0)$ is the current (Eulerian) position at time t of a material parcel that was at (Lagrangian) position x_0 at time t_0. The "deformation gradient," that is, the Jacobian matrix $Q = \partial x/\partial x_0$ of this "Lagrange-to-Euler map," is a function of only time, t,

$$\partial x/\partial x_0 = Q(t), \quad \text{with} \quad \det Q > 0.$$

The velocity of such a motion is given by

$$\dot{x}(t, x_0) = \dot{Q}(t)\, x_0 = \dot{Q}(t) Q^{-1}(t)\, x = u(t, x). \tag{67}$$

The kinetic energy for such a body occupying a reference volume $\mathcal{B}$ defines the quadratic form,

$$L = \tfrac{1}{2} \int_{\mathcal{B}} \rho(x_0) |\dot{x}(t, x_0)|^2 \, d^3 x_0 = \tfrac{1}{2} \operatorname{tr} \Big(\dot{Q}(t)^T I \, \dot{Q}(t) \Big) = \tfrac{1}{2} \dot{Q}^i_A \, I^{AB} \dot{Q}^i_B .$$

Here I is the constant symmetric tensor,

$$I^{AB} = \int_{\mathcal{B}} \rho(x_0) x_0^A x_0^B \, d^3 x_0,$$

which we will take as being proportional to the identity $I^{AB} = c_0^2 \delta^{AB}$ for the remainder of these considerations. This corresponds to taking an initially spherical reference configuration for the fluid. Hence, we are dealing with the Lagrangian consisting only of kinetic energy,[5]

$$L = \tfrac{1}{2} \operatorname{tr} \big(\dot{Q}(t)^T \dot{Q}(t) \big).$$

The Euler–Lagrange equations for this Lagrangian simply represent *free motion* on the group $GL_+(n, \mathbb{R})$,

$$\ddot{Q}(t) = 0,$$

which is immediately integrable as

$$Q(t) = Q(0) + \dot{Q}(0) t,$$

where $Q(0)$ and $\dot{Q}(0)$ are the values at the initial time $t = 0$. Legendre transforming this Lagrangian for free motion yields

$$P = \frac{\partial L}{\partial \dot{Q}^T} = \dot{Q}.$$

The corresponding Hamiltonian is expressed as

$$H(Q, P) = \tfrac{1}{2} \operatorname{tr} \big(P^T P \big) = \tfrac{1}{2} \| P \|^2 .$$

The canonical equations for this Hamiltonian are simply

$$\dot{Q} = P, \quad \text{with} \quad \dot{P} = 0.$$

[5] Riemann [58] considered the much more difficult problem of a *self-gravitating* ellipsoid deforming according to (66) in $\mathbb{R}^3$. See Chandrasekhar [11] for the history of this problem.

22.1 Polar decomposition of free motion on $GL_+(n,\mathbb{R})$

The deformation tensor $Q(t) \in GL_+(n,\mathbb{R})$ for such a body may be decomposed as

$$Q(t) = R^{-1}(t)D(t)S(t). \tag{68}$$

This is the polar decomposition of a matrix in $GL_+(n,\mathbb{R})$. The interpretations of the various components of the motion can be seen from equation (66). Namely,

- $R \in SO(n)$ rotates the x–coordinates,
- $S \in SO(n)$ rotates the x_0–coordinates in the reference configuration[6] and
- D is a diagonal matrix which represents stretching deformations along the principal axes of the body.

The two $SO(n)$ rotations lead to their corresponding angular frequencies, defined by

$$\Omega = \dot{R}R^{-1}, \quad \Lambda = \dot{S}S^{-1}. \tag{69}$$

Rigid body motion will result, when S restricts to the identity matrix and D is a constant diagonal matrix.

Remark 22.1 The combined motion of a set of fluid parcels governed by (66) along the curve $Q(t) \in GL_+(n,\mathbb{R})$ is called "ellipsoidal," because it can be envisioned in three dimensions as a fluid ellipsoid whose orientation in space is governed by $R \in SO(n)$, whose shape is determined by D consisting of its instantaneous principle axes lengths and whose internal circulation of material is described by $S \in SO(n)$. In addition, fluid parcels initially arranged along a straight line within the ellipse will remain on a straight line.

22.2 Euler–Poincaré dynamics of free Riemann ellipsoids

In Hamilton's principle, $\delta\int L\,dt = 0$, we chose a Lagrangian $L\colon TGL_+(n,\mathbb{R}) \to \mathbb{R}$ in the form

$$L(Q,\dot{Q}) = T(\Omega,\Lambda,D,\dot{D}), \tag{70}$$

in which the kinetic energy T is given by using the polar decomposition $Q(t) = R^{-1}(t)D(t)S(t)$ in (68), as follows.

$$\dot{Q} = R^{-1}(-\Omega D + \dot{D} + D\Lambda)S. \tag{71}$$

[6] This is the "particle relabeling map" for this class of motions.

Consequently, the kinetic energy for ellipsoidal motion becomes

$$
\begin{aligned}
T &= \tfrac{1}{2}\,\mathrm{trace}\big[-\Omega D^2\Omega - \Omega D\dot{D} + \Omega D\Lambda D + \dot{D}D\Omega + \dot{D}^2 \\
&\qquad\qquad - D\Lambda^2 D - \dot{D}\Lambda D + D\Lambda D\Omega + D\Lambda\dot{D}\big] \\
&= \tfrac{1}{2}\,\mathrm{trace}\big[-\Omega^2 D^2 - \Lambda^2 D^2 + \underbrace{2\Omega D\Lambda D}_{\text{Coriolis coupling}} + \dot{D}^2\big].
\end{aligned}
\tag{72}
$$

Remark 22.2 Note the discrete exchange symmetry of the kinetic energy: T is invariant under $\Omega \leftrightarrow \Lambda$.[7]

For $\Lambda = 0$ and D constant expression (72) for T reduces to the usual kinetic energy for the rigid-body,

$$
T\Big|_{\Lambda=0,\,D=\mathrm{const}} = -\tfrac{1}{4}\,\mathrm{trace}\big[\Omega(D\Omega + \Omega D)\big]. \tag{73}
$$

This Lagrangian (70) is invariant under the right action, $R \to Rg$ and $S \to Sg$, for $g \in SO(n)$. In taking variations we shall use the formulas[8]

$$
\delta\Omega = \dot{\Sigma} + [\Sigma, \Omega] \equiv \dot{\Sigma} - \mathrm{ad}_\Omega\,\Sigma, \qquad \Sigma \equiv \delta R\,R^{-1}, \tag{74}
$$

$$
\delta\Lambda = \dot{\Xi} + [\Xi, \Lambda] \equiv \dot{\Xi} - \mathrm{ad}_\Lambda\,\Xi, \qquad \Xi \equiv \delta S\,S^{-1}, \tag{75}
$$

in which the ad-operation is defined in terms of the Lie-algebra (matrix) commutator $[\cdot,\cdot]$ as, for example, $\mathrm{ad}_\Omega\,\Sigma \equiv [\Omega, \Sigma]$. Substituting these formulas into Hamilton's principle gives

$$
\begin{aligned}
0 = \delta\int L\,dt &= \int dt\,\frac{\partial L}{\partial\Omega}\cdot\delta\Omega + \frac{\partial L}{\partial\Lambda}\cdot\delta\Lambda + \frac{\partial L}{\partial D}\delta D + \frac{\partial L}{\partial\dot{D}}\delta\dot{D}, \\
&= \int dt\,\frac{\partial L}{\partial\Omega}\cdot[\dot{\Sigma} - \mathrm{ad}_\Omega\,\Sigma] + \frac{\partial L}{\partial\Lambda}\cdot[\dot{\Xi} - \mathrm{ad}_\Lambda\,\Xi] + \left[\frac{\partial L}{\partial D} - \frac{d}{dt}\frac{\partial L}{\partial\dot{D}}\right]\delta D, \\
&= -\int dt\left[\frac{d}{dt}\frac{\partial L}{\partial\Omega} - \mathrm{ad}^*_\Omega\,\frac{\delta L}{\delta\Omega}\right]\cdot\Sigma + \left[\frac{d}{dt}\frac{\partial L}{\partial\Lambda} - \mathrm{ad}^*_\Lambda\,\frac{\partial L}{\partial\Lambda}\right]\cdot\Xi + \left[\frac{d}{dt}\frac{\partial L}{\partial\dot{D}} - \frac{\partial L}{\partial D}\right]\delta D,
\end{aligned}
\tag{76}
$$

where the operation ad^*_Ω, for example, is defined by

$$
\mathrm{ad}^*_\Omega\,\frac{\partial L}{\partial\Omega}\cdot\Sigma = -\frac{\partial L}{\partial\Omega}\cdot\mathrm{ad}_\Omega\,\Sigma = -\frac{\partial L}{\partial\Omega}\cdot[\Omega, \Sigma], \tag{77}
$$

[7] According to Chandrasekhar [11] this discrete symmetry was first noticed by Riemann's friend, Dedekind [13].

[8] These variational formulas are obtained directly from the definitions of Ω and Λ.

and the dot '$\cdot$' denotes pairing between the Lie algebra and its dual. This could also have been written in the notation using $\langle\cdot,\cdot\rangle\colon \mathfrak{g}^*\times\mathfrak{g}\to\mathbb{R}$, as

$$\left\langle \mathrm{ad}^*_\Omega \frac{\partial L}{\partial\Omega}, \Sigma\right\rangle = -\left\langle \frac{\partial L}{\partial\Omega}, \mathrm{ad}_\Omega\,\Sigma\right\rangle = -\left\langle \frac{\partial L}{\partial\Omega}, [\Omega,\Sigma]\right\rangle. \tag{78}$$

The Euler–Poincaré dynamics is given by the stationarity conditions for Hamilton's principle,

$$\Sigma:\ \frac{d}{dt}\frac{\partial L}{\partial\Omega} - \mathrm{ad}^*_\Omega \frac{\partial L}{\partial\Omega} = 0, \tag{79}$$

$$\Xi:\ \frac{d}{dt}\frac{\partial L}{\partial\Lambda} - \mathrm{ad}^*_\Lambda \frac{\partial L}{\partial\Lambda} = 0, \tag{80}$$

$$\delta D:\ \frac{d}{dt}\frac{\partial L}{\partial\dot D} - \frac{\partial L}{\partial D} = 0. \tag{81}$$

These are the *Euler–Poincaré equations* for the ellipsoidal motions generated by Lagrangians of the form given in equation (70). For example, such Lagrangians determine the dynamics of the Riemann ellipsoids – circulating, rotating, self-gravitating fluid flows at constant density within an ellipsoidal boundary.

22.3 Left and right momentum maps: angular momentum versus circulation

The Euler–Poincaré equations ((79)–(81)) involve *angular momenta* defined in terms of the angular velocities Ω, Λ and the shape D by

$$M = \frac{\partial T}{\partial\Omega} = -\Omega D^2 - D^2\Omega + 2D\Lambda D, \tag{82}$$

$$N = \frac{\partial T}{\partial\Lambda} = -\Lambda D^2 - D^2\Lambda + 2D\Omega D. \tag{83}$$

These angular momenta are related to the original deformation gradient $Q = R^{-1}DS$ in equation (66) by the two momentum maps from Example 18.25

$$PQ^T - QP^T = \dot Q Q^T - Q\dot Q^T = R^{-1}MR, \tag{84}$$

$$P^TQ - Q^TP = \dot Q^T Q - Q^T\dot Q = S^{-1}NS. \tag{85}$$

To see that N is related to the *vorticity*, we consider the exterior derivative of the circulation one-form $\mathbf{u}\cdot d\mathbf{x}$ defined as

$$\begin{aligned} d(\mathbf{u}\cdot d\mathbf{x}) &= \mathrm{curl}\,\mathbf{u}\cdot d\mathbf{S} \\ &= \tfrac12\big(\dot Q^T Q - Q^T\dot Q\big)_{jk}\, dx_0^j\wedge dx_0^k = (S^{-1}NS)_{jk}dx_0^j\wedge dx_0^k. \end{aligned} \tag{86}$$

Thus, $S^{-1}NS$ is the fluid vorticity, referred to the Lagrangian coordinate frame. For Euler's fluid equations, Kelvin's circulation theorem implies $(S^{-1}NS)^{\cdot} = 0$.

Likewise, M is related to the *angular momentum* by considering

(87) $$u_i x_j - u_j x_i = \dot{Q}_{ik} x_0^k x_0^l Q_{lj}^T - Q_{ik} x_0^k x_0^l \dot{Q}_{lj}^T.$$

For spherical symmetry, we may choose $x_0^k x_0^l = \delta^{kl}$ and, in this case, the previous expression becomes

(88) $$u_i x_j - u_j x_i = \left[\dot{Q}Q^T - Q\dot{Q}^T\right]_{ij} = [R^{-1}MR]_{ij}.$$

Thus, $R^{-1}MR$ is the angular momentum of the motion, referred to the Lagrangian coordinate frame for spherical symmetry. In this case, the angular momentum is conserved, so that $(R^{-1}MR)^{\cdot} = 0$.

In terms of these angular momenta, the Euler–Poincaré–Lagrange equations ((79)–(81)) are expressed as

(89) $$\dot{M} = [\Omega, M],$$

(90) $$\dot{N} = [\Lambda, N],$$

(91) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{D}}\right) = \frac{\partial L}{\partial D}.$$

Perhaps not unexpectedly, because of the combined symmetries of the kinetic-energy Lagrangian (70) under both left and right actions of $SO(n)$, the first two equations are consistent with the conservation laws,

$$(R^{-1}MR)^{\cdot} = 0 \quad \text{and} \quad (S^{-1}NS)^{\cdot} = 0,$$

respectively. Thus, equation (89) is the angular momentum equation while (90) is the vorticity equation. (Fluids have both types of circulatory motions.) The remaining equation (91) for the diagonal matrix D determines the shape of the ellipsoid undergoing free motion on $GL(n,\mathbb{R})$.

22.4 Vector representation of free Riemann ellipsoids in 3D

In three dimensions these expressions may be written in vector form by using the *hat map*, written now using upper and lower case Greek letters as

$$\Omega_{ij} = \epsilon_{ijk}\omega_k, \quad \Lambda_{ij} = \epsilon_{ijk}\lambda_k,$$

with $\epsilon_{123} = 1$, and $D = \text{diag}\,\{d_1, d_2, d_3\}$.

Exercise 22.3 What is the analog of the hat map in four dimensions? Hint: locally the Lie algebra $so(4)$ is isomorphic to $so(3)\times so(3)$.

Hence, the angular-motion terms in the kinetic energy may be rewritten as

$$-\tfrac{1}{2}\,\mathrm{trace}\,\big(\Omega^2 D^2\big) = \tfrac{1}{2}\big[\big(d_1^2+d_2^2\big)\omega_3^2 + \big(d_2^2+d_3^2\big)\omega_1^2 + \big(d_3^2+d_1^2\big)\omega_2^2\big], \tag{92}$$

$$-\tfrac{1}{2}\,\mathrm{trace}\,\big(\Lambda^2 D^2\big) = \tfrac{1}{2}\big[\big(d_1^2+d_2^2\big)\lambda_3^2 + \big(d_2^2+d_3^2\big)\lambda_1^2 + \big(d_3^2+d_1^2\big)\lambda_2^2\big], \tag{93}$$

$$-\tfrac{1}{2}\,\mathrm{trace}(\Omega D\Lambda D) = \big[d_1 d_2(\omega_3\lambda_3) + d_2 d_3(\omega_1\lambda_1) + d_3 d_1(\omega_2\lambda_2)\big]. \tag{94}$$

On comparing equations (73) and (92) for the kinetic energy of the rigid body part of the motion, we identify the usual moments of inertia as

$$I_k = d_i^2 + d_j^2, \quad \text{with} \quad i,j,k \text{ cyclic}.$$

The antisymmetric matrices M and N have vector representations in 3D given by

$$M_k = \frac{\partial T}{\partial \omega_k} = (d_i^2 + d_j^2)\omega_k - 2 d_i d_j \lambda_k, \tag{95}$$

$$N_k = \frac{\partial T}{\partial \lambda_k} = (d_i^2 + d_j^2)\lambda_k - 2 d_i d_j \omega_k, \tag{96}$$

again with i,j,k cyclic permutations of $\{1,2,3\}$.

Vector representation in 3D In terms of their 3D vector representations of the angular momenta in equations (95) and (96), the two equations (89) and (90) become

$$\dot{\mathbf{M}} = (\dot R R^{-1})\mathbf{M} = \Omega\mathbf{M} = \boldsymbol{\omega}\times\mathbf{M}, \quad \dot{\mathbf{N}} = (\dot S S^{-1})\mathbf{N} = \Lambda\mathbf{N} = \boldsymbol{\lambda}\times\mathbf{N}. \tag{97}$$

Relative to the Lagrangian fluid frame of reference, these equations become

$$(R^{-1}\mathbf{M})^{\boldsymbol\cdot} = R^{-1}(\dot{\mathbf{M}} - \boldsymbol{\omega}\times\mathbf{M}) = 0, \tag{98}$$

$$(S^{-1}\mathbf{N})^{\boldsymbol\cdot} = S^{-1}(\dot{\mathbf{N}} - \boldsymbol{\lambda}\times\mathbf{N}) = 0. \tag{99}$$

So each of these degrees of freedom represents a rotating, deforming body, whose ellipsoidal shape is governed by the Euler–Lagrange equations (91) for the lengths of its three principal axes.

Exercise 22.4 (Elliptical motions with potential energy on $GL(2,\mathbb{R})$) Compute equations (89)–(91) for elliptical motion in the plane. Find what potentials $V(D)$ are solvable for $L = T(\Omega,\Lambda,D,\dot D) - V(D)$ by reducing these equations to the separated Newtonian forms,

$$\frac{d^2 r^2}{dt^2} = -\frac{dV(r)}{dr^2}, \qquad \frac{d^2\alpha}{dt^2} = -\frac{dW(\alpha)}{d\alpha},$$

for $r^2 = d_1^2 + d_2^2$ and $\alpha = \tan^{-1}(d_2/d_1)$ with $d_1(t)$ and $d_2(t)$ in two dimensions. Hint: consider the potential energy,

$$V(D) = V\big(\operatorname{tr} D^2, \det(D)\big),$$

for which the equations become homogeneous in $r^2(t)$.

Exercise 22.5 (Ellipsoidal motions with potential energy on $GL(3,\mathbb{R})$) Choose the Lagrangian in 3D,

$$L = \tfrac{1}{2}\operatorname{tr}\big(\dot{Q}^T\dot{Q}\big) - V\big(\operatorname{tr}\big(Q^TQ\big), \det(Q)\big),$$

where $Q(t) \in GL(3,\mathbb{R})$ is a 3×3 matrix function of time and the potential energy V is an arbitrary function of $\operatorname{tr}(Q^TQ)$ and $\det(Q)$.

(1) Legendre transform this Lagrangian. That is, find the momenta P_{ij} canonically conjugate to Q_{ij}, construct the Hamiltonian $H(Q,P)$ and write Hamilton's canonical equations of motion for this problem.

(2) Show that the Hamiltonian is invariant under $Q \to OQ$ where $O \in SO(3)$. Construct the cotangent lift of this action on P. Hence, construct the momentum map of this action.

(3) Construct another distinct action of $SO(3)$ on this system which also leaves its Hamiltonian $H(Q,P)$ invariant. Construct its momentum map. Do the two momentum maps Poisson commute? Why?

(4) How are these two momentum maps related to the angular momentum and circulation in equations (82) and (83)?

(5) How does the 2D restriction of this problem inform the previous one?

Exercise 22.6 ($GL(n,\mathbb{R})$–invariant motions) Begin with the Lagrangian

$$L = \tfrac{1}{2}\operatorname{tr}\big(\dot{S}S^{-1}\dot{S}S^{-1}\big) + \tfrac{1}{2}\dot{\mathbf{q}}^T S^{-1}\dot{\mathbf{q}}$$

where S is an $n\times n$ symmetric matrix and $\mathbf{q} \in \mathbb{R}^n$ is an n–component column vector.

(1) Legendre transform to construct the corresponding Hamiltonian and canonical equations.

(2) Show that the system is invariant under the group action

$$\mathbf{q} \to A\mathbf{q} \quad \text{and} \quad S \to ASA^T$$

for any constant invertible $n\times n$ matrix, A.

(3) Compute the infinitesimal generator for this group action and construct its corresponding momentum map. Is this momentum map equivariant?

(4) Verify directly that this momentum map is a conserved $n \times n$ matrix quantity by using the equations of motion.

(5) Is this system completely integrable for any value of $n > 2$?

23 Heavy top equations

23.1 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,

$$\dot{\mathbf{\Gamma}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}(t)\mathbf{\Gamma} = \mathbf{\Gamma} \times \mathbf{\Omega}.$$

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, $\mathbf{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\widehat{\ }$, where $\widehat{\ }\colon (so(3), [\cdot,\cdot]) \to (\mathbb{R}^3, \times)$ 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,

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega} + mg\,\mathbf{\Gamma} \times \boldsymbol{\chi}, \tag{100}$$

$$\dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}, \tag{101}$$

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,
- $\mathbb{I} = \operatorname{diag}(I_1, I_2, I_3)$ is the moment of inertia tensor, diagonalized in the body principle axes,
- $\mathbf{\Gamma} = R^{-1}(t)\hat{\mathbf{z}}$ 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 center of mass,
- m is the total mass of the body and g is the constant acceleration of gravity.

23.2 Heavy top action principle

Proposition 23.1 *The heavy top equations are equivalent to the* heavy top action principle *for a* reduced action

$$\delta S_{\text{red}} = 0, \quad \textit{with} \quad S_{\text{red}} = \int_a^b l(\boldsymbol{\Omega}, \boldsymbol{\Gamma})\, dt = \int_a^b \tfrac{1}{2}\langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle - \langle mg\boldsymbol{\chi}, \boldsymbol{\Gamma}\rangle\, dt, \tag{102}$$

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

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega}\times\boldsymbol{\Sigma} \quad \textit{and} \quad \delta\boldsymbol{\Gamma} = \boldsymbol{\Gamma}\times\boldsymbol{\Sigma}, \tag{103}$$

arising from variations of the definitions $\boldsymbol{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\hat{}$ *and* $\boldsymbol{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$ *in which* $\boldsymbol{\Sigma}(t) = \big(\mathbf{R}^{-1}\delta\mathbf{R}\big)\hat{}$ *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, we obtain the variation,

$$\begin{aligned}
\delta\int_a^b l(\boldsymbol{\Omega}, \boldsymbol{\Gamma})\, dt &= \int_a^b \langle \mathbb{I}\boldsymbol{\Omega}, \delta\boldsymbol{\Omega}\rangle - \langle mg\,\boldsymbol{\chi}, \delta\boldsymbol{\Gamma}\rangle\, dt \\
&= \int_a^b \langle \mathbb{I}\boldsymbol{\Omega}, \dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega}\times\boldsymbol{\Sigma}\rangle - \langle mg\,\boldsymbol{\chi}, \boldsymbol{\Gamma}\times\boldsymbol{\Sigma}\rangle\, dt \\
&= \int_a^b \Big\langle -\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Sigma}\Big\rangle + \langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega}\times\boldsymbol{\Sigma}\rangle - \langle mg\,\boldsymbol{\chi}, \boldsymbol{\Gamma}\times\boldsymbol{\Sigma}\rangle dt \\
&= \int_a^b \Big\langle -\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega} + \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega} + mg\,\boldsymbol{\Gamma}\times\boldsymbol{\chi}, \boldsymbol{\Sigma}\Big\rangle dt,
\end{aligned}$$

upon integrating by parts and using the endpoint conditions, $\boldsymbol{\Sigma}(b) = \boldsymbol{\Sigma}(a) = 0$. Since $\boldsymbol{\Sigma}$ is otherwise arbitrary, (102) is equivalent to

$$-\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega} + \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega} + mg\,\boldsymbol{\Gamma}\times\boldsymbol{\chi} = 0,$$

which is Euler's motion equation for the heavy top (100). This motion equation is completed by the auxiliary equation $\dot{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma}\times\boldsymbol{\Omega}$ in (101) arising from the definition of $\boldsymbol{\Gamma}$. □

The Legendre transformation for $l(\boldsymbol{\Omega}, \boldsymbol{\Gamma})$ gives the body angular momentum

$$\boldsymbol{\Pi} = \frac{\partial l}{\partial \boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega}.$$

The well known energy Hamiltonian for the heavy top then emerges as

$$h(\boldsymbol{\Pi}, \boldsymbol{\Gamma}) = \boldsymbol{\Pi}\cdot\boldsymbol{\Omega} - l(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) = \tfrac{1}{2}\langle \boldsymbol{\Pi}, \mathbb{I}^{-1}\boldsymbol{\Pi}\rangle + \langle mg\,\boldsymbol{\chi}, \boldsymbol{\Gamma}\rangle, \tag{104}$$

which is the sum of the kinetic and potential energies of the top.

The Lie–Poisson equations Let $f, h\colon \mathfrak{g}^* \to \mathbb{R}$ be real-valued functions on the dual space $\mathfrak{g}^*$. Denoting elements of $\mathfrak{g}^*$ by μ, the functional derivative of f at μ is defined as the unique element $\delta f/\delta\mu$ of $\mathfrak{g}$ defined by

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[f(\mu+\varepsilon\delta\mu)-f(\mu)] = \left\langle \delta\mu, \frac{\delta f}{\delta\mu}\right\rangle, \tag{105}$$

for all $\delta\mu \in \mathfrak{g}^*$, where $\langle\cdot,\cdot\rangle$ denotes the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$.

Definition 23.2 (Lie–Poisson brackets and Lie–Poisson equations) The $(\pm)$ *Lie–Poisson brackets* are defined by

$$\{f,h\}_\pm(\mu) = \pm\left\langle \mu, \left[\frac{\delta f}{\delta\mu}, \frac{\delta h}{\delta\mu}\right]\right\rangle = \mp\left\langle \mu, \mathrm{ad}_{\delta h/\delta\mu}\frac{\delta f}{\delta\mu}\right\rangle. \tag{106}$$

The corresponding *Lie–Poisson equations*, determined by $\dot f = \{f,h\}$ read

$$\dot\mu = \{\mu, h\} = \mp\,\mathrm{ad}^*_{\delta h/\delta\mu}\,\mu, \tag{107}$$

where one defines the ad^* operation in terms of the pairing $\langle\cdot,\cdot\rangle$, by

$$\{f,h\} = \left\langle \mu, \mathrm{ad}_{\delta h/\delta\mu}\frac{\delta f}{\delta\mu}\right\rangle = \left\langle \mathrm{ad}^*_{\delta h/\delta\mu}\,\mu, \frac{\delta f}{\delta\mu}\right\rangle.$$

The Lie–Poisson setting of mechanics is a special case of the general theory of systems on Poisson manifolds, for which there is now an extensive theoretical development. (See Marsden and Ratiu [2003] for a start on this literature.)

23.3 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 center 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 $\mathbf{\Pi}$'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. The map from T^*G to $\mathfrak{g}^*$ being a Poisson map *is a general fact about momentum maps*. The Hamiltonian point of view of all this is a standard subject.

Remark 23.3 (Lie–Poisson description of the heavy top) As it turns out, the underlying Lie algebra for the Lie–Poisson description of 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 actions of the Euclidean group of rotations and translations on the body, since 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. As we shall see later, a close parallel exists between this case and the Lie–Poisson structure for compressible fluids.

23.4 The heavy top Lie–Poisson brackets

The *Lie algebra of the special Euclidean group in 3D* is $se(3)=\mathbb{R}^3\times\mathbb{R}^3$ with the Lie bracket

$$[(\boldsymbol{\xi},\mathbf{u}),(\boldsymbol{\eta},\mathbf{v})]=(\boldsymbol{\xi}\times\boldsymbol{\eta},\boldsymbol{\xi}\times\mathbf{v}-\boldsymbol{\eta}\times\mathbf{u}). \tag{108}$$

We identify the dual space with pairs $(\mathbf{\Pi},\mathbf{\Gamma})$; the corresponding $(-)$ Lie–Poisson bracket called the *heavy top bracket* is

$$\{f,h\}(\mathbf{\Pi},\mathbf{\Gamma})=-\mathbf{\Pi}\cdot\nabla_{\Pi}f\times\nabla_{\Pi}h-\mathbf{\Gamma}\cdot\big(\nabla_{\Pi}f\times\nabla_{\Gamma}h-\nabla_{\Pi}h\times\nabla_{\Gamma}f\big). \tag{109}$$

This Lie–Poisson bracket and the Hamiltonian (104) recover the equations (100) and (101) for the heavy top, as

$$\begin{aligned}\dot{\mathbf{\Pi}}&=\{\mathbf{\Pi},h\}=\mathbf{\Pi}\times\nabla_{\Pi}h+\mathbf{\Gamma}\times\nabla_{\Gamma}h=\mathbf{\Pi}\times\mathbb{I}^{-1}\mathbf{\Pi}+\mathbf{\Gamma}\times mg\,\boldsymbol{\chi},\\ \dot{\mathbf{\Gamma}}&=\{\mathbf{\Gamma},h\}=\mathbf{\Gamma}\times\nabla_{\Pi}h=\mathbf{\Gamma}\times\mathbb{I}^{-1}\mathbf{\Pi}.\end{aligned}$$

Remark 23.4 (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 $so(3)\circledS\mathbb{R}^3$ 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 $\mathfrak{g} \circledS V$ of a Lie algebra $\mathfrak{g}$ on a vector space V is given by

$$\big[(X,a),(\bar{X},\overline{a})\big] = \big([X,\bar{X}\,],\bar{X}(a) - X(\overline{a})\big)$$

in which $X,\bar{X} \in \mathfrak{g}$ and $a,\overline{a} \in V$. Here, the action of the Lie algebra on the vector space is denoted, for example, $X(\overline{a})$. Usually, this action would be the Lie derivative.

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, for example, 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 isotropy of $\mathbf{R}^3$ from $SO(3)$ to $SO(2)$. The general theory for semidirect products is reviewed in a variety of places, including Marsden, Ratiu and Weinstein [47; 46]. 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 magnetized fluids, to multiphase fluids, to super fluids, to Yang–Mills fluids, relativistic, or not, and to liquid crystals. See, for example, the papers by Gibbons, Holm and Kupershmidt [20; 27; 28; 29]. For discussions of many of these theories from the Euler–Poincaré viewpoint, see Holm, Marsden and Ratiu [31] and Holm [23].

23.5 The heavy top formulation by the Kaluza–Klein construction

The Lagrangian in the heavy top action principle (102) may be transformed into a 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* $L_{\mathrm{KK}}\colon TQ_{\mathrm{KK}} \simeq TSO(3)\times T\mathbb{R}^3 \mapsto \mathbb{R}$ as

$$L_{\mathrm{KK}}(\mathbf{R},\mathbf{q},\dot{\mathbf{R}},\dot{\mathbf{q}};\hat{\mathbf{z}}) = L_{\mathrm{KK}}(\boldsymbol{\Omega},\boldsymbol{\Gamma},\mathbf{q},\dot{\mathbf{q}}) = \tfrac{1}{2}\langle\,\mathbb{I}\boldsymbol{\Omega},\boldsymbol{\Omega}\,\rangle + \tfrac{1}{2}|\boldsymbol{\Gamma}+\dot{\mathbf{q}}|^2, \tag{110}$$

with $\boldsymbol{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\hat{}$ 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 the kinetic energy of a metric, the *Kaluza–Klein metric* on TQ_{KK}.

The Legendre transformation for L_{KK} gives the momenta

$$\boldsymbol{\Pi} = \mathbb{I}\boldsymbol{\Omega} \qquad \text{and} \qquad \mathbf{p} = \boldsymbol{\Gamma} + \dot{\mathbf{q}}. \tag{111}$$

Since L_{KK} does not depend on $\mathbf{q}$, the Euler–Lagrange equation

$$\frac{d}{dt}\frac{\partial L_{\mathrm{KK}}}{\partial \dot{\mathbf{q}}} = \frac{\partial L_{\mathrm{KK}}}{\partial \mathbf{q}} = 0,$$

shows that $\mathbf{p} = \partial L_{\mathrm{KK}}/\partial \dot{\mathbf{q}}$ 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 23.1 with the Kaluza–Klein Lagrangian returns Euler's motion equation for the heavy top (100).

The Hamiltonian H_{KK} associated to L_{KK} by the Legendre transformation (111) is

$$\begin{aligned} H_{\mathrm{KK}}(\boldsymbol{\Pi}, \boldsymbol{\Gamma}, \mathbf{q}, \mathbf{p}) &= \boldsymbol{\Pi}\cdot\boldsymbol{\Omega} + \mathbf{p}\cdot\dot{\mathbf{q}} - L_{\mathrm{KK}}(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \mathbf{q}, \dot{\mathbf{q}}) \\ &= \tfrac{1}{2}\boldsymbol{\Pi}\cdot\mathbb{I}^{-1}\boldsymbol{\Pi} - \mathbf{p}\cdot\boldsymbol{\Gamma} + \tfrac{1}{2}|\mathbf{p}|^2 \\ &= \tfrac{1}{2}\boldsymbol{\Pi}\cdot\mathbb{I}^{-1}\boldsymbol{\Pi} + \tfrac{1}{2}|\mathbf{p}-\boldsymbol{\Gamma}|^2 - \tfrac{1}{2}|\boldsymbol{\Gamma}|^2. \end{aligned}$$

Recall that $\boldsymbol{\Gamma}$ is a unit vector. On the constant level set $|\boldsymbol{\Gamma}|^2 = 1$, the Kaluza–Klein Hamiltonian H_{KK} is a positive quadratic function, shifted by a constant. Likewise, on the constant level set $\mathbf{p} = -mg\,\boldsymbol{\chi}$, the Kaluza–Klein Hamiltonian H_{KK} is a function of only the variables $(\boldsymbol{\Pi}, \boldsymbol{\Gamma})$ and is equal to the Hamiltonian (104) for the heavy top up to an additive constant. Consequently, the Lie–Poisson equations for the Kaluza–Klein Hamiltonian H_{KK} now reproduce Euler's motion equation for the heavy top (100).

Exercise 23.5 Write the Kaluza–Klein construction on $SE(3) = SO(3)\circledS \mathbb{R}^3$.

24 Euler–Poincaré (EP) reduction theorem

Remark 24.1 (Geodesic motion) As emphasized by Arnold [2], in many interesting cases, the Euler–Poincaré equations on the dual of a Lie algebra $\mathfrak{g}^*$ correspond to *geodesic motion* on the corresponding group G. The relationship between the equations on $\mathfrak{g}^*$ and on G is the content of the basic Euler–Poincaré theorem discussed later. Similarly, on the Hamiltonian side, the preceding paragraphs described the relation between the Hamiltonian equations on T^*G and the Lie–Poisson equations on $\mathfrak{g}^*$. The issue of geodesic motion is especially simple: if either the Lagrangian on $\mathfrak{g}$ or the Hamiltonian on $\mathfrak{g}^*$ is purely quadratic, then the corresponding motion on the group is geodesic motion.

24.1 We were already speaking prose (EP)

Many of our previous considerations may be recast immediately as Euler–Poincaré equations.

- Rigid bodies $\simeq$ $\big(\mathrm{EP}SO(n)\big)$,
- Deforming bodies $\simeq\big(\mathrm{EP}GL_+(n,\mathbb{R})\big)$,
- Heavy tops $\simeq$ $\big(\mathrm{EP}SO(3)\times\mathbb{R}^3\big)$,
- EPDiff

24.2 Euler–Poincaré reduction

This lecture applies reduction by symmetry to Hamilton's principle. For a G–invariant Lagrangian defined on TG, this reduction takes Hamilton's principle from TG to $TG/G \simeq \mathfrak{g}$. Stationarity of the symmetry-reduced Hamilton's principle yields the Euler–Poincaré equations on $\mathfrak{g}^*$. The corresponding reduced Legendre transformation yields the Lie–Poisson Hamiltonian formulation of these equations.

Euler–Poincaré Reduction starts with a right (respectively, left) invariant Lagrangian $L\colon TG\to\mathbb{R}$ on the tangent bundle of a Lie group G. This means that $L(T_gR_h(v)) = L(v)$, respectively $L(T_gL_h(v)) = L(v)$, for all $g,h\in G$ and all $v\in T_gG$. In shorter notation, right invariance of the Lagrangian may be written as

$$L(g(t),\dot g(t)) = L(g(t)h,\dot g(t)h),$$

for all $h\in G$.

Theorem 24.2 (Euler–Poincaré Reduction) *Let G be a Lie group, $L\colon TG\to\mathbb{R}$ a right-invariant Lagrangian, and $l := L|_{\mathfrak{g}}\colon \mathfrak{g}\to\mathbb{R}$ be its restriction to $\mathfrak{g}$. For a curve $g(t)\in G$, let*

$$\xi(t) = \dot g(t)\cdot g(t)^{-1} := T_{g(t)}R_{g(t)^{-1}}\dot g(t)\in\mathfrak{g}.$$

Then the following four statements are equivalent:

(i) *$g(t)$ satisfies the Euler–Lagrange equations for Lagrangian L defined on G.*

(ii) *The variational principle*

$$\delta\int_a^b L(g(t),\dot g(t))dt = 0$$

holds, for variations with fixed endpoints.

(iii) *The (right invariant)* Euler–Poincaré equations *hold:*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = -\operatorname{ad}^*_{\xi}\frac{\delta l}{\delta \xi}.$$

(iv) *The variational principle*

$$\delta \int_a^b l(\xi(t))dt = 0$$

holds on $\mathfrak{g}$*, using variations of the form* $\delta\xi = \dot{\eta} - [\xi, \eta]$*, where* $\eta(t)$ *is an arbitrary path in* $\mathfrak{g}$ *which vanishes at the endpoints, that is,* $\eta(a) = \eta(b) = 0$.

Proof The proof consists of three steps.

Step I: Proof that (i) $\Longleftrightarrow$ (ii) This is Hamilton's principle: the Euler–Lagrange equations follow from stationary action for variations δg which vanish at the endpoints. (See Lecture 13.)

Step II: Proof that (ii) $\Longleftrightarrow$ (iv) Proving equivalence of the variational principles (ii) on TG and (iv) on $\mathfrak{g}$ for a right-invariant Lagrangian requires calculation of the variations $\delta\xi$ of $\xi = \dot{g}g^{-1}$ induced by δg. To simplify the exposition, the calculation will be done first for matrix Lie groups, then generalized to arbitrary Lie groups.

Step IIA: Proof that (ii) $\Longleftrightarrow$ (iv) for a matrix Lie group For $\xi = \dot{g}g^{-1}$, define $g_\epsilon(t)$ to be a family of curves in G such that $g_0(t) = g(t)$ and denote

$$\delta g := \left.\frac{dg_\epsilon(t)}{d\epsilon}\right|_{\epsilon=0.}$$

The variation of ξ is computed in terms of δg as

$$\delta\xi = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0}(\dot{g}_\epsilon g_\epsilon^{-1}) = \left.\frac{d^2 g}{dt\,d\epsilon}\right|_{\epsilon=0} g^{-1} - \dot{g}g^{-1}(\delta g)g^{-1}. \tag{112}$$

Set $\eta := g^{-1}\delta g$. That is, $\eta(t)$ is an arbitrary curve in $\mathfrak{g}$ which vanishes at the endpoints. The time derivative of η is computed as

$$\dot{\eta} = \frac{d\eta}{dt} = \frac{d}{dt}\left(\left(\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} g_\epsilon\right)g^{-1}\right) = \left.\frac{d^2 g}{dt\,d\epsilon}\right|_{\epsilon=0} g^{-1} - (\delta g)g^{-1}\dot{g}g^{-1}. \tag{113}$$

Taking the difference of (112) and (113) implies

$$\delta\xi - \dot{\eta} = -\dot{g}g^{-1}(\delta g)g^{-1} + (\delta g)g^{-1}\dot{g}g^{-1} = -\xi\eta + \eta\xi = -[\xi, \eta].$$

That is, for matrix Lie algebras,

$$\delta\xi = \dot{\eta} - [\xi, \eta],$$

where $[\xi, \eta]$ is the matrix commutator. Next, we notice that right invariance of L allows one to change variables in the Lagrangian by applying $g^{-1}(t)$ from the right, as

$$L(g(t), \dot{g}(t)) = L(e, \dot{g}(t)g^{-1}(t)) =: l(\xi(t)).$$

Combining this definition of the symmetry-reduced Lagrangian $l\colon \mathfrak{g} \to \mathbb{R}$ together with the formula for variations $\delta\xi$ just deduced proves the equivalence of (ii) and (iv) for matrix Lie groups.

Step IIB: Proof that (ii) $\Longleftrightarrow$ (iv) for an arbitrary Lie group The same proof extends to any Lie group G by using the following lemma.

Lemma 24.3 *Let $g\colon U \subset \mathbb{R}^2 \to G$ be a smooth map and denote its partial derivatives by*

$$\xi(t,\varepsilon) := T_{g(t,\varepsilon)} R_{g(t,\varepsilon)^{-1}} \frac{\partial g(t,\varepsilon)}{\partial t}, \qquad \eta(t,\varepsilon) := T_{g(t,\varepsilon)} R_{g(t,\varepsilon)^{-1}} \frac{\partial g(t,\varepsilon)}{\partial \varepsilon}. \tag{114}$$

Then

$$\frac{\partial \xi}{\partial \varepsilon} - \frac{\partial \eta}{\partial t} = -[\xi, \eta], \tag{115}$$

where $[\xi, \eta]$ is the Lie algebra bracket on $\mathfrak{g}$. Conversely, if $U \subset \mathbb{R}^2$ is simply connected and $\xi, \eta\colon U \to \mathfrak{g}$ are smooth functions satisfying (115), then there exists a smooth function $g\colon U \to G$ such that (114) holds.

Proof of Lemma 24.3 Write $\xi = \dot{g}g^{-1}$ and $\eta = g'g^{-1}$ in natural notation and express the partial derivatives $\dot{g} = \partial g/\partial t$ and $g' = \partial g/\partial \epsilon$ using the right translations as

$$\dot{g} = \xi \circ g \quad \text{and} \quad g' = \eta \circ g.$$

By the chain rule, these definitions have mixed partial derivatives

$$\dot{g}' = \xi' = \nabla\xi \cdot \eta \quad \text{and} \quad \dot{g}' = \dot{\eta} = \nabla\eta \cdot \xi.$$

The difference of the mixed partial derivatives implies the desired formula (115),

$$\xi' - \dot{\eta} = \nabla\xi \cdot \eta - \nabla\eta \cdot \xi = -[\xi, \eta] = -\operatorname{ad}_\xi \eta.$$

(Note the minus sign in the last two terms.) □

Step III: Proof of equivalence (iii) $\Longleftrightarrow$ (iv) Let us show that the reduced variational principle produces the Euler–Poincaré equations. We write the functional derivative of the reduced action $S_{\text{red}} = \int_a^b l(\xi)\,dt$ with Lagrangian $l(\xi)$ in terms of the natural pairing $\langle\cdot,\cdot\rangle$ between $\mathfrak{g}^*$ and $\mathfrak{g}$ as

$$\begin{aligned}\delta\int_a^b l(\xi(t))dt &= \int_a^b \left\langle \frac{\delta l}{\delta\xi}, \delta\xi \right\rangle dt = \int_a^b \left\langle \frac{\delta l}{\delta\xi}, \dot\eta - \mathrm{ad}_\xi\,\eta \right\rangle dt\\ &= \int_a^b \left\langle \frac{\delta l}{\delta\xi}, \dot\eta \right\rangle dt - \int_a^b \left\langle \frac{\delta l}{\delta\xi}, \mathrm{ad}_\xi\,\eta \right\rangle dt\\ &= -\int_a^b \left\langle \frac{d}{dt}\frac{\delta l}{\delta\xi} + \mathrm{ad}^*_\xi\,\frac{\delta l}{\delta\xi}, \eta \right\rangle dt.\end{aligned}$$

The last equality follows from integration by parts and vanishing of the variation $\eta(t)$ at the endpoints. Thus, stationarity $\delta\int_a^b l(\xi(t))dt = 0$ for any $\eta(t)$ that vanishes at the endpoints is equivalent to

$$\frac{d}{dt}\frac{\delta l}{\delta\xi} = -\,\mathrm{ad}^*_\xi\,\frac{\delta l}{\delta\xi},$$

which are the Euler–Poincaré equations. □

Remark 24.4 (Left-invariant Euler–Poincaré equations) The same theorem holds for left invariant Lagrangians on TG, except for a sign in the Euler–Poincaré equations,

$$\frac{d}{dt}\frac{\delta l}{\delta\xi} = +\,\mathrm{ad}^*_\xi\,\frac{\delta l}{\delta\xi},$$

which arises because left-invariant variations satisfy $\delta\xi = \dot\eta + [\xi,\ \eta]$ (with the opposite sign).

Exercise 24.5 Write out the corresponding proof of the Euler–Poincaré reduction theorem for left-invariant Lagrangians defined on the tangent space TG of a group G.

Reconstruction The procedure for reconstructing the solution $v(t) \in T_{g(t)}G$ of the Euler–Lagrange equations with initial conditions $g(0) = g_0$ and $\dot g(0) = v_0$ starting from the solution of the Euler–Poincaré equations is as follows. First, solve the initial value problem for the right-invariant Euler–Poincaré equations:

$$\frac{d}{dt}\frac{\delta l}{\delta\xi} = -\,\mathrm{ad}^*_\xi\,\frac{\delta l}{\delta\xi} \quad \text{with} \quad \xi(0) = \xi_0 := v_0 g_0^{-1}$$

Then from the solution for $\xi(t)$ reconstruct the curve $g(t)$ on the group by solving the “linear differential equation with time-dependent coefficients"

$$\dot g(t) = \xi(t)g(t) \quad \text{with} \quad g(0) = g_0.$$

The Euler–Poincaré reduction theorem guarantees then that $v(t) = \dot g(t) = \xi(t)\cdot g(t)$ is a solution of the Euler–Lagrange equations with initial condition $v_0 = \xi_0 g_0$.

Remark 24.6 A similar statement holds, with obvious changes for left-invariant Lagrangian systems on TG.

24.3 Reduced Legendre transformation

As in the equivalence relation between the Lagrangian and Hamiltonian formulations discussed earlier, the relationship between symmetry-reduced Euler–Poincaré and Lie–Poisson formulations is determined by the Legendre transformation.

Definition 24.7 The Legendre transformation $\mathbb{F}l\colon \mathfrak{g} \to \mathfrak{g}^*$ is defined by

$$\mathbb{F}l(\xi) = \frac{\delta l}{\delta \xi} = \mu.$$

Lie–Poisson Hamiltonian formulation Let $h(\mu) := \langle \mu, \xi \rangle - l(\xi)$. Assuming that $\mathbb{F}l$ is a diffeomorphism yields

$$\frac{\delta h}{\delta \mu} = \xi + \left\langle \mu, \frac{\delta \xi}{\delta \mu} \right\rangle - \left\langle \frac{\delta l}{\delta \xi}, \frac{\delta \xi}{\delta \mu} \right\rangle = \xi.$$

So the Euler–Poincaré equations for l are equivalent to the Lie–Poisson equations for h:

$$\frac{d}{dt}\left(\frac{\delta l}{\delta \xi}\right) = -\mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} \iff \dot{\mu} = -\mathrm{ad}^*_{\delta h/\delta \mu} \mu.$$

The Lie–Poisson equations may be written in the Poisson bracket form

$$\dot{f} = \{f, h\}, \tag{116}$$

where $f\colon \mathfrak{g}^* \to \mathbb{R}$ is an arbitrary smooth function and the bracket is the (right) Lie–Poisson bracket given by

$$\{f, h\}(\mu) = \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta h}{\delta \mu} \right] \right\rangle = -\left\langle \mu, \mathrm{ad}_{\delta h/\delta \mu} \frac{\delta f}{\delta \mu} \right\rangle = -\left\langle \mathrm{ad}^*_{\delta h/\delta \mu}\, \mu, \frac{\delta f}{\delta \mu} \right\rangle. \tag{117}$$

In the important case when ℓ is quadratic, the Lagrangian L is the quadratic form associated to a right invariant Riemannian metric on G. In this case, the Euler–Lagrange equations for L on G describe geodesic motion relative to this metric and these geodesics are then equivalently described by either the Euler–Poincaré, or the Lie–Poisson equations.

Exercise 24.8 Exercise 22.6 requires an extension of the pure EP reduction theorem for a Lagrangian $L\colon (TG \times TQ) \to \mathbb{R}$. Following the proof of the EP reduction theorem, make this extension.

Exercise 24.9 Compute the pure EP equations for geodesic motion on $SE(3)$. These equations turn out to be applicable to the motion of an ellipsoidal body through a fluid.

25 EPDiff: the Euler–Poincaré equation on the diffeomorphisms

25.1 The n–dimensional EPDiff equation and its properties

Eulerian geodesic motion of a fluid in n dimensions is generated as an EP equation via Hamilton's principle, when the Lagrangian is given by the kinetic energy. The kinetic energy defines a norm $\|\mathbf{u}\|^2$ for the Eulerian fluid velocity, taken as $\mathbf{u}(\mathbf{x},t)$: $R^n \times R^1 \to R^n$. The choice of the kinetic energy as a positive functional of fluid velocity $\mathbf{u}$ is a modeling step that depends upon the physics of the problem being studied. We shall choose the Lagrangian,

$$\|\mathbf{u}\|^2 = \int \mathbf{u}\cdot Q_{\rm op}\mathbf{u}\, d^n x = \int \mathbf{u}\cdot\mathbf{m}\, d^n x, \tag{118}$$

so that the positive-definite, symmetric, operator $Q_{\rm op}$ defines the norm $\|\mathbf{u}\|$, for appropriate (homogeneous, say, or periodic) boundary conditions. The EPDiff equation is the Euler–Poincaré equation for this Eulerian geodesic motion of a fluid. Namely,

$$\frac{d}{dt}\frac{\delta \ell}{\delta \mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}}\frac{\delta \ell}{\delta \mathbf{u}} = 0, \quad \text{with} \quad \ell[\mathbf{u}] = \tfrac{1}{2}\|\mathbf{u}\|^2. \tag{119}$$

Here ad^* is the dual of the vector-field ad-operation (the commutator) under the natural L^2 pairing $\langle\cdot,\cdot\rangle$ induced by the variational derivative $\delta\ell[\mathbf{u}] = \langle \delta\ell/\delta\mathbf{u}\,,\delta\mathbf{u}\rangle$. This pairing provides the definition of ad^*,

$$\langle \mathrm{ad}^*_{\mathbf{u}}\,\mathbf{m},\mathbf{v}\rangle = -\langle \mathbf{m}, \mathrm{ad}_{\mathbf{u}}\,\mathbf{v}\rangle, \tag{120}$$

where $\mathbf{u}$ and $\mathbf{v}$ are vector fields, $\mathrm{ad}_{\mathbf{u}}\mathbf{v} = [\mathbf{u},\mathbf{v}]$ is the commutator, that is, the *Lie bracket* given in components by (summing on repeated indices)

$$[\mathbf{u},\mathbf{v}]^i = u^j\frac{\partial v^i}{\partial x^j} - v^j\frac{\partial u^i}{\partial x^j}, \quad \text{or} \quad [\mathbf{u},\mathbf{v}] = \mathbf{u}\cdot\nabla\mathbf{v} - \mathbf{v}\cdot\nabla\mathbf{u}. \tag{121}$$

The notation $\mathrm{ad}_{\mathbf{u}}\,\mathbf{v} := [\mathbf{u},\mathbf{v}]$ formally denotes the adjoint action of the *right* Lie algebra of $\mathrm{Diff}(\mathcal{D})$ on itself, and $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ is the fluid momentum, a *one-form density* whose co-vector components are also denoted as $\mathbf{m}$.

If $\mathbf{u} = u^j \partial/\partial x^j$, $\mathbf{m} = m_i dx^i \otimes dV$, then the preceding formula for $\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV)$ has the *coordinate expression* in $\mathbb{R}^n$,

$$\left(\mathrm{ad}^*_{\mathbf{u}}\,\mathbf{m}\right)_i dx^i \otimes dV = \left(\frac{\partial}{\partial x^j}(u^j m_i) + m_j \frac{\partial u^j}{\partial x^i}\right) dx^i \otimes dV. \tag{122}$$

In this notation, the abstract EPDiff equation (119) may be written explicitly in Euclidean coordinates as a partial differential equation for a co-vector function

$$\mathbf{m}(\mathbf{x}, t)\colon\ R^n \times R^1 \to R^n.$$

Namely,

$$\frac{\partial}{\partial t}\mathbf{m} + \underbrace{\mathbf{u}\cdot\nabla\mathbf{m}}_{\text{Convection}} + \underbrace{\nabla\mathbf{u}^T\cdot\mathbf{m}}_{\text{Stretching}} + \underbrace{\mathbf{m}(\mathrm{div}\,\mathbf{u})}_{\text{Expansion}} = 0, \quad \text{with} \quad \mathbf{m} = \frac{\delta\ell}{\delta\mathbf{u}} = Q_{\mathrm{op}}\mathbf{u}. \tag{123}$$

To explain the terms in underbraces, we rewrite EPDiff as preservation of the one-form density of momentum along the characteristic curves of the velocity. Namely,

$$\frac{d}{dt}\big(\mathbf{m}\cdot d\mathbf{x}\otimes dV\big) = 0 \quad \text{along} \quad \frac{d\mathbf{x}}{dt} = \mathbf{u} = G * \mathbf{m}. \tag{124}$$

This form of the EPDiff equation also emphasizes its nonlocality, since the velocity is obtained from the momentum density by convolution against the Green's function G of the operator Q_{op}. Thus, $\mathbf{u} = G * \mathbf{m}$ with $Q_{\mathrm{op}}G = \delta(\mathbf{x})$, the Dirac measure. We may check that this "characteristic form" of EPDiff recovers its Eulerian form by computing directly,

$$\begin{aligned}
\frac{d}{dt}\big(\mathbf{m}\cdot d\mathbf{x}\otimes dV\big) &= \frac{d\mathbf{m}}{dt}\cdot d\mathbf{x}\otimes dV + \mathbf{m}\cdot d\frac{d\mathbf{x}}{dt}\otimes dV + \mathbf{m}\cdot d\mathbf{x}\otimes\left(\frac{d}{dt}dV\right)\\
&\qquad \text{along } \frac{d\mathbf{x}}{dt} = \mathbf{u} = G * \mathbf{m}\\
&= \left(\frac{\partial}{\partial t}\mathbf{m} + \mathbf{u}\cdot\nabla\mathbf{m} + \nabla\mathbf{u}^T\cdot\mathbf{m} + \mathbf{m}(\mathrm{div}\,\mathbf{u})\right)\cdot d\mathbf{x}\otimes dV = 0.
\end{aligned}$$

Exercise 25.1 Show that EPDiff may be written as

$$\left(\frac{\partial}{\partial t} + \mathcal{L}_{\mathbf{u}}\right)\big(\mathbf{m}\cdot d\mathbf{x}\otimes dV\big) = 0, \tag{125}$$

where $\mathcal{L}_{\mathbf{u}}$ is the Lie derivative with respect to the vector field with components $\mathbf{u} = G * \mathbf{m}$. Hint: How does this Lie-derivative form of EPDiff in (125) differ from its characteristic form (124)?

EPDiff may also be written equivalently in terms of the operators div, grad and curl in 2D and 3D as

$$\frac{\partial}{\partial t}\mathbf{m} - \mathbf{u}\times \operatorname{curl}\mathbf{m} + \nabla(\mathbf{u}\cdot\mathbf{m}) + \mathbf{m}(\operatorname{div}\mathbf{u}) = 0. \tag{126}$$

Thus, for example, its numerical solution would require an algorithm which has the capability to deal with the distinctions and relationships among the operators div, grad and curl.

25.2 Derivation of the n–dimensional EPDiff equation as geodesic flow

Let's derive the EPDiff equation (123) by following the proof of the EP reduction theorem leading to the Euler–Poincaré equations for right invariance in the form (119). Following this calculation for the present case yields

$$\begin{aligned}\delta\int_a^b l(\mathbf{u})dt &= \int_a^b \left\langle \frac{\delta l}{\delta \mathbf{u}}, \delta\mathbf{u}\right\rangle dt = \int_a^b \left\langle \frac{\delta l}{\delta \mathbf{u}}, \dot{\mathbf{v}} - \mathrm{ad}_{\mathbf{u}}\,\mathbf{v}\right\rangle dt\\ &= \int_a^b \left\langle \frac{\delta l}{\delta \mathbf{u}}, \dot{\mathbf{v}}\right\rangle dt - \int_a^b \left\langle \frac{\delta l}{\delta \mathbf{u}}, \mathrm{ad}_{\mathbf{u}}\,\mathbf{v}\right\rangle dt = -\int_a^b \left\langle \frac{d}{dt}\frac{\delta l}{\delta \mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}}\frac{\delta l}{\delta \mathbf{u}}, \mathbf{v}\right\rangle dt,\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ is the pairing between elements of the Lie algebra and its dual. In our case, this is the L^2 pairing, for example,

$$\left\langle \frac{\delta l}{\delta \mathbf{u}}, \delta\mathbf{u}\right\rangle = \int \frac{\delta l}{\delta u^i}\,\delta u^i\, d^n x$$

This pairing allows us to compute the coordinate form of the EPDiff equation explicitly, as

$$\begin{aligned}\int_a^b \left\langle \frac{\delta l}{\delta \mathbf{u}}, \delta\mathbf{u}\right\rangle dt &= \int_a^b dt \int \frac{\delta l}{\delta u^i}\left(\frac{\partial v^i}{\partial t} + u^j\frac{\partial v^i}{\partial x^j} - v^j\frac{\partial u^i}{\partial x^j}\right) d^n x\\ &= -\int_a^b dt \int \left\{\frac{\partial}{\partial t}\frac{\delta l}{\delta u^i} + \frac{\partial}{\partial x^j}\left(\frac{\delta l}{\delta u^i}u^j\right) + \frac{\delta l}{\delta u^j}\frac{\partial u^j}{\partial x^i}\right\} v^i\, d^n x\end{aligned}$$

Substituting $\mathbf{m} = \delta l/\delta\mathbf{u}$ now recovers the coordinate forms for the coadjoint action of vector fields in (122) and the EPDiff equation itself in (123). When $\ell[\mathbf{u}] = \frac{1}{2}\|\mathbf{u}\|^2$, EPDiff describes geodesic motion on the diffeomorphisms with respect to the norm $\|\mathbf{u}\|$.

Lemma 25.2 *In Step IIB of the proof of the Euler–Poincaré reduction theorem (that (ii) $\Longleftrightarrow$ (iv) for an arbitrary Lie group) a certain formula for the variations for time-dependent vector fields was employed. That formula was employed again in the* *calculation above as*

$$\delta\mathbf{u} = \dot{\mathbf{v}} - \mathrm{ad}_{\mathbf{u}}\,\mathbf{v}. \tag{127}$$

This formula may be rederived as follows in the present context. We write $\mathbf{u} = \dot{g}g^{-1}$ *and* $\mathbf{v} = g'g^{-1}$ *in natural notation and express the partial derivatives* $\dot{g} = \partial g/\partial t$ *and* $g' = \partial g/\partial\epsilon$ *using the right translations as*

$$\dot{g} = \mathbf{u}\circ g \quad \text{and} \quad g' = \mathbf{v}\circ g.$$

To compute the mixed partials, consider the chain rule for say $\mathbf{u}(g(t,\epsilon)\mathbf{x}_0)$ *and set* $\mathbf{x}(t,\epsilon) = g(t,\epsilon)\cdot\mathbf{x}_0$. *Then,*

$$\mathbf{u}' = \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\cdot\frac{\partial\mathbf{x}}{\partial\epsilon} = \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\cdot g'(t,\epsilon)\mathbf{x}_0 = \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\cdot g'g^{-1}\mathbf{x} = \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\cdot\mathbf{v}(\mathbf{x}).$$

The chain rule for $\dot{\mathbf{v}}$ *gives a similar formula with* $\mathbf{u}$ *and* $\mathbf{v}$ *exchanged. Thus, the chain rule gives two expressions for the mixed partial derivative* $\dot{g}'$ *as*

$$\dot{g}' = \mathbf{u}' = \nabla\mathbf{u}\cdot\mathbf{v} \quad \text{and} \quad \dot{g}' = \dot{\mathbf{v}} = \nabla\mathbf{v}\cdot\mathbf{u}.$$

The difference of the mixed partial derivatives then implies the desired formula (127), since

$$\mathbf{u}' - \dot{\mathbf{v}} = \nabla\mathbf{u}\cdot\mathbf{v} - \nabla\mathbf{v}\cdot\mathbf{u} = -[\mathbf{u},\mathbf{v}] = -\mathrm{ad}_{\mathbf{u}}\,\mathbf{v}.$$

26 EPDiff: the Euler–Poincaré equation on the diffeomorphisms

In this lecture, we shall discuss the solutions of EPDiff for pressureless compressible geodesic motion in one spatial dimension. This is the *EPDiff equation* in 1D, [9]

$$\partial_t m + \mathrm{ad}^*_u\, m = 0, \quad \text{or, equivalently,} \tag{128}$$

$$\partial_t m + u m_x + 2u_x m = 0, \quad \text{with } m = Q_{\mathrm{op}}u. \tag{129}$$

[9] A one-form density in 1D takes the form $m(dx)^2$ and the EP equation is given by

$$\frac{d}{dt}\big(m\,(dx)^2\big) = \frac{dm}{dt}(dx)^2 + 2m\,(du)(dx) = 0 \quad \text{with} \quad \frac{d}{dt}dx = du = u_x dx \quad \text{and} \quad u = G*m,$$

where $G*m$ denotes convolution with a function G on the real line.

- The EPDiff equation describes geodesic motion on the diffeomorphism group with respect to a family of metrics for the fluid velocity $u(t,x)$, with notation,

$$m = \frac{\delta \ell}{\delta u} = Q_{\mathrm{op}} u \quad \text{for a kinetic-energy Lagrangian} \tag{130}$$

$$\ell(u) = \tfrac{1}{2} \int u \, Q_{\mathrm{op}} u \, dx = \tfrac{1}{2} \|u\|^2. \tag{131}$$

- In one dimension, Q_{op} in equation (130) is a positive, symmetric operator that defines the kinetic energy metric for the velocity.
- The EPDiff equation (129) is written in terms of the variable $m = \delta\ell/\delta u$. It is appropriate to call this variational derivative m, because it is the momentum density associated with the fluid velocity u.
- Physically, the first nonlinear term in the EPDiff equation (129) is fluid transport.
- The coefficient 2 arises in the second nonlinear term, because, in one dimension, two of the summands in $\mathrm{ad}^*_u\, m = u m_x + 2 u_x m$ are the *same*, cf. equation (122).
- The momentum is expressed in terms of the velocity by $m = \delta\ell/\delta u = Q_{\mathrm{op}} u$. Equivalently, for solutions that vanish at spatial infinity, one may think of the velocity as being obtained from the convolution,

$$u(x) = G * m(x) = \int G(x-y) m(y) \, dy, \tag{132}$$

  where G is the Green's function for the operator Q_{op} on the real line.
- The operator Q_{op} and its Green's function G are chosen to be even under reflection, $G(-x) = G(x)$, so that u and m have the same parity. Moreover, the EPDiff equation (129) conserves the total momentum $M = \int m(y)\, dy$, for any even Green's function.

Exercise 26.1 Show that equation (129) conserves $M = \int m(y)\, dy$ for any even Green's function $G(-x) = G(x)$, for either periodic, or homogeneous boundary conditions.

- The traveling wave solutions of 1D EPDiff when the Green's function G is chosen to be even under reflection are the "pulsons,"

$$u(x,t) = c\, G(x - ct).$$

Exercise 26.2 Prove this statement, that the traveling wave solutions of 1D EPDiff are pulsons when the Green's function is even. What role is played in the solution by the Green's function being even? Hint: Evaluate the derivative of an even function at $x = 0$.

- See Fringer and Holm [19] and references therein for further discussions and numerical simulations of the pulson solutions of the 1D EPDiff equation.

26.1 Pulsons

The EPDiff equation (129) on the real line has the remarkable property that its solutions *collectivize*[10] into the finite dimensional solutions of the "N–pulson" form that was discovered for a special form of G in Camassa and Holm [10], then was extended for *any* even G in Fringer and Holm [19],

$$u(x,t)=\sum_{i=1}^{N} p_i(t)\,G(x-q_i(t)). \tag{133}$$

Since $G(x)$ is the Green's function for the operator $Q_{\rm op}$, the corresponding solution for the momentum $m=Q_{\rm op}u$ is given by a sum of delta functions,

$$m(x,t)=\sum_{i=1}^{N} p_i(t)\,\delta(x-q_i(t)). \tag{134}$$

Thus, the time-dependent "collective coordinates" $q_i(t)$ and $p_i(t)$ are the positions and velocities of the N pulses in this solution. These parameters satisfy the finite dimensional geodesic motion equations obtained as canonical Hamiltonian equations

$$\dot{q}_i=\frac{\partial H_N}{\partial p_i}=\sum_{j=1}^{N} p_j\;G(q_i-q_j), \tag{135}$$

$$\dot{p}_i=-\frac{\partial H_N}{\partial q_i}=-p_i\sum_{j=1}^{N} p_j\;G'(q_i-q_j), \tag{136}$$

in which the Hamiltonian is given by the quadratic form,

$$H_N=\tfrac{1}{2}\sum_{i,j=1}^{N} p_i\;p_j\;G(q_i-q_j). \tag{137}$$

Remark 26.3 In a certain sense, equations (135)–(136) comprise the analog for the peakon momentum relation (134) of the "symmetric generalized rigid body equations" in (65).

[10] See Guillemin and Sternberg [21] for discussions of the concept of collective variables for Hamiltonian theories. We will discuss the collectivization for the EPDiff equation later from the viewpoint of momentum maps.

Thus, the canonical equations for the Hamiltonian H_N describe the nonlinear collective interactions of the N–pulson solutions of the EPDiff equation (129) as finite-dimensional geodesic motion of a particle on an N–dimensional surface whose co-metric is

$$G^{ij}(q) = G(q_i - q_j). \tag{138}$$

Fringer and Holm [19] showed numerically that the N–pulson solutions describe the emergent patterns in the solution of the initial value problem for EPDiff equation (129) with spatially confined initial conditions.

Exercise 26.4 Equations (135)–(136) describe geodesic motion.

(1) Write the Lagrangian and Euler–Lagrange equations for this motion.

(2) Solve equations (135)–(136) for $N = 2$ when $\lim_{|x|\to\infty} G(x) = 0$.
   - (a) Why should the solution be described as exchange of momentum in elastic collisions?
   - (b) Consider both head-on and overtaking collisions.
   - (c) Consider the antisymmetric case, when the total momentum vanishes.

Integrability Calogero and Francoise [8; 9] found that for any finite number N the Hamiltonian equations for H_N in (137) are completely integrable in the Liouville sense[11] for

$$\begin{aligned} G &\equiv G_1(x) = \lambda + \mu\cos(\nu x) + \mu_1 \sin(\nu|x|) \\ \text{and} \qquad G &\equiv G_2(x) = \alpha + \beta|x| + \gamma x^2, \end{aligned}$$

with λ, μ, μ_1, ν, and α, β, γ being arbitrary constants, such that λ and μ are real and μ_1 and ν both real or both imaginary.[12] Particular cases of G_1 and G_2 are the peakons $G_1(x) = e^{-|x|/\alpha}$ of Camassa and Holm [10] and the compactons $G_2(x) = \max(1-|x|, 0)$ of the Hunter–Saxton equation, (see Hunter and Zheng [34]). The latter is the EPDiff equation (129), with $\ell(u) = \frac{1}{2}\int u_x^2\,dx$ and thus $m = -u_{xx}$.

Lie–Poisson Hamiltonian form of EPDiff In terms of m, the conserved energy Hamiltonian for the EPDiff equation (129) is obtained by Legendre transforming the kinetic energy Lagrangian, as

$$h = \left\langle \frac{\delta\ell}{\delta u}, u \right\rangle - \ell(u).$$

[11] A Hamiltonian system is integrable in the Liouville sense, if the number of independent constants of motion in involution is the same as the number of its degrees of freedom.

[12] This choice of the constants keeps H_N real in (137).

Thus, the Hamiltonian depends on m, as

$$h(m) = \tfrac{1}{2}\int m(x)G(x-y)m(y)\,dxdy,$$

which also reveals the geodesic nature of the EPDiff equation (129) and the role of $G(x)$ in the kinetic energy metric on the Hamiltonian side.

The corresponding *Lie–Poisson bracket* for EPDiff as a Hamiltonian evolution equation is given by,

$$\partial_t m = \{m,h\} = -(\partial m + m\partial)\frac{\delta h}{\delta m} \quad \text{and} \quad \frac{\delta h}{\delta m} = u,$$

which recovers the starting equation and indicates some of its connections with fluid equations on the Hamiltonian side. For any two smooth functionals f,h of m in the space for which the solutions of EPDiff exist, this Lie–Poisson bracket may be expressed as

$$\{f,h\} = -\int \frac{\delta f}{\delta m}(\partial m + m\partial)\frac{\delta h}{\delta m}dx = -\int m\left[\frac{\delta f}{\delta m},\frac{\delta h}{\delta m}\right]dx$$

where $[\cdot,\cdot]$ denotes the Lie algebra bracket of vector fields. That is,

$$\left[\frac{\delta f}{\delta m},\frac{\delta h}{\delta m}\right] = \frac{\delta f}{\delta m}\partial\frac{\delta h}{\delta m} - \frac{\delta h}{\delta m}\partial\frac{\delta f}{\delta m}.$$

Exercise 26.5 What is the Casimir for this Lie–Poisson bracket? What does it mean from the viewpoint of coadjoint orbits?

26.2 Peakons

The case $G(x) = e^{-|x|/\alpha}$ with a constant length scale α is the Green's function for which the operator in the kinetic energy Lagrangian (130) is $Q_{\rm op} = 1-\alpha^2\partial_x^2$. For this (Helmholtz) operator $Q_{\rm op}$, the Lagrangian and corresponding kinetic energy norm are given by,

$$\ell[u] = \tfrac{1}{2}\|u\|^2 = \tfrac{1}{2}\int uQ_{\rm op}u\,dx = \tfrac{1}{2}\int u^2+\alpha^2u_x^2\,dx, \quad \text{for} \quad \lim_{|x|\to\infty} u = 0.$$

This Lagrangian is the H^1 norm of the velocity in one dimension. In this case, the EPDiff equation (129) is also the zero-dispersion limit of the completely integrable CH equation for unidirectional shallow water waves first derived in Camassa and Holm [10],

$$m_t + um_x + 2mu_x = -c_0u_x + \gamma u_{xxx}, \qquad m = u - \alpha^2u_{xx}. \tag{139}$$

This equation describes shallow water dynamics as completely integrable soliton motion at quadratic order in the asymptotic expansion for unidirectional shallow water waves on a free surface under gravity. See Dullin, Gottwald and Holm [14; 15; 16] for more details and explanations of this asymptotic expansion for unidirectional shallow water waves to quadratic order.

Because of the relation $m = u - \alpha^2 u_{xx}$, equation (139) is nonlocal. In other words, it is an integral-partial differential equation. In fact, after writing equation (139) in the equivalent form,

$$(1-\alpha^2\partial^2)(u_t + uu_x) = -\partial\big(u^2 + \tfrac{1}{2}\alpha^2 u_x^2\big) - c_0 u_x + \gamma u_{xxx}, \tag{140}$$

one sees the interplay between local and nonlocal linear dispersion in its phase velocity relation,

$$\frac{\omega}{k} = \frac{c_0 - \gamma k^2}{1+\alpha^2 k^2}, \tag{141}$$

for waves with frequency ω and wave number k linearized around $u = 0$. For $\gamma/c_0 < 0$, short waves and long waves travel in the same direction. Long waves travel faster than short ones (as required in shallow water) provided $\gamma/c_0 > -\alpha^2$. Then the phase velocity lies in the interval $\omega/k \in (-\gamma/\alpha^2, c_0]$.

The famous Korteweg–de Vries (KdV) soliton equation,

$$u_t + 3uu_x = -c_0 u_x + \gamma u_{xxx}, \tag{142}$$

emerges at *linear* order in the asymptotic expansion for shallow water waves, in which one takes $\alpha^2 \to 0$ in (140) and (141). In KdV, the parameters c_0 and γ are seen as deformations of the *Riemann equation*,

$$u_t + 3uu_x = 0.$$

The parameters c_0 and γ represent linear wave dispersion, which modifies and eventually balances the tendency for nonlinear waves to steepen and break. The parameter α, which introduces nonlocality, also regularizes this nonlinear tendency, even in the absence of c_0 and γ.

27 Diffeons – singular momentum solutions of the EPDiff equation for geodesic motion in higher dimensions

As an example of the EP theory in higher dimensions, we shall generalize the one-dimensional pulson solutions of the previous section to n–dimensions. The corresponding singular momentum solutions of the EPDiff equation in higher dimensions are called "diffeons."

27.1 n–dimensional EPDiff equation

Eulerian geodesic motion of a fluid in n–dimensions is generated as an EP equation via Hamilton's principle, when the Lagrangian is given by the kinetic energy. The kinetic energy defines a norm $\|\mathbf{u}\|^2$ for the Eulerian fluid velocity, $\mathbf{u}(\mathbf{x},t)$: $R^n \times R^1 \to R^n$. As mentioned earlier, the choice of the kinetic energy as a positive functional of fluid velocity $\mathbf{u}$ is a modeling step that depends upon the physics of the problem being studied. Following our earlier procedure, as in equations (118) and (119), we shall choose the Lagrangian,

(143) $$\|\mathbf{u}\|^2 = \int \mathbf{u}\cdot Q_{\rm op}\mathbf{u}\, d^n x = \int \mathbf{u}\cdot\mathbf{m}\, d^n x,$$

so that the positive-definite, symmetric, operator $Q_{\rm op}$ defines the norm $\|\mathbf{u}\|$, for appropriate boundary conditions and the EPDiff equation for Eulerian geodesic motion of a fluid emerges,

(144) $$\frac{d}{dt}\frac{\delta\ell}{\delta\mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}}\frac{\delta\ell}{\delta\mathbf{u}} = 0, \quad \text{with} \quad \ell[\mathbf{u}] = \tfrac{1}{2}\|\mathbf{u}\|^2.$$

Legendre transforming to the Hamiltonian side The corresponding Legendre transform yields the following invertible relations between momentum and velocity,

(145) $$\mathbf{m} = Q_{\rm op}\mathbf{u} \quad \text{and} \quad \mathbf{u} = G*\mathbf{m},$$

where G is the *Green's function* for the operator $Q_{\rm op}$, assuming appropriate boundary conditions (on $\mathbf{u}$) that allow inversion of the operator $Q_{\rm op}$ to determine $\mathbf{u}$ from $\mathbf{m}$.

The corresponding *Hamiltonian* is,

(146) $$h[\mathbf{m}] = \langle\mathbf{m},\mathbf{u}\rangle - \tfrac{1}{2}\|\mathbf{u}\|^2 = \tfrac{1}{2}\int \mathbf{m}\cdot G*\mathbf{m}\; d^n x \equiv \tfrac{1}{2}\|\mathbf{m}\|^2,$$

which also defines a norm $\|\mathbf{m}\|$ via a convolution kernel G that is symmetric and positive, when the Lagrangian $\ell[\mathbf{u}]$ is a norm. As expected, the norm $\|\mathbf{m}\|$ given by the Hamiltonian $h[\mathbf{m}]$ specifies the velocity $\mathbf{u}$ in terms of its Legendre-dual momentum $\mathbf{m}$ by the variational operation,

(147) $$\mathbf{u} = \frac{\delta h}{\delta\mathbf{m}} = G*\mathbf{m} \equiv \int G(\mathbf{x}-\mathbf{y})\,\mathbf{m}(\mathbf{y})\, d^n y.$$

We shall choose the kernel $G(\mathbf{x}-\mathbf{y})$ to be translation-invariant (so Noether's theorem implies that total momentum $\mathbf{M} = \int \mathbf{m}\, d^n x$ is conserved) and symmetric under spatial reflections (so that $\mathbf{u}$ and $\mathbf{m}$ have the same parity).

After the Legendre transformation (146), the EPDiff equation (144) appears in its equivalent *Lie–Poisson Hamiltonian form*,

$$\frac{\partial}{\partial t}\mathbf{m} = \{\mathbf{m}, h\} = -\operatorname{ad}^*_{\delta h/\delta \mathbf{m}}\mathbf{m}. \tag{148}$$

Here the operation $\{\cdot,\cdot\}$ denotes the Lie–Poisson bracket dual to the (right) action of vector fields amongst themselves by vector-field commutation

$$\{f, h\} = -\left\langle \mathbf{m}, \left[\frac{\delta f}{\delta \mathbf{m}}, \frac{\delta h}{\delta \mathbf{m}}\right]\right\rangle$$

For more details and additional background concerning the relation of classical EP theory to Lie–Poisson Hamiltonian equations, see Holm, Marsden and Ratiu [44; 31].

In a moment we will also consider the momentum maps for EPDiff.

27.2 Diffeons: n–dimensional analogs of pulsons for the EPDiff equation

The momentum for the one-dimensional pulson solutions (134) on the real line is supported at points via the Dirac delta measures in its solution ansatz,

$$m(x,t) = \sum_{i=1}^{N} p_i(t)\,\delta\big(x - q_i(t)\big), \quad m \in R^1. \tag{149}$$

We shall develop n–dimensional analogs of these one-dimensional pulson solutions for the Euler–Poincaré equation (126) by generalizing this solution ansatz to allow measure-valued n–dimensional vector solutions $\mathbf{m} \in R^n$ for which the Euler–Poincaré momentum is supported on co-dimension–k *subspaces* R^{n-k} with integer $k \in [1, n]$. For example, one may consider the two-dimensional vector momentum $\mathbf{m} \in R^2$ in the plane that is supported on one-dimensional curves (momentum fronts). Likewise, in three dimensions, one could consider two-dimensional momentum surfaces (sheets), one-dimensional momentum filaments, etc. The corresponding vector momentum ansatz that we shall use is the following, cf. the pulson solutions (149),

$$\mathbf{m}(\mathbf{x},t) = \sum_{i=1}^{N} \int \mathbf{P}_i(s,t)\,\delta\big(\mathbf{x} - \mathbf{Q}_i(s,t)\big)ds, \quad \mathbf{m} \in R^n. \tag{150}$$

Here, $\mathbf{P}_i, \mathbf{Q}_i \in R^n$ for $i = 1, 2, \ldots, N$. For example, when $n - k = 1$, so that $s \in R^1$ is one-dimensional, the delta function in solution (150) supports an evolving family of vector-valued curves, called *momentum filaments*. (For simplicity of notation, we suppress the implied subscript i in the arclength s for each $\mathbf{P}_i$ and $\mathbf{Q}_i$.) The Legendre-dual relations (145) imply that the velocity corresponding to the momentum filament ansatz (150) is,

$$\mathbf{u}(\mathbf{x},t) = G * \mathbf{m} = \sum_{j=1}^{N} \int \mathbf{P}_j(s',t)\, G\big(\mathbf{x} - \mathbf{Q}_j(s',t)\big) ds'. \tag{151}$$

Just as for the 1D case of the pulsons, we shall show that substitution of the nD solution ansatz (150) and (151) into the EPDiff equation (123) produces canonical geodesic Hamiltonian equations for the n–dimensional vector parameters $\mathbf{Q}_i(s,t)$ and $\mathbf{P}_i(s,t)$, $i = 1,2,\dots,N$.

27.2.1 Canonical Hamiltonian dynamics of diffeon momentum filaments in R^n

For definiteness in what follows, we shall consider the example of momentum filaments $\mathbf{m} \in R^n$ supported on one-dimensional space curves in R^n, so $s \in R^1$ is the arclength parameter of one of these curves. This solution ansatz is reminiscent of the Biot–Savart Law for vortex filaments, although the flow is not incompressible. The dynamics of momentum surfaces, for $s \in R^k$ with $k < n$, follow a similar analysis.

Substituting the momentum filament ansatz (150) for $s \in R^1$ and its corresponding velocity (151) into the Euler–Poincaré equation (123), then integrating against a smooth test function $\phi(\mathbf{x})$ implies the following canonical equations (denoting explicit summation on $i, j \in 1,2,\dots N$),

$$\frac{\partial}{\partial t}\mathbf{Q}_i(s,t) = \sum_{j=1}^{N} \int \mathbf{P}_j(s',t)\, G(\mathbf{Q}_i(s,t) - \mathbf{Q}_j(s',t))\big) ds' = \frac{\delta H_N}{\delta \mathbf{P}_i}, \tag{152}$$

$$\begin{aligned} \frac{\partial}{\partial t}\mathbf{P}_i(s,t) &= -\sum_{j=1}^{N} \int \big(\mathbf{P}_i(s,t)\cdot\mathbf{P}_j(s',t)\big) \frac{\partial}{\partial \mathbf{Q}_i(s,t)} G\big(\mathbf{Q}_i(s,t) - \mathbf{Q}_j(s',t)\big)\, ds' \\ &= -\frac{\delta H_N}{\delta \mathbf{Q}_i}, \qquad \text{(sum on } j\text{, no sum on } i\text{).} \end{aligned} \tag{153}$$

The dot product $\mathbf{P}_i \cdot \mathbf{P}_j$ denotes the inner, or scalar, product of the two vectors $\mathbf{P}_i$ and $\mathbf{P}_j$ in R^n. Thus, the solution ansatz (150) yields a closed set of *integro-partial-differential equations (IPDEs)* given by (152) and (153) for the vector parameters $\mathbf{Q}_i(s,t)$ and $\mathbf{P}_i(s,t)$ with $i = 1,2\dots N$. These equations are generated canonically by the following Hamiltonian function $H_N\colon (R^n \times R^n)^{\otimes N} \to R$,

$$H_N = \tfrac{1}{2} \iint \sum_{i,j=1}^{N} \big(\mathbf{P}_i(s,t)\cdot\mathbf{P}_j(s',t)\big)\, G\big(\mathbf{Q}_i(s,t) - \mathbf{Q}_j(s',t)\big)\, ds\, ds'. \tag{154}$$

This Hamiltonian arises by substituting the momentum ansatz (150) into the Hamiltonian (146) obtained from the Legendre transformation of the Lagrangian corresponding to the kinetic energy norm of the fluid velocity. Thus, the evolutionary IPDE system (152) and (153) represents canonically Hamiltonian geodesic motion on the space of curves in R^n with respect to the co-metric given on these curves in (154). The Hamiltonian $H_N = \frac{1}{2}\|\mathbf{P}\|^2$ in (154) defines the norm $\|\mathbf{P}\|$ in terms of this co-metric that combines convolution using the Green's function G and sum over filaments with the scalar product of momentum vectors in R^n.

Remark 27.1 Note the Lagrangian property of the s coordinate, since

$$\frac{\partial}{\partial t}\mathbf{Q}_i(s,t) = \mathbf{u}(\mathbf{Q}_i(s,t),t).$$

28 Singular solution momentum map $\mathbf{J}_{\mathrm{Sing}}$ for diffeons

The diffeon momentum filament ansatz (150) reduces, and *collectivizes* the solution of the geodesic EP PDE (123) in $n+1$ dimensions into the system (152) and (153) of $2N$ canonical evolutionary IPDEs. One can summarize the mechanism by which this process occurs, by saying that the map that implements the canonical $(\mathbf{Q},\mathbf{P})$ variables in terms of singular solutions is a (cotangent bundle) momentum map. Such momentum maps are Poisson maps; so the canonical Hamiltonian nature of the dynamical equations for $(\mathbf{Q},\mathbf{P})$ fits into a general theory which also provides a framework for suggesting other avenues of investigation.

Theorem 28.1 *The momentum ansatz* (150) *for measure-valued solutions of the* EPDiff *equation* (123)*, defines an equivariant momentum map*

$$\mathbf{J}_{\mathrm{Sing}}\colon T^*\operatorname{Emb}(S,\mathbb{R}^n) \to \mathfrak{X}(\mathbb{R}^n)^*$$

that is called the singular solution momentum map *in Holm and Marsden [30].*

We shall explain the notation used in the theorem's statement in the course of its proof. Right away, however, we note that the sense of "defines" is that the momentum solution ansatz (150) expressing $\mathbf{m}$ (a vector function of spatial position $\mathbf{x}$) in terms of $\mathbf{Q},\mathbf{P}$ (which are functions of s) can be regarded as a map from the space of $(\mathbf{Q}(s),\mathbf{P}(s))$ to the space of $\mathbf{m}$'s. This will turn out to be the Lagrange-to-Euler map for the fluid description of the singular solutions.

Following Holm and Marsden [30], we shall give two proofs of this result from two rather different viewpoints. The first proof below uses the formula for a momentum map for a cotangent lifted action, while the second proof focuses on a Poisson bracket computation. Each proof also explains the context in which one has a momentum map. (See Marsden and Ratiu [44] for general background on momentum maps.)

First proof For simplicity and without loss of generality, let us take $N = 1$ and so suppress the index a. That is, we shall take the case of an isolated singular solution. As the proof will show, this is not a real restriction.

To set the notation, fix a k–dimensional manifold S with a given volume element and whose points are denoted $s \in S$. Let $\mathrm{Emb}(S, \mathbb{R}^n)$ denote the set of smooth embeddings $\mathbf{Q}\colon S \to \mathbb{R}^n$. (If the EPDiff equations are taken on a manifold M, replace $\mathbb{R}^n$ with M.) Under appropriate technical conditions, which we shall just treat formally here, $\mathrm{Emb}(S, \mathbb{R}^n)$ is a smooth manifold. (See, for example, Ebin and Marsden [17], and Marsden and Hughes [43] for a discussion and references.)

The tangent space $T_{\mathbf{Q}}\mathrm{Emb}(S, \mathbb{R}^n)$ to $\mathrm{Emb}(S, \mathbb{R}^n)$ at the point $\mathbf{Q} \in \mathrm{Emb}(S, \mathbb{R}^n)$ is given by the space of *material velocity fields*, namely the linear space of maps $\mathbf{V}\colon S \to \mathbb{R}^n$ that are vector fields over the map $\mathbf{Q}$. The dual space to this space will be identified with the space of one-form densities over $\mathbf{Q}$, which we shall regard as maps $\mathbf{P}\colon S \to (\mathbb{R}^n)^*$. In summary, the cotangent bundle $T^*\mathrm{Emb}(S, \mathbb{R}^n)$ is identified with the space of pairs of maps $(\mathbf{Q}, \mathbf{P})$.

These give us the domain space for the singular solution momentum map. Now we consider the action of the symmetry group. Consider the group $\mathfrak{G} = \mathrm{Diff}$ of diffeomorphisms of the space $\mathfrak{S}$ in which the EPDiff equations are operating, concretely in our case $\mathbb{R}^n$. Let it act on $\mathfrak{S}$ by composition on the *left*. Namely for $\eta \in \mathrm{Diff}(\mathbb{R}^n)$, we let

$$\eta \cdot \mathbf{Q} = \eta \circ \mathbf{Q}. \tag{155}$$

Now lift this action to the cotangent bundle $T^*\mathrm{Emb}(S, \mathbb{R}^n)$ in the standard way (see, for instance, Marsden and Ratiu [44] for this construction). This lifted action is a symplectic (and hence Poisson) action and has an equivariant momentum map. *We claim that this momentum map is precisely given by the ansatz* (150).

To see this, one only needs to recall and then apply the general formula for the momentum map associated with an action of a general Lie group $\mathfrak{G}$ on a configuration manifold Q and cotangent lifted to T^*Q.

First let us recall the general formula. Namely, the momentum map is the map $\mathbf{J}\colon T^*Q \to \mathfrak{g}^*$ ($\mathfrak{g}^*$ denotes the dual of the Lie algebra $\mathfrak{g}$ of $\mathfrak{G}$) defined by

$$\mathbf{J}(\alpha_q) \cdot \xi = \left\langle \alpha_q, \xi_Q(q) \right\rangle, \tag{156}$$

where $\alpha_q \in T_q^*Q$ and $\xi \in \mathfrak{g}$, where ξ_Q is the infinitesimal generator of the action of $\mathfrak{G}$ on Q associated to the Lie algebra element ξ, and where $\left\langle \alpha_q, \xi_Q(q) \right\rangle$ is the natural pairing of an element of T_q^*Q with an element of T_qQ.

Now we apply this formula to the special case in which the group $\mathfrak{G}$ is the diffeomorphism group $\mathrm{Diff}(\mathbb{R}^n)$, the manifold Q is $\mathrm{Emb}(S,\mathbb{R}^n)$ and where the action of the group on $\mathrm{Emb}(S,\mathbb{R}^n)$ is given by (155). The sense in which the Lie algebra of $\mathfrak{G} = \mathrm{Diff}$ is the space $\mathfrak{g} = \mathfrak{X}$ of vector fields is well-understood. Hence, its dual is naturally regarded as the space of one-form densities. The momentum map is thus a map $\mathbf{J}\colon T^*\,\mathrm{Emb}(S,\mathbb{R}^n) \to \mathfrak{X}^*$.

With $\mathbf{J}$ given by (156), we only need to work out this formula. First, we shall work out the infinitesimal generators. Let $X \in \mathfrak{X}$ be a Lie algebra element. By differentiating the action (155) with respect to η in the direction of X at the identity element we find that the infinitesimal generator is given by

$$X_{\mathrm{Emb}(S,\mathbb{R}^n)}(\mathbf{Q}) = X \circ \mathbf{Q}.$$

Thus, taking α_q to be the cotangent vector $(\mathbf{Q},\mathbf{P})$, equation (156) gives

$$\begin{aligned}\langle \mathbf{J}(\mathbf{Q},\mathbf{P}), X\rangle &= \langle (\mathbf{Q},\mathbf{P}), X\circ \mathbf{Q}\rangle \\ &= \int_S P_i(s) X^i(\mathbf{Q}(s)) d^k s.\end{aligned}$$

On the other hand, note that the right hand side of (150) (again with the index a suppressed, and with t suppressed as well), when paired with the Lie algebra element X is

$$\begin{aligned}\left\langle \int_S \mathbf{P}(s)\,\delta\,(\mathbf{x}-\mathbf{Q}(s))\,d^k s,\, X\right\rangle &= \int_{\mathbb{R}^n}\int_S \left(P_i(s)\,\delta\,(\mathbf{x}-\mathbf{Q}(s))\,d^k s\right) X^i(\mathbf{x}) d^n x \\ &= \int_S P_i(s) X^i(\mathbf{Q}(s) d^k s.\end{aligned}$$

This shows the expression given by (150) is equal to $\mathbf{J}$ and so the result is proved. □

Second proof As is standard (see, for example, Marsden and Ratiu [44]), one can characterize momentum maps by means of the following relation, required to hold for all functions F on $T^*\,\mathrm{Emb}(S,\mathbb{R}^n)$; that is, functions of $\mathbf{Q}$ and $\mathbf{P}$:

$$\{F, \langle \mathbf{J}, \xi\rangle\} = \xi_P[F]. \tag{157}$$

In our case, we shall take $\mathbf{J}$ to be given by the solution ansatz and verify that it satisfies this relation. To do so, let $\xi \in \mathfrak{X}$ so that the left side of (157) becomes

$$\left\{F, \int_S P_i(s)\xi^i(\mathbf{Q}(s)) d^k s\right\} = \int_S \left[\frac{\delta F}{\delta Q^i}\xi^i(\mathbf{Q}(s)) - P_i(s)\frac{\delta F}{\delta P_j}\frac{\delta}{\delta Q^j}\xi^i(\mathbf{Q}(s))\right] d^k s.$$

On the other hand, one can directly compute from the definitions that the infinitesimal generator of the action on the space $T^*\,\mathrm{Emb}(S,\mathbb{R}^n)$ corresponding to the vector field

$\xi^i(\mathbf{x})\frac{\partial}{\partial Q^i}$ (a Lie algebra element), is given by

$$\delta\mathbf{Q} = \xi \circ \mathbf{Q}, \quad \delta\mathbf{P} = -P_i(s)\frac{\partial}{\partial\mathbf{Q}}\xi^{\,i}(\mathbf{Q}(s)),$$

(see Marsden and Ratiu [44, (12.1.14)]) which verifies that (157) holds.

An important element left out in this proof so far is that it does not make clear that the momentum map is *equivariant*, a condition needed for the momentum map to be Poisson. The first proof took care of this automatically since *momentum maps for cotangent lifted actions are always equivariant* and hence are Poisson.

Thus, to complete the second proof, we need to check directly that the momentum map is equivariant. Actually, we shall only check that it is infinitesimally invariant by showing that it is a Poisson map from $T^*\,\mathrm{Emb}(S,\mathbb{R}^n)$ to the space of $\mathbf{m}$'s (the dual of the Lie algebra of $\mathfrak{X}$) with its Lie–Poisson bracket. This sort of approach to characterize equivariant momentum maps is discussed in an interesting way by Weinstein [63].

The following direct computation shows that the singular solution momentum map (150) is Poisson. This is accomplished by using the canonical Poisson brackets for $\{\mathbf{P}\}, \{\mathbf{Q}\}$ and applying the chain rule to compute $\{m_i(\mathbf{x}), m_j(\mathbf{y})\}$, with notation $\delta\,'_k(\mathbf{y}) \equiv \partial\delta(\mathbf{y})/\partial y^k$.

We get

$$\begin{aligned}
\{m_i(\mathbf{x}),m_j(\mathbf{y})\} &= \Big\{\sum_{a=1}^N \int ds\; P_i^a(s,t)\delta\big(\mathbf{x}-\mathbf{Q}^a(s,t)\big), \sum_{b=1}^N \int ds'\, P_j^b(s',t)\,\delta\big(\mathbf{y}-\mathbf{Q}^b(s',t)\big)\Big\} \\
&= \sum_{a,b=1}^N \iint ds\,ds' \Big[\{P_i^a(s),\, P_j^b(s')\}\delta\big(\mathbf{x}-\mathbf{Q}^a(s)\big)\delta\big(\mathbf{y}-\mathbf{Q}^b(s')\big) \\
&\qquad - \{P_i^a(s),Q_k^b(s')\}\,P_j^b(s')\delta\big(\mathbf{x}-\mathbf{Q}^a(s)\big)\delta'_k\big(\mathbf{y}-\mathbf{Q}^b(s')\big) \\
&\qquad - \{Q_k^a(s),P_j^b(s')\}\,P_i^a(s)\delta'_k\big(\mathbf{x}-\mathbf{Q}^a(s)\big)\delta\big(\mathbf{y}-\mathbf{Q}^b(s')\big) \\
&\qquad + \{Q_k^a(s),Q_\ell^b(s')\}\,P_i^a(s)\,P_j^b(s')\delta'_k\big(\mathbf{x}-\mathbf{Q}^a(s)\big)\delta'_\ell\big(\mathbf{y}-\mathbf{Q}^b(s')\big)\Big].
\end{aligned}$$

Substituting the canonical Poisson bracket relations

$$\begin{aligned}
&\{P_i^a(s),\, P_j^b(s')\} = 0 \\
&\{Q_k^a(s),\, Q_\ell^b(s')\} = 0, \\
\text{and}\qquad &\{Q_k^a(s),\, P_j^b(s')\} = \delta^{ab}\delta_{kj}\delta(s-s')
\end{aligned}$$

into the preceding computation yields

$$\begin{aligned}\{m_i(\mathbf{x}),m_j(\mathbf{y})\}&=\Big\{\sum_{a=1}^{N}\int ds\,P_i^a(s,t)\,\delta(\mathbf{x}-\mathbf{Q}^a(s,t)),\sum_{b=1}^{N}\int ds'\,P_j^b(s',t)\,\delta(\mathbf{y}-\mathbf{Q}^b(s',t))\Big\}\\ &=\sum_{a=1}^{N}\int ds\,P_j^a(s)\delta(\mathbf{x}-\mathbf{Q}^a(s))\delta_i'(\mathbf{y}-\mathbf{Q}^a(s))\\ &\qquad-\sum_{a=1}^{N}\int ds\,P_i^a(s)\delta_j'(\mathbf{x}-\mathbf{Q}^a(s))\delta(\mathbf{y}-\mathbf{Q}^a(s))\\ &=-\left(m_j(\mathbf{x})\frac{\partial}{\partial x^i}+\frac{\partial}{\partial x^j}m_i(\mathbf{x})\right)\delta(\mathbf{x}-\mathbf{y}).\end{aligned}$$

Thus,

$$\{m_i(\mathbf{x}),m_j(\mathbf{y})\}=-\left(m_j(\mathbf{x})\frac{\partial}{\partial x^i}+\frac{\partial}{\partial x^j}m_i(\mathbf{x})\right)\delta(\mathbf{x}-\mathbf{y}),\tag{158}$$

which is readily checked to be the Lie–Poisson bracket on the space of $\mathbf{m}$'s, restricted to their singular support. This completes the second proof of theorem. □

Each of these proofs has shown the following basic fact.

Corollary 28.2 *The singular solution momentum map defined by the singular solution ansatz (150), namely,*

$$\mathbf{J}_{\mathrm{Sing}}\colon T^*\operatorname{Emb}(S,\mathbb{R}^n)\to\mathfrak{X}(\mathbb{R}^n)^*$$

is a Poisson map from the canonical Poisson structure on $T^*\operatorname{Emb}(S,\mathbb{R}^n)$ to the Lie–Poisson structure on $\mathfrak{X}(\mathbb{R}^n)^*$.

This is perhaps the most basic property of the singular solution momentum map. Some of its more sophisticated properties are outlined by Holm and Marsden [30].

Pulling back the equations Since the solution ansatz (150) has been shown in the preceding Corollary to be a Poisson map, the pull back of the Hamiltonian from $\mathfrak{X}^*$ to $T^*\operatorname{Emb}(S,\mathbb{R}^n)$ gives equations of motion on the latter space that project to the equations on $\mathfrak{X}^*$.

> Thus, the basic fact that the momentum map $\mathbf{J}_{\mathrm{Sing}}$ is Poisson explains why the functions $\mathbf{Q}^a(s,t)$ and $\mathbf{P}^a(s,t)$ satisfy canonical Hamiltonian equations.

Note that the coordinate $s \in \mathbb{R}^k$ that labels these functions is a "Lagrangian coordinate" in the sense that it does not evolve in time but rather labels the solution.

In terms of the pairing

$$\langle \cdot, \cdot \rangle \colon \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}, \tag{159}$$

between the Lie algebra $\mathfrak{g}$ (vector fields in $\mathbb{R}^n$) and its dual $\mathfrak{g}^*$ (one-form densities in $\mathbb{R}^n$), the following relation holds for measure-valued solutions under the momentum map (150),

$$\begin{aligned}
\langle \mathbf{m}, \mathbf{u} \rangle &= \int \mathbf{m} \cdot \mathbf{u} \, d^n\mathbf{x}, \quad L^2 \text{ pairing for } \mathbf{m}, \mathbf{u} \in \mathbb{R}^n, \\
&= \int\!\!\int \sum_{a,b=1}^{N} \big(\mathbf{P}^a(s,t) \cdot \mathbf{P}^b(s',t)\big) G\big(\mathbf{Q}^a(s,t) - \mathbf{Q}^b(s',t)\big) ds\, ds' \\
&= \int \sum_{a=1}^{N} \mathbf{P}^a(s,t) \cdot \frac{\partial \mathbf{Q}^a(s,t)}{\partial t} ds \\
&\equiv \langle\!\langle \mathbf{P}, \dot{\mathbf{Q}} \rangle\!\rangle,
\end{aligned} \tag{160}$$

which is the natural pairing between the points $(\mathbf{Q}, \mathbf{P}) \in T^* \operatorname{Emb}(S, \mathbb{R}^n)$ and $(\mathbf{Q}, \dot{\mathbf{Q}}) \in T \operatorname{Emb}(S, \mathbb{R}^n)$. This corresponds to preservation of the action of the Lagrangian $\ell[\mathbf{u}]$ under cotangent lift of $\operatorname{Diff}(\mathbb{R}^n)$.

The pull-back of the Hamiltonian $H[\mathbf{m}]$ defined on the dual of the Lie algebra $\mathfrak{g}^*$, to $T^* \operatorname{Emb}(S, \mathbb{R}^n)$ is easily seen to be consistent with what we had before:

$$H[\mathbf{m}] \equiv \tfrac{1}{2} \langle \mathbf{m}, G * \mathbf{m} \rangle = \tfrac{1}{2} \langle\!\langle \mathbf{P}, G * \mathbf{P} \rangle\!\rangle \equiv H_N[\mathbf{P}, \mathbf{Q}]. \tag{161}$$

In summary, in concert with the Poisson nature of the singular solution momentum map, we see that the singular solutions in terms of $\mathbf{Q}$ and $\mathbf{P}$ satisfy Hamiltonian equations and also define an invariant solution set for the EPDiff equations. In fact:

> This invariant solution set is a special coadjoint orbit for the diffeomorphism group, as we shall discuss in the next section.

29 The geometry of the momentum map

In this section we explore the geometry of the singular solution momentum map discussed earlier in a little more detail. The treatment is formal, in the sense that there are a number of technical issues in the infinite dimensional case that will be left open. We will mention a few of these as we proceed.

29.1 Coadjoint orbits

We claim that *the image of the singular solution momentum map is a coadjoint orbit in* $\mathfrak{X}^*$. This means that (modulo some issues of connectedness and smoothness, which we do not consider here) the solution ansatz given by (150) defines a coadjoint orbit in the space of all one-form densities, regarded as the dual of the Lie algebra of the diffeomorphism group. These coadjoint orbits should be thought of as singular orbits—that is, due to their special nature, they are not generic.

Recognizing them as coadjoint orbits is one way of gaining further insight into why the singular solutions form dynamically invariant sets—it is a general fact that coadjoint orbits in $\mathfrak{g}^*$ are *symplectic submanifolds* of the Lie–Poisson manifold $\mathfrak{g}^*$ (in our case $\mathfrak{X}(\mathbb{R}^n)^*$) and, correspondingly, are dynamically invariant for any Hamiltonian system on $\mathfrak{g}^*$.

The idea of the proof of our claim is simply this: whenever one has an equivariant momentum map $\mathbf{J}\colon P \to \mathfrak{g}^*$ for the action of a group G on a symplectic or Poisson manifold P, and that action is transitive, then the image of $\mathbf{J}$ is an orbit (or at least a piece of an orbit). This general result, due to Kostant, is stated more precisely by Marsden and Ratiu [44, Theorem 14.4.5]. Roughly speaking, the reason that transitivity holds in our case is because one can "move the images of the manifolds S around at will with arbitrary velocity fields" using diffeomorphisms of $\mathbb{R}^n$.

29.2 The momentum map $\mathbf{J}_S$ and the Kelvin circulation theorem

The momentum map $\mathbf{J}_{\text{Sing}}$ involves $\text{Diff}(\mathbb{R}^n)$, the left action of the diffeomorphism group on the space of embeddings $\text{Emb}(S, \mathbb{R}^n)$ by smooth maps of the target space $\mathbb{R}^n$, namely,

$$\text{Diff}(\mathbb{R}^n)\colon \mathbf{Q} \cdot \eta = \eta \circ \mathbf{Q}, \tag{162}$$

where, recall, $\mathbf{Q}\colon S \to \mathbb{R}^n$. As above, the cotangent bundle $T^* \text{Emb}(S, \mathbb{R}^n)$ is identified with the space of pairs of maps $(\mathbf{Q}, \mathbf{P})$, with $\mathbf{Q}\colon S \to \mathbb{R}^n$ and $\mathbf{P}\colon S \to T^*\mathbb{R}^n$.

However, there is another momentum map $\mathbf{J}_S$ associated with the *right action* of the diffeomorphism group of S on the embeddings $\text{Emb}(S, \mathbb{R}^n)$ by smooth maps of the "Lagrangian labels" S (fluid particle relabeling by $\eta\colon S \to S$). This action is given by

$$\text{Diff}(S)\colon \mathbf{Q} \cdot \eta = \mathbf{Q} \circ \eta. \tag{163}$$

The infinitesimal generator of this right action is

$$X_{\text{Emb}(S,\mathbb{R}^n)}(\mathbf{Q}) = \left.\frac{d}{dt}\right|_{t=0} \mathbf{Q} \circ \eta_t = T\mathbf{Q} \circ X. \tag{164}$$

where $X \in \mathfrak{X}$ is tangent to the curve η_t at $t = 0$. Thus, again taking $N = 1$ (so we suppress the index a) and also letting α_q in the momentum map formula (156) be the cotangent vector $(\mathbf{Q}, \mathbf{P})$, one computes $\mathbf{J}_S$:

$$\begin{aligned}
\langle \mathbf{J}_S(\mathbf{Q}, \mathbf{P}), X \rangle &= \langle (\mathbf{Q}, \mathbf{P}), T\mathbf{Q} \cdot X \rangle \\
&= \int_S P_i(s) \frac{\partial Q^i(s)}{\partial s^m} X^m(s)\, d^k s \\
&= \int_S X\Big(\mathbf{P}(s) \cdot d\mathbf{Q}(s)\Big)\, d^k s \\
&= \left(\int_S \mathbf{P}(s) \cdot d\mathbf{Q}(s) \otimes d^k s, X(s) \right) \\
&= \langle \mathbf{P} \cdot d\mathbf{Q}, X \rangle .
\end{aligned}$$

Consequently, the momentum map formula (156) yields

$$\mathbf{J}_S(\mathbf{Q}, \mathbf{P}) = \mathbf{P} \cdot d\mathbf{Q}, \tag{165}$$

with the indicated pairing of the one-form density $\mathbf{P} \cdot d\mathbf{Q}$ with the vector field X.

We have set things up so that the following is true.

Proposition 29.1 *The momentum map $\mathbf{J}_S$ is preserved by the evolution equations (152)–(153) for $\mathbf{Q}$ and $\mathbf{P}$.*

Proof It is enough to notice that the Hamiltonian H_N in equation (154) is invariant under the cotangent lift of the action of $\mathrm{Diff}(S)$; it merely amounts to the invariance of the integral over S under reparametrization; that is, the change of variables formula; keep in mind that $\mathbf{P}$ includes a density factor. □

Remark 29.2

- This result is similar to the Kelvin–Noether theorem for circulation Γ of an ideal fluid, which may be written as $\Gamma = \oint_{c(s)} D(s)^{-1} \mathbf{P}(s) \cdot d\mathbf{Q}(s)$ for each Lagrangian circuit $c(s)$, where D is the mass density and $\mathbf{P}$ is again the canonical momentum density. This similarity should come as no surprise, because the Kelvin–Noether theorem for ideal fluids arises from invariance of Hamilton's principle under fluid parcel relabeling by the *same* right action of the diffeomorphism group, as in (163).
- Note that, being an equivariant momentum map, the map $\mathbf{J}_S$, as with $\mathbf{J}_{\mathrm{Sing}}$, is also a Poisson map. That is, substituting the canonical Poisson bracket into

relation (165); that is, the relation $\mathbf{M}(\mathbf{x}) = \sum_i P_i(\mathbf{x})\nabla Q^i(\mathbf{x})$ yields the Lie–Poisson bracket on the space of $\mathbf{M}$'s. We use the different notations $\mathbf{m}$ and $\mathbf{M}$ because these quantities are analogous to the body and spatial angular momentum for rigid body mechanics. In fact, the quantity $\mathbf{m}$ given by the solution Ansatz; specifically, $\mathbf{m} = \mathbf{J}_{\mathrm{Sing}}(\mathbf{Q}, \mathbf{P})$ gives the singular solutions of the EPDiff equations, while $\mathbf{M}(\mathbf{x}) = \mathbf{J}_S(\mathbf{Q}, \mathbf{P}) = \sum_i P_i(\mathbf{x})\nabla Q^i(\mathbf{x})$ is a conserved quantity.

- In the language of fluid mechanics, the expression of $\mathbf{m}$ in terms of $(\mathbf{Q}, \mathbf{P})$ is an example of a *Clebsch representation*, which expresses the solution of the EPDiff equations in terms of canonical variables that evolve by standard canonical Hamilton equations. This has been known in the case of fluid mechanics for more than 100 years. For modern discussions of the Clebsch representation for ideal fluids, see, for example, Holm and Kupershmidt [28] and Marsden and Weinstein [49].

- One more remark is in order; namely the special case in which $S = M$ is of course allowed. In this case, $\mathbf{Q}$ corresponds to the map η itself and $\mathbf{P}$ just corresponds to its conjugate momentum. The quantity $\mathbf{m}$ corresponds to the spatial (dynamic) momentum density (that is, right translation of $\mathbf{P}$ to the identity), while $\mathbf{M}$ corresponds to the conserved "body" momentum density (that is, left translation of $\mathbf{P}$ to the identity).

29.3 Brief summary

$\mathrm{Emb}(S, \mathbb{R}^n)$ admits two group actions. These are: the group $\mathrm{Diff}(S)$ of diffeomorphisms of S, which acts by composition on the *right*; and the group $\mathrm{Diff}(\mathbb{R}^n)$ which acts by composition on the *left*. The group $\mathrm{Diff}(\mathbb{R}^n)$ acting from the left produces the singular solution momentum map, $\mathbf{J}_{\mathrm{Sing}}$. The action of $\mathrm{Diff}(S)$ from the right produces the conserved momentum map $\mathbf{J}_S\colon T^*\mathrm{Emb}(S, \mathbb{R}^n) \to \mathfrak{X}(S)^*$. We now assemble both momentum maps into one figure as follows:

$$
\begin{tikzcd}
 & T^*\mathrm{Emb}(S, M) \arrow[dl, "\mathbf{J}_{\mathrm{Sing}}"'] \arrow[dr, "\mathbf{J}_S"] & \\
\mathfrak{X}(M)^* & & \mathfrak{X}(S)^*
\end{tikzcd}
$$

30 The Euler–Poincaré framework: fluids à la Holm, Marsden and Ratiu [31]

Almost all fluid models of interest admit the following general assumptions. These assumptions form the basis of the Euler–Poincaré theorem for Continua that we shall state later in this section, after introducing the notation necessary for dealing geometrically with the reduction of Hamilton's Principle from the material (or Lagrangian) picture of fluid dynamics, to the spatial (or Eulerian) picture. This theorem was first stated and proved by Holm, Marsden and Ratiu [31], to which we refer for additional details, as well as for abstract definitions and proofs.

Basic assumptions underlying the Euler–Poincaré theorem for continua

- There is a *right* representation of a Lie group G on the vector space V and G acts in the natural way on the *right* on $TG \times V^*$: $(U_g, a)h = (U_g h, ah)$.
- The Lagrangian function L: $TG \times V^* \to \mathbb{R}$ is right G–invariant under the isotropy group of $a_0 \in V^*$.[13]
- In particular, if $a_0 \in V^*$, define the Lagrangian L_{a_0}: $TG \to \mathbb{R}$ by $L_{a_0}(U_g) = L(U_g, a_0)$. Then L_{a_0} is right invariant under the lift to TG of the right action of G_{a_0} on G, where G_{a_0} is the isotropy group of a_0.
- Right G–invariance of L permits one to define the Lagrangian on the Lie algebra $\mathfrak{g}$ of the group G. Namely, ℓ: $\mathfrak{g} \times V^* \to \mathbb{R}$ is defined by,

  $$\ell(u, a) = L\big(U_g g^{-1}(t), a_0 g^{-1}(t)\big) = L(U_g, a_0),$$

  where $u = U_g g^{-1}(t)$ and $a = a_0 g^{-1}(t)$. Conversely, this relation defines for any ℓ: $\mathfrak{g} \times V^* \to \mathbb{R}$ a right G–invariant function L: $TG \times V^* \to \mathbb{R}$.
- For a curve $g(t) \in G$, let $u(t) := \dot{g}(t) g(t)^{-1}$ and define the curve $a(t)$ as the unique solution of the linear differential equation with time dependent coefficients $\dot{a}(t) = -a(t)u(t)$, where the action of an element of the Lie algebra $u \in \mathfrak{g}$ on an advected quantity $a \in V^*$ is denoted by concatenation from the right. The solution with initial condition $a(0) = a_0 \in V^*$ can be written as $a(t) = a_0 g(t)^{-1}$.

Notation for reduction of Hamilton's Principle by symmetries

- Let $\mathfrak{g}(\mathcal{D})$ denote the space of vector fields on $\mathcal{D}$ of some fixed differentiability class. These vector fields are endowed with the *Lie bracket* given in components

[13] For fluid dynamics, right G–invariance of the Lagrangian function L is traditionally called "particle relabeling symmetry."

by (summing on repeated indices)

$$[\mathbf{u},\mathbf{v}]^i = u^j\frac{\partial v^i}{\partial x^j} - v^j\frac{\partial u^i}{\partial x^j}. \tag{166}$$

The notation $\mathrm{ad}_{\mathbf{u}}\mathbf{v} := [\mathbf{u},\mathbf{v}]$ formally denotes the adjoint action of the *right* Lie algebra of $\mathrm{Diff}(\mathcal{D})$ on itself.

- Identify the Lie algebra of vector fields $\mathfrak{g}$ with its dual $\mathfrak{g}^*$ by using the L^2 pairing

$$\langle \mathbf{u},\mathbf{v}\rangle = \int_{\mathcal{D}} \mathbf{u}\cdot\mathbf{v}\, dV. \tag{167}$$

- Let $\mathfrak{g}(\mathcal{D})^*$ denote the geometric dual space of $\mathfrak{g}(\mathcal{D})$, that is, $\mathfrak{g}(\mathcal{D})^* := \Lambda^1(\mathcal{D})\otimes \mathrm{Den}(\mathcal{D})$. This is the space of one–form densities on $\mathcal{D}$. If $\mathbf{m}\otimes dV \in \Lambda^1(\mathcal{D})\otimes \mathrm{Den}(\mathcal{D})$, then the pairing of $\mathbf{m}\otimes dV$ with $\mathbf{u}\in\mathfrak{g}(\mathcal{D})$ is given by the L^2 pairing,

$$\langle \mathbf{m}\otimes dV, \mathbf{u}\rangle = \int_{\mathcal{D}} \mathbf{m}\cdot\mathbf{u}\, dV \tag{168}$$

where $\mathbf{m}\cdot\mathbf{u}$ is the standard contraction of a one–form $\mathbf{m}$ with a vector field $\mathbf{u}$.

- For $\mathbf{u}\in\mathfrak{g}(\mathcal{D})$ and $\mathbf{m}\otimes dV \in \mathfrak{g}(\mathcal{D})^*$, the dual of the adjoint representation is defined by

$$\langle \mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV), \mathbf{v}\rangle = -\int_{\mathcal{D}} \mathbf{m}\cdot \mathrm{ad}_{\mathbf{u}}\mathbf{v}\, dV = -\int_{\mathcal{D}} \mathbf{m}\cdot[\mathbf{u},\mathbf{v}]\, dV \tag{169}$$

and its expression is

$$\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV) = (\pounds_{\mathbf{u}}\mathbf{m} + (\mathrm{div}_{dV}\,\mathbf{u})\mathbf{m})\otimes dV = \pounds_{\mathbf{u}}(\mathbf{m}\otimes dV), \tag{170}$$

where $\mathrm{div}_{dV}\,\mathbf{u}$ is the divergence of $\mathbf{u}$ relative to the measure dV, that is, $\pounds_{\mathbf{u}} dV = (\mathrm{div}_{dV}\,\mathbf{u})dV$. Hence, $\mathrm{ad}^*_{\mathbf{u}}$ coincides with the Lie-derivative $\pounds_{\mathbf{u}}$ for one-form densities.

- If $\mathbf{u} = u^j\,\partial/\partial x^j$, $\mathbf{m} = m_i dx^i$, then the one–form factor in the preceding formula for $\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV)$ has the *coordinate expression*

$$\left(\mathrm{ad}^*_{\mathbf{u}}\,\mathbf{m}\right)_i dx^i = \left(u^j\frac{\partial m_i}{\partial x^j} + m_j\frac{\partial u^j}{\partial x^i} + (\mathrm{div}_{dV}\,\mathbf{u})m_i\right) dx^i \tag{171}$$

$$= \left(\frac{\partial}{\partial x^j}(u^j m_i) + m_j\frac{\partial u^j}{\partial x^i}\right) dx^i. \tag{172}$$

The last equality assumes that the divergence is taken relative to the standard measure $dV = d^n\mathbf{x}$ in $\mathbb{R}^n$. (On a Riemannian manifold the metric divergence needs to be used.)

Conventions and terminology in continuum mechanics Throughout the rest of the lecture notes, we shall follow Holm, Marsden and Ratiu [31] in using the conventions and terminology for the standard quantities in continuum mechanics.

Definition 30.1 Elements of $\mathcal{D}$ representing the material particles of the system are denoted by X; their coordinates X^A, $A = 1, \dots, n$ may thus be regarded as the *particle labels*.

- A *configuration*, which we typically denote by η, or g, is an element of $\mathrm{Diff}(\mathcal{D})$.
- A *motion*, denoted as η_t or alternatively as $g(t)$, is a time dependent curve in $\mathrm{Diff}(\mathcal{D})$.

Definition 30.2 The *Lagrangian*, or *material velocity* $\mathbf{U}(X,t)$ of the continuum along the motion η_t or $g(t)$ is defined by taking the time derivative of the motion keeping the particle labels X fixed:

$$\mathbf{U}(X,t) := \frac{d\eta_t(X)}{dt} := \frac{\partial}{\partial t}\bigg|_X \eta_t(X) := \dot{g}(t)\cdot X.$$

These are convenient shorthand notations for the time derivative at fixed Lagrangian coordinate X.

Consistent with this definition of material velocity, the tangent space to $\mathrm{Diff}(\mathcal{D})$ at $\eta \in \mathrm{Diff}(\mathcal{D})$ is given by

$$T_\eta \,\mathrm{Diff}(\mathcal{D}) = \{\mathbf{U}_\eta \colon \mathcal{D} \to T\mathcal{D} \mid \mathbf{U}_\eta(X) \in T_{\eta(X)}\mathcal{D}\}.$$

Elements of $T_\eta \,\mathrm{Diff}(\mathcal{D})$ are usually thought of as vector fields on $\mathcal{D}$ covering η. The tangent lift of right translations on $T\,\mathrm{Diff}(\mathcal{D})$ by $\varphi \in \mathrm{Diff}(\mathcal{D})$ is given by

$$\mathbf{U}_\eta\varphi := T_\eta R_\varphi(\mathbf{U}_\eta) = \mathbf{U}_\eta \circ \varphi.$$

Definition 30.3 During a motion η_t or $g(t)$, the particle labeled by X describes a path in $\mathcal{D}$, whose points

$$x(X,t) := \eta_t(X) := g(t)\cdot X,$$

are called the *Eulerian* or *spatial points* of this path, which is also called the *Lagrangian trajectory*, because a Lagrangian fluid parcel follows this path in space. The derivative

$\mathbf{u}(x,t)$ of this path, evaluated at fixed Eulerian point x, is called the *Eulerian* or *spatial velocity* of the system:

$$\mathbf{u}(x,t) := \mathbf{u}(\eta_t(X),t) := \mathbf{U}(X,t) := \frac{\partial}{\partial t}\bigg|_X \eta_t(X) := \dot{g}(t)\cdot X := \dot{g}(t)g^{-1}(t)\cdot x.$$

Thus the Eulerian velocity $\mathbf{u}$ is a time dependent vector field on $\mathcal{D}$, denoted as $\mathbf{u}_t \in \mathfrak{g}(\mathcal{D})$, where $\mathbf{u}_t(x) := \mathbf{u}(x,t)$. We also have the fundamental relationships

$$\mathbf{U}_t = \mathbf{u}_t \circ \eta_t \quad \text{and} \quad \mathbf{u}_t = \dot{g}(t)g^{-1}(t),$$

where we denote $\mathbf{U}_t(X) := \mathbf{U}(X,t)$.

Definition 30.4 The *representation space* V^* of $\mathrm{Diff}(\mathcal{D})$ in continuum mechanics is often some subspace of the tensor field densities on $\mathcal{D}$, denoted as $\mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, and the representation is given by pull back. It is thus a *right* representation of $\mathrm{Diff}(\mathcal{D})$ on $\mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$. The right action of the Lie algebra $\mathfrak{g}(\mathcal{D})$ on V^* is denoted as *concatenation from the right*. That is, we denote

$$a\mathbf{u} := \pounds_{\mathbf{u}} a,$$

which is the Lie derivative of the tensor field density a along the vector field $\mathbf{u}$.

Definition 30.5 The *Lagrangian of a continuum mechanical system* is a function

$$L\colon T\,\mathrm{Diff}(\mathcal{D}) \times V^* \to \mathbb{R},$$

which is right invariant relative to the tangent lift of right translation of $\mathrm{Diff}(\mathcal{D})$ on itself and pull back on the tensor field densities. Invariance of the Lagrangian L induces a function $\ell\colon \mathfrak{g}(\mathcal{D}) \times V^* \to \mathbb{R}$ given by

$$\ell(\mathbf{u},a) = L(\mathbf{u}\circ\eta, \eta^* a) = L(\mathbf{U}, a_0),$$

where $\mathbf{u} \in \mathfrak{g}(\mathcal{D})$ and $a \in V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, and where $\eta^* a$ denotes the pull back of a by the diffeomorphism η and $\mathbf{u}$ is the Eulerian velocity. That is,

$$\mathbf{U} = \mathbf{u}\circ\eta \quad \text{and} \quad a_0 = \eta^* a. \tag{173}$$

The evolution of a is by right action, given by the equation

$$\dot{a} = -\pounds_{\mathbf{u}}\, a = -a\mathbf{u}. \tag{174}$$

The solution of this equation, for the initial condition a_0, is

$$a(t) = \eta_{t*} a_0 = a_0 g^{-1}(t), \tag{175}$$

where the lower star denotes the push forward operation and η_t is the flow of $\mathbf{u} = \dot{g}g^{-1}(t)$.

Definition 30.6 *Advected Eulerian quantities* are defined in continuum mechanics to be those variables which are Lie transported by the flow of the Eulerian velocity field. Using this standard terminology, equation (174), or its solution (175) states that the tensor field density $a(t)$ (which may include mass density and other Eulerian quantities) is advected.

Remark 30.7 (Dual tensors) As we mentioned, typically $V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$ for continuum mechanics. On a general manifold, tensors of a given type have natural duals. For example, symmetric covariant tensors are dual to symmetric contravariant tensor densities, the pairing being given by the integration of the natural contraction of these tensors. Likewise, k–forms are naturally dual to $(n-k)$–forms, the pairing being given by taking the integral of their wedge product.

Definition 30.8 The *diamond operation* $\diamond$ between elements of V and V^* produces an element of the dual Lie algebra $\mathfrak{g}(\mathcal{D})^*$ and is defined as

(176) $$\langle b \diamond a, \mathbf{w} \rangle = -\int_{\mathcal{D}} b \cdot \pounds_{\mathbf{w}}\, a\,,$$

where $b \cdot \pounds_{\mathbf{w}}\, a$ denotes the contraction, as described above, of elements of V and elements of V^* and $\mathbf{w} \in \mathfrak{g}(\mathcal{D})$. (These operations do *not* depend on a Riemannian structure.)

For a path $\eta_t \in \mathrm{Diff}(\mathcal{D})$, let $\mathbf{u}(x,t)$ be its Eulerian velocity and consider the curve $a(t)$ with initial condition a_0 given by the equation

(177) $$\dot{a} + \pounds_{\mathbf{u}} a = 0.$$

Let the Lagrangian $L_{a_0}(\mathbf{U}) := L(\mathbf{U}, a_0)$ be right-invariant under $\mathrm{Diff}(\mathcal{D})$. We can now state the Euler–Poincaré Theorem for Continua of Holm, Marsden and Ratiu [31].

Theorem 30.9 (Euler–Poincaré Theorem for Continua) *Given a path η_t in* $\mathrm{Diff}(\mathcal{D})$ *with Lagrangian velocity* $\mathbf{U}$ *and Eulerian velocity* $\mathbf{u}$, *the following are equivalent:*

(i) *Hamilton's variational principle*

(178) $$\delta \int_{t_1}^{t_2} L\left(X, \mathbf{U}_t(X), a_0(X)\right)\, dt = 0$$

holds, for variations $\delta\eta_t$ vanishing at the endpoints.

(ii) *η_t satisfies the Euler–Lagrange equations for L_{a_0} on* $\mathrm{Diff}(\mathcal{D})$.

(iii) *The constrained variational principle in Eulerian coordinates*

$$\delta \int_{t_1}^{t_2} \ell(\mathbf{u}, a)\, dt = 0 \tag{179}$$

holds on $\mathfrak{g}(\mathcal{D}) \times V^*$*, using variations of the form*

$$\delta \mathbf{u} = \frac{\partial \mathbf{w}}{\partial t} + [\mathbf{u}, \mathbf{w}] = \frac{\partial \mathbf{w}}{\partial t} + \mathrm{ad}_{\mathbf{u}} \mathbf{w}, \qquad \delta a = -\pounds_{\mathbf{w}}\, a, \tag{180}$$

where $\mathbf{w}_t = \delta\eta_t \circ \eta_t^{-1}$ *vanishes at the endpoints.*

(iv) *The Euler–Poincaré equations for continua*

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{u}} = -\,\mathrm{ad}^*_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a = -\pounds_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a, \tag{181}$$

hold, with auxiliary equations $(\partial_t + \pounds_{\mathbf{u}})a = 0$ *for each advected quantity* $a(t)$*. The* $\diamond$ *operation defined in (176) needs to be determined on a case by case basis, depending on the nature of the tensor* $a(t)$*. The variation* $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ *is a one–form density and we have used relation (170) in the last step of equation (181).*

We refer to Holm, Marsden and Ratiu [31] for the proof of this theorem in the abstract setting. We shall see some of the features of this result in the concrete setting of continuum mechanics shortly.

Discussion of the Euler–Poincaré equations

The following string of equalities shows *directly* that (iii) is equivalent to (iv):

$$\begin{aligned} 0 = \delta \int_{t_1}^{t_2} l(\mathbf{u}, a) dt &= \int_{t_1}^{t_2} \left(\frac{\delta l}{\delta \mathbf{u}} \cdot \delta \mathbf{u} + \frac{\delta l}{\delta a} \cdot \delta a \right) dt \\ &= \int_{t_1}^{t_2} \left[\frac{\delta l}{\delta \mathbf{u}} \cdot \left(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\, \mathbf{w} \right) - \frac{\delta l}{\delta a} \cdot \pounds_{\mathbf{w}}\, a \right] dt \\ &= \int_{t_1}^{t_2} \mathbf{w} \cdot \left[-\frac{\partial}{\partial t} \frac{\delta l}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}} \frac{\delta l}{\delta \mathbf{u}} + \frac{\delta l}{\delta a} \diamond a \right] dt. \end{aligned} \tag{182}$$

The rest of the proof follows essentially the same track as the proof of the pure Euler–Poincaré theorem, modulo slight changes to accommodate the advected quantities.

In the absence of dissipation, most Eulerian fluid equations[14] can be written in the EP form in equation (181),

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}}\frac{\delta \ell}{\delta \mathbf{u}} = \frac{\delta \ell}{\delta a}\diamond a, \quad \text{with} \quad \big(\partial_t + \pounds_{\mathbf{u}}\big)a = 0. \tag{183}$$

Equation (183) is *Newton's Law*: The Eulerian time derivative of the momentum density $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ (a one-form density dual to the velocity $\mathbf{u}$) is equal to the force density $(\delta\ell/\delta a)\diamond a$, with the $\diamond$ operation defined in (176). Thus, Newton's Law is written in the Eulerian fluid representation as[15]

$$\frac{d}{dt}\Big|_{\mathrm{Lag}}\mathbf{m} := \big(\partial_t + \pounds_{\mathbf{u}}\big)\mathbf{m} = \frac{\delta \ell}{\delta a}\diamond a, \quad \text{with} \quad \frac{d}{dt}\Big|_{\mathrm{Lag}} a := \big(\partial_t + \pounds_{\mathbf{u}}\big)a = 0. \tag{184}$$

- The left side of the EP equation in (184) describes the fluid's dynamics due to its kinetic energy. A fluid's kinetic energy typically defines a norm for the Eulerian fluid velocity, $KE = \frac{1}{2}\|\mathbf{u}\|^2$. The left side of the EP equation is the *geodesic* part of its evolution, with respect to this norm. See Arnold and Khesin [4] for discussions of this interpretation of ideal incompressible flow and references to the literature. However, in a gravitational field, for example, there will also be dynamics due to potential energy. And this dynamics will by governed by the right side of the EP equation.
- The right side of the EP equation in (184) modifies the geodesic motion. Naturally, the right side of the EP equation is also a geometrical quantity. The diamond operation $\diamond$ represents the dual of the Lie algebra action of vectors fields on the tensor a. Here $\delta\ell/\delta a$ is the dual tensor, under the natural pairing (usually, L^2 pairing) $\langle\cdot,\cdot\rangle$ that is induced by the variational derivative of the Lagrangian $\ell(\mathbf{u},a)$. The diamond operation $\diamond$ is defined in terms of this pairing in (176). For the L^2 pairing, this is integration by parts of (minus) the Lie derivative in (176).

[14] Exceptions to this statement are certain multiphase fluids, and complex fluids with active internal degrees of freedom such as liquid crystals. These require a further extension, not discussed here.

[15] In coordinates, a one-form density takes the form $\mathbf{m}\cdot d\mathbf{x}\otimes dV$ and the EP equation (181) is given mnemonically by

$$\frac{d}{dt}\Big|_{\mathrm{Lag}}\big(\mathbf{m}\cdot d\mathbf{x}\otimes dV\big) = \underbrace{\frac{d\mathbf{m}}{dt}\Big|_{\mathrm{Lag}}\cdot d\mathbf{x}\otimes dV}_{\text{Advection}} + \underbrace{\mathbf{m}\cdot d\mathbf{u}\otimes dV}_{\text{Stretching}} + \underbrace{\mathbf{m}\cdot d\mathbf{x}\otimes(\nabla\cdot\mathbf{u})dV}_{\text{Expansion}} = \frac{\delta\ell}{\delta a}\diamond a$$

with $\frac{d}{dt}\big|_{\mathrm{Lag}} d\mathbf{x} := \big(\partial_t + \pounds_{\mathbf{u}}\big)d\mathbf{x} = d\mathbf{u} = \mathbf{u}_{,j}dx^j$, upon using commutation of Lie derivative and exterior derivative. Compare this formula with the definition of $\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV)$ in equation (171).

- The quantity a is typically a tensor (for example, a density, a scalar, or a differential form) and we shall sum over the various types of tensors a that are involved in the fluid description. The second equation in (184) states that each tensor a is carried along by the Eulerian fluid velocity $\mathbf{u}$. Thus, a is for fluid "attribute," and its Eulerian evolution is given by minus its Lie derivative, $-\pounds_{\mathbf{u}}a$. That is, a stands for the set of fluid attributes that each Lagrangian fluid parcel carries around (advects), such as its buoyancy, which is determined by its individual salt, or heat content, in ocean circulation.
- Many examples of how equation (184) arises in the dynamics of continuous media are given by Holm, Marsden and Ratiu [31]. The EP form of the Eulerian fluid description in (184) is analogous to the classical dynamics of rigid bodies (and tops, under gravity) in body coordinates. Rigid bodies and tops are also governed by Euler–Poincaré equations, as Poincaré showed in a two-page paper with no references, over a century ago [55]. For modern discussions of the EP theory, see, for example, Marsden and Ratiu [44], or Holm, Marsden and Ratiu [31].

Exercise 30.10 For what types of tensors a_0 can one recast the EP equations for continua (181) as geodesic motion, by using a version of the Kaluza–Klein construction?

30.1 Corollary of the EP theorem: the Kelvin–Noether circulation theorem

Corollary 30.11 (Kelvin–Noether Circulation Theorem) *Assume $\mathbf{u}(x,t)$ satisfies the Euler–Poincaré equations for continua:*

$$\frac{\partial}{\partial t}\left(\frac{\delta \ell}{\delta \mathbf{u}}\right) = -\pounds_{\mathbf{u}}\left(\frac{\delta \ell}{\delta \mathbf{u}}\right) + \frac{\delta \ell}{\delta a}\diamond a$$

and the quantity a satisfies the advection relation

$$\frac{\partial a}{\partial t} + \pounds_{\mathbf{u}} a = 0. \tag{185}$$

Let η_t be the flow of the Eulerian velocity field $\mathbf{u}$, that is, $\mathbf{u} = (d\eta_t/dt)\circ\eta_t^{-1}$. Define the advected fluid loop $\gamma_t := \eta_t\circ\gamma_0$ and the circulation map $I(t)$ by

$$I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}}. \tag{186}$$

In the circulation map $I(t)$ the advected mass density D_t satisfies the push forward relation $D_t = \eta_* D_0$. This implies the advection relation (185) with $a = D$, namely, *the continuity equation,*

$$\partial_t D + \operatorname{div} D\mathbf{u} = 0.$$

Then the map $I(t)$ *satisfies the* Kelvin circulation relation,

$$\frac{d}{dt} I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta a} \diamond a\,. \tag{187}$$

Both an abstract proof of the Kelvin–Noether Circulation Theorem and a proof tailored for the case of continuum mechanical systems are given in Holm, Marsden and Ratiu [31]. We provide a version of the latter below.

Proof First we change variables in the expression for $I(t)$:

$$I(t) = \oint_{\gamma_t} \frac{1}{D_t}\frac{\delta l}{\delta \mathbf{u}} = \oint_{\gamma_0} \eta_t^* \left[\frac{1}{D_t}\frac{\delta l}{\delta \mathbf{u}}\right] = \oint_{\gamma_0} \frac{1}{D_0}\eta_t^* \left[\frac{\delta l}{\delta \mathbf{u}}\right].$$

Next, we use the Lie derivative formula, namely

$$\frac{d}{dt}\left(\eta_t^* \alpha_t\right) = \eta_t^* \left(\frac{\partial}{\partial t}\alpha_t + \pounds_{\mathbf{u}}\alpha_t\right),$$

applied to a one–form density α_t. This formula gives

$$\begin{aligned}\frac{d}{dt} I(t) = \frac{d}{dt}\oint_{\gamma_0} \frac{1}{D_0}\eta_t^* \left[\frac{\delta l}{\delta \mathbf{u}}\right] &= \oint_{\gamma_0} \frac{1}{D_0}\frac{d}{dt}\left(\eta_t^* \left[\frac{\delta l}{\delta \mathbf{u}}\right]\right) \\ &= \oint_{\gamma_0} \frac{1}{D_0}\eta_t^* \left[\frac{\partial}{\partial t}\left(\frac{\delta l}{\delta \mathbf{u}}\right) + \pounds_{\mathbf{u}}\left(\frac{\delta l}{\delta \mathbf{u}}\right)\right].\end{aligned}$$

By the Euler–Poincaré equations (181), this becomes

$$\frac{d}{dt} I(t) = \oint_{\gamma_0} \frac{1}{D_0}\eta_t^* \left[\frac{\delta l}{\delta a}\diamond a\right] = \oint_{\gamma_t} \frac{1}{D_t}\left[\frac{\delta l}{\delta a}\diamond a\right],$$

again by the change of variables formula. □

Corollary 30.12 *Since the last expression holds for every loop* γ_t*, we may write it as*

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right)\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} = \frac{1}{D}\frac{\delta l}{\delta a}\diamond a. \tag{188}$$

Remark 30.13 The Kelvin–Noether theorem is called so here because its derivation relies on the invariance of the Lagrangian L under the particle relabeling symmetry, and Noether's theorem is associated with this symmetry. However, the result (187) is the *Kelvin circulation theorem*: the circulation integral $I(t)$ around any fluid loop (γ_t, moving with the velocity of the fluid parcels $\mathbf{u}$) is invariant under the fluid motion.

These two statements are equivalent. We note that *two velocities* appear in the integrand $I(t)$: the fluid velocity $\mathbf{u}$ and $D^{-1}\delta\ell/\delta\mathbf{u}$. The latter velocity is the momentum density $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ divided by the mass density D. These two velocities are the basic ingredients for performing modeling and analysis in any ideal fluid problem. One simply needs to put these ingredients together in the Euler–Poincaré theorem and its corollary, the Kelvin–Noether theorem.

31 Euler–Poincaré theorem and GFD (geophysical fluid dynamics)

31.1 Variational formulae in three dimensions

We compute explicit formulae for the variations δa in the cases that the set of tensors a is drawn from a set of scalar fields and densities on $\mathbb{R}^3$. We shall denote this symbolically by writing

$$a \in \{b, D\, d^3x\}. \tag{189}$$

We have seen that invariance of the set a in the Lagrangian picture under the dynamics of $\mathbf{u}$ implies in the Eulerian picture that

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) a = 0,$$

where $\pounds_{\mathbf{u}}$ denotes Lie derivative with respect to the velocity vector field $\mathbf{u}$. Hence, for a fluid dynamical Eulerian action $\mathfrak{S} = \int dt\ \ell(\mathbf{u}; b, D)$, the advected variables b and D satisfy the following Lie-derivative relations,

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) b = 0, \qquad \text{or} \qquad \frac{\partial b}{\partial t} = -\,\mathbf{u}\cdot\nabla\, b, \tag{190}$$

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) D\, d^3x = 0, \qquad \text{or} \qquad \frac{\partial D}{\partial t} = -\,\nabla\cdot(D\mathbf{u}). \tag{191}$$

In fluid dynamical applications, the advected Eulerian variables b and $D\, d^3x$ represent the buoyancy b (or specific entropy, for the compressible case) and volume element (or mass density) $D\, d^3x$, respectively. According to Theorem 30.9, equation (179), the variations of the tensor functions a at fixed $\mathbf{x}$ and t are also given by Lie derivatives, namely $\delta a = -\pounds_{\mathbf{w}}\, a$, or

$$\begin{aligned} \delta b &= -\pounds_{\mathbf{w}}\, b = -\mathbf{w}\cdot\nabla\, b, \\ \delta D\ d^3x &= -\pounds_{\mathbf{w}}\,(D\, d^3x) = -\nabla\cdot(D\mathbf{w})\ d^3x. \end{aligned} \tag{192}$$

Hence, Hamilton's principle (179) with this dependence yields

$$
\begin{aligned}
0 &= \delta \int dt\ \ell(\mathbf{u}; b, D) \\
&= \int dt \left[\frac{\delta \ell}{\delta \mathbf{u}} \cdot \delta \mathbf{u} + \frac{\delta \ell}{\delta b}\, \delta b + \frac{\delta \ell}{\delta D}\, \delta D \right] \\
&= \int dt \left[\frac{\delta \ell}{\delta \mathbf{u}} \cdot \Big(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\, \mathbf{w} \Big) - \frac{\delta \ell}{\delta b}\, \mathbf{w} \cdot \nabla\, b - \frac{\delta \ell}{\delta D}\, \Big(\nabla \cdot (D\mathbf{w}) \Big) \right] \\
&= \int dt\ \mathbf{w} \cdot \left[- \frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}}\, \frac{\delta \ell}{\delta \mathbf{u}} - \frac{\delta \ell}{\delta b}\, \nabla\, b + D\, \nabla \frac{\delta \ell}{\delta D} \right] \\
&= - \int dt\ \mathbf{w} \cdot \left[\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}} \Big) \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta b}\, \nabla\, b - D\, \nabla \frac{\delta \ell}{\delta D} \right],
\end{aligned}
\tag{193}
$$

where we have consistently dropped boundary terms arising from integrations by parts, by invoking natural boundary conditions. Specifically, we may impose $\hat{\mathbf{n}} \cdot \mathbf{w} = 0$ on the boundary, where $\hat{\mathbf{n}}$ is the boundary's outward unit normal vector and $\mathbf{w} = \delta \eta_t \circ \eta_t^{-1}$ vanishes at the endpoints.

31.2 Euler–Poincaré framework for GFD

The Euler–Poincaré equations for continua (181) may now be summarized in vector form for advected Eulerian variables a in the set (189). We adopt the notational convention of the circulation map I in equations (186) and (187) that a one form density can be made into a one form (no longer a density) by dividing it by the mass density D and we use the Lie-derivative relation for the continuity equation $(\partial/\partial t + \pounds_{\mathbf{u}}) D d^3 x = 0$. Then, the Euclidean components of the Euler–Poincaré equations for continua in equation (193) are expressed in Kelvin theorem form (188) with a slight abuse of notation as

$$
\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}} \Big) \Big(\frac{1}{D} \frac{\delta \ell}{\delta \mathbf{u}} \cdot d\mathbf{x} \Big) \ + \ \frac{1}{D} \frac{\delta \ell}{\delta b} \nabla b \cdot d\mathbf{x} - \nabla \Big(\frac{\delta \ell}{\delta D} \Big) \cdot d\mathbf{x} = 0,
\tag{194}
$$

in which the variational derivatives of the Lagrangian ℓ are to be computed according to the usual physical conventions, that is, as Fréchet derivatives. Formula (194) is the Kelvin–Noether form of the equation of motion for ideal continua. Hence, we have the explicit Kelvin theorem expression, cf. equations (186) and (187),

$$
\frac{d}{dt} \oint_{\gamma_t(\mathbf{u})} \frac{1}{D} \frac{\delta \ell}{\delta \mathbf{u}} \cdot d\mathbf{x} = - \oint_{\gamma_t(\mathbf{u})} \frac{1}{D} \frac{\delta \ell}{\delta b} \nabla b \cdot d\mathbf{x}\,,
\tag{195}
$$

where the curve $\gamma_t(\mathbf{u})$ moves with the fluid velocity $\mathbf{u}$. Then, by Stokes' theorem, the Euler equations generate circulation of $\mathbf{v} := (D^{-1} \delta l / \delta \mathbf{u})$ whenever the gradients ∇b

and $\nabla(D^{-1}\delta l/\delta b)$ are not collinear. The corresponding *conservation of potential vorticity* q on fluid parcels is given by

$$\frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q = 0, \quad \text{where} \quad q = \frac{1}{D}\nabla b\cdot\operatorname{curl}\left(\frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}}\right). \tag{196}$$

This is also called *PV convection*. Equations (194)–(196) embody most of the panoply of equations for GFD. The vector form of equation (194) is,

$$\underbrace{\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\left(\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}}\right) + \frac{1}{D}\frac{\delta l}{\delta u^j}\nabla u^j}_{\text{Geodesic Nonlinearity: Kinetic energy}} = \underbrace{\nabla\frac{\delta l}{\delta D} - \frac{1}{D}\frac{\delta l}{\delta b}\nabla b}_{\text{Potential energy}} \tag{197}$$

In geophysical applications, the Eulerian variable D represents the frozen-in volume element and b is the buoyancy. In this case, *Kelvin's theorem* is

$$\frac{dI}{dt} = \int\int_{S(t)} \nabla\left(\frac{1}{D}\frac{\delta l}{\delta b}\right)\times\nabla b\cdot d\mathbf{S},$$

with circulation integral

$$I = \oint_{\gamma(t)} \frac{1}{D}\frac{\delta l}{\delta \mathbf{u}}\cdot d\mathbf{x}.$$

31.3 Euler's equations for a rotating stratified ideal incompressible fluid

The Lagrangian In the Eulerian velocity representation, we consider Hamilton's principle for fluid motion in a three dimensional domain with action functional $S = \int l\, dt$ and Lagrangian $l(\mathbf{u}, b, D)$ given by

$$l(\mathbf{u}, b, D) = \int \rho_0 D(1+b)\left(\tfrac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - gz\right) - p(D-1)\, d^3x, \tag{198}$$

where $\rho_{tot} = \rho_0 D(1+b)$ is the total mass density, ρ_0 is a dimensional constant and $\mathbf{R}$ is a given function of $\mathbf{x}$. This variations at fixed $\mathbf{x}$ and t of this Lagrangian are the following,

$$\begin{aligned} \frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} &= \rho_0(1+b)(\mathbf{u}+\mathbf{R}), & \frac{\delta l}{\delta b} &= \rho_0 D\left(\tfrac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\right), \\ \frac{\delta l}{\delta D} &= \rho_0(1+b)\left(\tfrac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\right) - p, & \frac{\delta l}{\delta p} &= -(D-1). \end{aligned} \tag{199}$$

Hence, from the Euclidean component formula (197) for Hamilton principles of this type and the fundamental vector identity,

$$(\mathbf{b}\cdot\nabla)\mathbf{a} + a_j\nabla b^j = -\,\mathbf{b}\times(\nabla\times\mathbf{a}) + \nabla(\mathbf{b}\cdot\mathbf{a}), \tag{200}$$

we find the motion equation for an Euler fluid in three dimensions,

$$\frac{d\mathbf{u}}{dt} - \mathbf{u}\times\operatorname{curl}\mathbf{R} + g\hat{\mathbf{z}} + \frac{1}{\rho_0(1+b)}\nabla p = 0, \tag{201}$$

where $\operatorname{curl}\mathbf{R} = 2\boldsymbol{\Omega}(\mathbf{x})$ is the Coriolis parameter (that is, twice the local angular rotation frequency). In writing this equation, we have used advection of buoyancy,

$$\frac{\partial b}{\partial t} + \mathbf{u}\cdot\nabla b = 0,$$

from equation (190). The pressure p is determined by requiring preservation of the constraint $D = 1$, for which the continuity equation (191) implies $\operatorname{div}\mathbf{u} = 0$. The Euler motion equation (201) is Newton's Law for the acceleration of a fluid due to three forces: Coriolis, gravity and pressure gradient. The dynamic balances among these three forces produce the many circulatory flows of geophysical fluid dynamics. The *conservation of potential vorticity* q on fluid parcels for these Euler GFD flows is given by

$$\frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q = 0, \quad \text{where, on using } D = 1, \quad q = \nabla b\cdot\operatorname{curl}\big(\mathbf{u}+\mathbf{R}\big). \tag{202}$$

Exercise 31.1 (Semidirect-product Lie–Poisson bracket for compressible ideal fluids)

(1) Compute the Legendre transform for the Lagrangian,

$$l(\mathbf{u}, b, D)\colon \mathfrak{X}\times\Lambda^0\times\Lambda^3 \mapsto \mathbb{R}$$

whose advected variables satisfy the auxiliary equations,

$$\frac{\partial b}{\partial t} = -\,\mathbf{u}\cdot\nabla\, b, \qquad \frac{\partial D}{\partial t} = -\,\nabla\cdot(D\mathbf{u}).$$

(2) Compute the Hamiltonian, assuming the Legendre transform is a linear invertible operator on the velocity $\mathbf{u}$. For definiteness in computing the Hamiltonian, assume the Lagrangian is given by

$$l(\mathbf{u}, b, D) = \int D\big(\tfrac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - e(D,b)\big)\, d^{\,3}x, \tag{203}$$

with prescribed function $\mathbf{R}(\mathbf{x})$ and specific internal energy $e(D,b)$ satisfying the First Law of Thermodynamics,

$$de = \frac{p}{D^2}dD + T\,db,$$

where p is pressure, T temperature.

(3) Find the semidirect-product Lie–Poisson bracket for the Hamiltonian formulation of these equations.

(4) Does this Lie–Poisson bracket have Casimirs? If so, what are the corresponding symmetries and momentum maps?

32 Hamilton–Poincaré reduction and Lie–Poisson equations

In the Euler–Poincaré framework one starts with a Lagrangian defined on the tangent bundle of a Lie group G

$$L\colon TG \to \mathbb{R}$$

and the dynamics is given by Euler–Lagrange equations arising from the variational principle

$$\delta \int_{t_0}^{t_1} L(g, \dot{g}) dt = 0$$

The Lagrangian L is taken left/right invariant and because of this property one can *reduce* the problem obtaining a new system which is defined on the Lie algebra $\mathfrak{g}$ of G, obtaining a new set of equations, the Euler–Poincaré equations, arising from a reduced variational principle

$$\delta \int_{t_0}^{t_1} l(\xi) dt = 0$$

where $l(\xi)$ is the reduced lagrangian and $\xi \in \mathfrak{g}$.

Problem 32.1 Is there a similar procedure for Hamiltonian systems? More precisely: given a Hamiltonian function

$$H\colon T^*G \to \mathbb{R}$$

defined on the cotangent bundle T^*G, one wants to perform a similar procedure of reduction and derive the equations of motion on the dual of the Lie algebra $\mathfrak{g}^*$, provided the Hamiltonian is again left/right invariant.

Hamilton–Poincaré reduction gives a positive answer to this problem, in the context of variational principles as it is done in the Euler–Poincaré framework: we are going to explain how this procedure is performed.

More in general, we will also consider advected quantities belonging to a vector space V on which G acts, so that the Hamiltonian is written in this case as

$$H\colon T^*G \times V^* \to \mathbb{R}$$

(see Holm, Marsden and Ratiu [31; 32]). The space V is regarded here exactly the same as in the Euler–Poincaré theory.
The equations of motion, that is, Hamilton's equations, may be derived from the following variational principle

$$\delta \int_{t_0}^{t_1} \{\langle p(t), \dot{g}(t)\rangle - H_{a_0}(g(t), p(t))\}\, dt = 0$$

as it is well know from ordinary classical mechanics ($\dot{g}(t)$ has to be considered as the tangent vector to the curve $g(t)$, so that $\dot{g}(t) \in T_{g(t)}G$).

Problem 32.2 What happens if H_{a_0} is left/right invariant?

It turns out that in this case the whole function

$$F(g, \dot{g}, p) = \langle p, \dot{g}\rangle - H_{a_0}(g, p)$$

is also invariant. The proof is straightforward once the action is specified (from now on we consider only left invariance):

$$h\,(g, \dot{g}, p) = (hg, T_g L_h\, \dot{g}, T^*_{hg} L_{h^{-1}}\, p)$$

where $T_g L_h \colon T_g G \to T_{hg} G$ is the tangent of the left translation map $L_h\, g = hg \in G$ at the point g and $T^*_{hg} L_{h^{-1}} \colon T^*_g G \to T^*_{hg} G$ is the dual of the map $T_{hg} L_{h^{-1}} \colon T_{hg} G \to T_g G$.

We now check that

$$\begin{aligned}
\langle h\, p, h\, \dot{g}\rangle &= \langle T^*_{hg} L_{h^{-1}}\, p, T_g L_h\, \dot{g}\rangle \\
&= \langle p, T_{hg} L_{h^{-1}} \circ T_g L_h\, \dot{g}\rangle \\
&= \langle p, T_g (L_{h^{-1}} \circ L_h)\, \dot{g}\rangle = \langle p, \dot{g}\rangle
\end{aligned}$$

where the chain rule for the tangent map has been used. The same result holds for the right action.
Due to this invariance property, one can write the variational principle as

$$\delta \int_{t_0}^{t_1} \{\langle \mu, \xi\rangle - h(\mu, a)\}\, dt = 0$$

with

$$\mu(t) = g^{-1}(t)\, p(t) \in \mathfrak{g}^*, \qquad \xi(t) = g^{-1}(t)\, \dot{g}(t) \in \mathfrak{g}, \qquad a(t) = g^{-1}(t)\, a_0 \in V^*$$

In particular $a(t)$ is the solution of

$$\dot{a}(t) = -\xi(t)\, a_0.$$

where a Lie algebra action of $\mathfrak{g}$ on V^* is implicitly defined. In order to find the equations of motion one calculates the variations

$$\delta\int_{t_0}^{t_1}\{\langle\mu,\xi\rangle - h(\mu,a)\}\,dt = \int_{t_0}^{t_1}\left\{\langle\delta\mu,\xi\rangle + \langle\mu,\delta\xi\rangle - \left\langle\delta\mu,\frac{\delta h}{\delta\mu}\right\rangle - \left\langle\delta a,\frac{\delta h}{\delta a}\right\rangle\right\}\,dt$$

As in the Euler–Poincaré theorem, we use the following expressions for the variations

$$\delta\xi = \dot\eta + [\xi,\eta], \qquad \delta a = -\eta a$$

and using the definition of the diamond operator we find

$$\begin{aligned}
&\int_{t_0}^{t_1}\left\{\langle\delta\mu,\xi\rangle + \langle\mu,\delta\xi\rangle - \left\langle\delta\mu,\frac{\delta h}{\delta\mu}\right\rangle - \left\langle\delta a,\frac{\delta h}{\delta a}\right\rangle\right\}\,dt\\
&= \int_{t_0}^{t_1}\left\{\left\langle\delta\mu,\xi - \frac{\delta h}{\delta\mu}\right\rangle + \langle\mu,\dot\eta + \mathrm{ad}_\xi\eta\rangle + \left\langle\eta a,\frac{\delta h}{\delta a}\right\rangle\right\}\,dt\\
&= \int_{t_0}^{t_1}\left\{\left\langle\delta\mu,\xi - \frac{\delta h}{\delta\mu}\right\rangle + \langle -\dot\mu + \mathrm{ad}^*_\xi\mu,\eta\rangle - \left\langle\frac{\delta h}{\delta a}\diamond a,\eta\right\rangle\right\}\,dt
\end{aligned}$$

so that

$$\xi = \frac{\delta h}{\delta\mu}$$

and the equations of motion are

$$\dot\mu = \mathrm{ad}^*_\xi\mu - \frac{\delta h}{\delta a}\diamond a$$

together with

$$\dot a = -\frac{\delta h}{\delta\mu}\,a.$$

This equations of motion written on the dual Lie algebra $\mathfrak{g}$ are called *Lie–Poisson* equations. We have now proven the following:

Theorem 32.3 (Hamilton–Poincaré reduction theorem) *With the preceding notation, the following statements are equivalent:*

(1) *With a_0 held fixed, the variational principle*

$$\delta\int_{t_0}^{t_1}\{\langle p(t),\dot g(t)\rangle - H_{a_0}(g(t),p(t))\}\,dt = 0$$

holds, for variations $\delta g(t)$ of $g(t)$ vanishing at the endpoints.

(2) *$(g(t),p(t))$ satisfies Hamilton's equations for H_{a_0} on G.*

(3) *The constrained variational principle*

$$\delta \int_{t_0}^{t_1} \{\langle \mu(t), \xi(t)\rangle - h(\mu(t), a(t))\}\, dt = 0$$

holds for $\mathfrak{g} \times V^*$, *using variations of* ξ *and* a *of the form*

$$\delta\xi = \dot\eta + [\xi, \eta], \qquad \delta a = -\eta a$$

where $\eta(t) \in \mathfrak{g}$ *vanishes at the endpoints*

(4) *The* Lie–Poisson *equations hold on* $\mathfrak{g} \times V^*$

$$(\dot\mu, \dot a) = \left(\mathrm{ad}^*_\xi \mu - \frac{\delta h}{\delta a} \diamond a, -\frac{\delta h}{\delta \mu}\, a\right)$$

Remark 32.4 More exactly one should start with an invariant Hamiltonian defined on

$$T^*(G \times V) = T^*G \times V \times V^*$$

However, as mentioned by Holm, Marsden and Ratiu [31; 32], such an approach turns out to be equivalent to the treatment presented here.

Remark 32.5 (Legendre transform) Lie–Poisson equations may arise from the Euler–Poincaré setting by Legendre transform

$$\mu = \frac{\delta l}{\delta \xi}.$$

If this is a diffeomorphism, then the Hamilton–Poincaré theorem is equivalent to the Euler–Poincaré theorem.

Remark 32.6 (Lie–Poisson structure) One shows that $\mathfrak{g}^* \times V^*$ is a Poisson manifold:

$$\begin{aligned}
\dot F(\mu, a) &= \left\langle \dot\mu, \frac{\delta F}{\delta \mu}\right\rangle + \left\langle \dot a, \frac{\delta F}{\delta a}\right\rangle \\
&= \left\langle \mathrm{ad}^*_{\delta H/\delta\mu}\mu - \frac{\delta H}{\delta a} \diamond a, \frac{\delta F}{\delta \mu}\right\rangle - \left\langle \frac{\delta H}{\delta \mu} a, \frac{\delta F}{\delta a}\right\rangle \\
&= \left\langle \mu, \left[\frac{\delta H}{\delta \mu}, \frac{\delta F}{\delta \mu}\right]\right\rangle - \left\langle \frac{\delta H}{\delta a} \diamond a, \frac{\delta F}{\delta \mu}\right\rangle - \left\langle \frac{\delta H}{\delta \mu} a, \frac{\delta F}{\delta a}\right\rangle \\
&= -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\right\rangle - \left\langle \frac{\delta H}{\delta a} \diamond a, \right\rangle - \left\langle \frac{\delta H}{\delta \mu} a, \frac{\delta F}{\delta a}\right\rangle \\
&= -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\right\rangle - \left\langle a, \frac{\delta F}{\delta \mu}\frac{\delta H}{\delta a} - \frac{\delta H}{\delta \mu}\frac{\delta F}{\delta a}\right\rangle
\end{aligned}$$

In fact it can be easily shown that this structure

$$\{F,H\}(\mu,a) = -\left\langle \mu, \left[\frac{\delta F}{\delta\mu}, \frac{\delta H}{\delta\mu}\right]\right\rangle - \left\langle a, \frac{\delta F}{\delta\mu}\frac{\delta H}{\delta a} - \frac{\delta H}{\delta\mu}\frac{\delta F}{\delta a}\right\rangle$$

satisfies the definition of a Poisson structure. In particular one finds that *any dual Lie algebra* $\mathfrak{g}$ *is a Poisson manifold.*

Note this structure has been found during lectures for the simpler case without advected quantities.

Remark 32.7 (Right invariance) It can be shown that for a right invariant Hamiltonian one has

$$\{F,H\}(\mu,a) = +\left\langle \mu, \left[\frac{\delta F}{\delta\mu}, \frac{\delta H}{\delta\mu}\right]\right\rangle + \left\langle a, \frac{\delta F}{\delta\mu}\frac{\delta H}{\delta a} - \frac{\delta H}{\delta\mu}\frac{\delta F}{\delta a}\right\rangle$$

$$(\dot\mu, \dot a) = -\left(\mathrm{ad}^*_\xi \mu - \frac{\delta h}{\delta a}\diamond a, -\frac{\delta h}{\delta\mu}\, a\right)$$

with all signs changed respect to the case of left invariance.

33 Two applications

33.1 The Vlasov equation

In plasma physics a main topic is collisionless particle dynamics, whose main equation, the Vlasov equation, will be heuristically derived here. In this context a central role is held by the distribution function on phase space $f(\mathbf{q},\mathbf{p},t)$, basically expressing the particle density on phase space. Intended as a density one defines $F := f(\mathbf{q},\mathbf{p},t)\, d\mathbf{q}\, d\mathbf{p}$: because of the conservation of particles, one writes the continuity equation just as one does as in the context of fluid dynamics

$$\dot F + \nabla\cdot(\mathbf{u}\, F) = 0$$

where $\mathbf{u}$ is a "velocity" vector field on phase space, which is given by the single particle motion

$$\mathbf{u} = (\dot{\mathbf{q}}, \dot{\mathbf{p}}) \in \mathfrak{X}\big(T^*\mathbb{R}^N\big)$$

if we now assume that the generic single particle undergoes a Hamiltonian motion, the Hamiltonian function $h(\mathbf{q},\mathbf{p},t)$ can be introduced directly by means of the single particle Hamilton's equations

$$(\dot{\mathbf{q}}, \dot{\mathbf{p}}) = \left(\frac{\partial h}{\partial \mathbf{p}}, -\frac{\partial h}{\partial \mathbf{q}}\right)$$

which shows that $\mathbf{u}$ has zero divergence, assuming the Hessian of h is symmetric. Therefore, the Vlasov equation written in terms of the distribution function $f(\mathbf{q},\mathbf{p},t)$ is

$$\dot{f} + \mathbf{u}\cdot\nabla f = 0$$

Expanding now the Hamiltonian h as the total single particle energy

$$h(\mathbf{q},\mathbf{p},t) = \frac{1}{2m}\mathbf{p}^2 + V(\mathbf{q},\mathbf{p},t)$$

one obtains the more common form

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m}\cdot\frac{\partial f}{\partial \mathbf{q}} - \frac{\partial V}{\partial \mathbf{q}}\cdot\frac{\partial f}{\partial \mathbf{p}} = 0$$

Problem 33.1 Can the Vlasov equation be cast in Lie–Poisson form?

We show here why the answer is yes. First we write the Vlasov equation in terms of a generic single particle Hamiltonian h as

$$\dot{f} + \{f,h\} = 0$$

where we recall the canonical Poisson bracket

$$\{f,h\} = \frac{\partial f}{\partial \mathbf{q}}\cdot\frac{\partial h}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}}\cdot\frac{\partial h}{\partial \mathbf{q}}$$

The main point of this discussion is that the canonical Poisson bracket provides the set $\mathcal{F}(T^*\mathbb{R}^N)$ of the functions on the phase space with a Lie algebra structure

$$[k,h] = \{k,h\}$$

At this point, in order to look for a Lie–Poisson equation, one calculates the coadjoint operator such that

$$\langle f, \{h,k\}\rangle = \langle f, \mathrm{ad}_h k\rangle = \langle \mathrm{ad}^*_h f, k\rangle = \langle -\{h,f\}, k\rangle$$

where the last equality is justified by the Leibniz property of the Poisson bracket, with the pairing defined as

$$\langle f,g\rangle = \int f\, g\, d\mathbf{q}d\mathbf{p}.$$

In conclusion, the argument above shows that the Vlasov equation can in fact be written in the Lie–Poisson form

$$\dot{f} + \mathrm{ad}^*_h f = 0$$

33.2 Ideal barotropic compressible fluids

The reduced Lagrangian for ideal compressible fluids is written as

$$l(\mathbf{u}, D) = \int \frac{D}{2}|\mathbf{u}|^2 - De(D)\, d\mathbf{x}$$

where $\mathbf{u} \in \mathfrak{X}(M \subset \mathbb{R}^3)$ is tangential on the boundary ∂M and D is the advected density, which satisfies the *continuity equation*

$$\partial_t D + \mathcal{L}_{\mathbf{u}} D = 0.$$

Moreover, the internal energy satisfies the barotropic First Law of Thermodynamics

$$de = -p(D)d(D^{-1}) = \frac{p(D)}{D^2}\, dD$$

for the pressure $p(D)$. The "reduced" Legendre transform on this Lie algebra $\mathfrak{X}(\mathbb{R}^3)$ is given by

$$\mathbf{m} = D\mathbf{u}$$

and the Hamiltonian is then written as

$$h(\mathbf{m}, D) = \langle \mathbf{m}, \mathbf{u} \rangle - l(\mathbf{u}, D)$$

that is

$$h(\mathbf{m}, D) = \int \frac{1}{2D}\, |\mathbf{m}|^2 \, + De(D)\, d\mathbf{x}$$

The Lie–Poisson equations in this case are as from the general theory

$$\begin{aligned} \partial_t \mathbf{m} &= -\mathrm{ad}^*_{\delta h/\delta \mathbf{m}}\, \mathbf{m} - \frac{\delta h}{\delta D} \diamond D \\ \partial_t D &= -\mathcal{L}_{\delta h/\delta \mathbf{m}}\, D \end{aligned}$$

Earlier we found that the coadjoint action is given by the Lie derivative. On the other hand we may calculate the expression of the diamond operation from its definition

$$\left\langle \frac{\delta h}{\delta D}, -\pounds_\eta D \right\rangle = \left\langle \frac{\delta h}{\delta D} \diamond D, \eta \right\rangle$$

to be

$$\left\langle \frac{\delta h}{\delta D}, -\operatorname{div} D\eta \right\rangle = \left\langle D\nabla \frac{\delta h}{\delta D}, \eta \right\rangle$$

Therefore, we have

$$\frac{\delta h}{\delta D} \diamond D = D\nabla \frac{\delta h}{\delta D}$$

where

$$\delta h/\delta D = -\frac{|\mathbf{m}|^2}{2D^2} + \Big(e + \frac{p}{D}\Big)$$

Substituting into the momentum equation and using the First Law to find $d(e+p/D) = (1/D)dp$ yields

$$\partial_t \mathbf{m} = -\mathcal{L}_{\mathbf{u}}\,\mathbf{m} - \nabla p$$

Upon expanding the Lie derivative for the momentum density $\mathbf{m}$ and using the continuity equation for the density, this quickly becomes

$$\partial_t \mathbf{u} = -\mathbf{u}\cdot\nabla\mathbf{u} - \frac{1}{D}\nabla p$$

which is Euler's equation for a barotropic fluid.

33.3 Euler's equations for ideal incompressible fluid motion

The barotropic equations recover Euler's equations for ideal incompressible fluid motion when the internal energy in the reduced Lagrangian for ideal compressible fluids is replaced by the constraint $D = 1$, as

$$l(\mathbf{u}, D) = \int \frac{D}{2}|\mathbf{u}|^2 - p(D-1)\,d\mathbf{x}$$

where again $\mathbf{u} \in \mathfrak{X}(M \subset \mathbb{R}^3)$ is tangential on the boundary ∂M and the advected density D satisfies the continuity equation,

$$\partial_t D + \operatorname{div} D\mathbf{u} = 0.$$

This equation enforces incompressibility $\operatorname{div}\mathbf{u} = 0$ when evaluated on the constraint $D = 1$. The pressure p is now a Lagrange multiplier, which is determined by the condition that incompressibility be preserved by the dynamics.

References

[1] **R Abraham**, **J E Marsden**, *Foundations of mechanics*, Benjamin/Cummings Publishing Co. Advanced Book Program, Reading, MA (1978) MR515141 Second edition, revised and enlarged, With the assistance of Tudor Raţiu and Richard Cushman

[2] **V Arnold**, *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 (1966) 319–361 MR0202082

[3] **V I Arnold**, *Mathematical methods of classical mechanics*, second edition, Graduate Texts in Mathematics 60, Springer, New York (1989) MR997295 Translated from the Russian by K Vogtmann and A Weinstein

[4] **V I Arnold**, **B A Khesin**, *Topological methods in hydrodynamics*, Applied Mathematical Sciences 125, Springer, New York (1998) MR1612569

[5] **A M Bloch**, *Nonholonomic mechanics and control*, Interdisciplinary Applied Mathematics 24, Springer, New York (2003) MR1978379 With the collaboration of J Baillieul, P Crouch and J Marsden. With scientific input from P S Krishnaprasad, R M Murray and D Zenkov

[6] **A M Bloch**, **R W Brockett**, **P E Crouch**, *Double bracket equations and geodesic flows on symmetric spaces*, Comm. Math. Phys. 187 (1997) 357–373 MR1463833

[7] **A M Bloch**, **P E Crouch**, *Optimal control and geodesic flows*, Systems Control Lett. 28 (1996) 65–72 MR1404249

[8] **F Calogero**, *An integrable Hamiltonian system*, Phys. Lett. A 201 (1995) 306–310 MR1331829

[9] **F Calogero**, **J-P Françoise**, *A completely integrable Hamiltonian system*, J. Math. Phys. 37 (1996) 2863–2871 MR1390240

[10] **R Camassa**, **D D Holm**, *An integrable shallow water equation with peaked solitons*, Phys. Rev. Lett. 71 (1993) 1661–1664 MR1234453

[11] **S Chandrasekhar**, *Ellipsoidal Figures of Equilibrium*, Yale University Press (1969)

[12] **E A Coddington**, **N Levinson**, *Theory of Ordinary Differential Equations*, Robert E Krieger (1984)

[13] **R Dedekind**, *Zusatz zu der vorstehenden Abhandlung*, J. Reine Angew. Math. 58 (1860) 217–228

[14] **H Dullin**, **G Gottwald**, **D D Holm**, *An integrable shallow water equation with linear and nonlinear dispersion*, Phys. Rev. Lett. 87 (2001) 194501–04

[15] **H R Dullin**, **G A Gottwald**, **D D Holm**, *Camassa–Holm, Korteweg-de Vries-5 and other asymptotically equivalent equations for shallow water waves*, Fluid Dynam. Res. 33 (2003) 73–95 MR1995028 In memoriam Prof. Philip Gerald Drazin 1934–2002

[16] **H R Dullin**, **G A Gottwald**, **D D Holm**, *On asymptotically equivalent shallow water wave equations*, Phys. D 190 (2004) 1–14 MR2043789

[17] **D G Ebin**, **J Marsden**, *Groups of diffeomorphisms and the motion of an incompressible fluid.*, Ann. of Math. (2) 92 (1970) 102–163 MR0271984

[18] **H Ertel**, *Ein neuer hydrodynamischer Wirbelsatz*, Meteorol. Z. 59 (1942) 271–281

[19] **O B Fringer**, **D D Holm**, *Integrable vs. nonintegrable geodesic soliton behavior*, Phys. D 150 (2001) 237–263 MR1820736

[20] **J Gibbons**, **D D Holm**, **B Kupershmidt**, *Gauge-invariant Poisson brackets for chromohydrodynamics*, Phys. Lett. A 90 (1982) 281–283 MR668288

[21] **V Guillemin**, **S Sternberg**, *Symplectic techniques in physics*, Cambridge University Press, Cambridge (1984) MR770935

[22] **L Haine**, *The algebraic complete integrability of geodesic flow on* $\mathrm{SO}(N)$, Comm. Math. Phys. 94 (1984) 271–287 MR761797

[23] **D D Holm**, *Euler–Poincaré dynamics of perfect complex fluids*, from: "Geometry, mechanics, and dynamics", Springer, New York (2002) 113–167 MR1919828

[24] **D Holm**, *The Euler–Poincaré variational framework for modeling fluid dynamics*, from: "Geometric mechanics and symmetry", London Math. Soc. Lecture Note Ser. 306, Cambridge Univ. Press, Cambridge (2005) 157–209 MR2172979

[25] **D D Holm**, *Geometric Mechanics Part I: Dynamics and Symmetry*, Imperial College Press, London (2008) MR2419209

[26] **D D Holm**, *Geometric Mechanics Part II: Rotating, Translating and Rolling*, Imperial College Press, London (2008) MR2419210

[27] **D D Holm**, **B A Kupershmidt**, *Poisson structures of superfluids*, Phys. Lett. A 91 (1982) 425–430 MR677628

[28] **D D Holm**, **B A Kupershmidt**, *Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity*, Phys. D 6 (1983) 347–363 MR709394

[29] **D D Holm**, **B A Kupershmidt**, *The analogy between spin glasses and Yang–Mills fluids*, J. Math. Phys. 29 (1988) 21–30 MR921760

[30] **D D Holm**, **J E Marsden**, *Momentum maps and measure-valued solutions (peakons, filaments, and sheets) for the EPDiff equation*, from: "The breadth of symplectic and Poisson geometry", Progr. Math. 232, Birkhäuser, Boston (2005) 203–235 MR2103008

[31] **D D Holm**, **J E Marsden**, **T S Ratiu**, *The Euler–Poincaré equations and semidirect products with applications to continuum theories*, Adv. Math. 137 (1998) 1–81 MR1627802

[32] **D D Holm**, **J E Marsden**, **T S Ratiu**, *Euler–Poincaré models of ideal fluids with nonlinear dispersion*, Phys. Rev. Lett. 80 (1998) 4173–4177

[33] **D D Holm**, **T Schmah**, **C Stoica**, *Geometric Mechanics and symmetry: From Finite to Infinite Dimensions*, Oxford Texts in Applied and Engineering Mathematics 12, Oxford University Press, Oxford (2009) MR2548736 With solutions to selected exercises by David C P Ellis

[34] **J K Hunter**, **Y X Zheng**, *On a completely integrable nonlinear hyperbolic variational equation*, Phys. D 79 (1994) 361–386 MR1306466

[35] **J V José**, **E J Saletan**, *Classical dynamics: A contemporary approach*, Cambridge University Press, Cambridge (1998) MR1640663

[36] **C H Kimberling**, *Emmy Noether*, Amer. Math. Monthly 79 (1972) 136–149 MR0299446

[37] **B Kostant**, *Orbits, symplectic structures and representation theory*, from: "Proc. U.S.-Japan Seminar in Differential Geometry (Kyoto, 1965)", Nippon Hyoronsha, Tokyo (1966) p. 71 MR0213476

[38] **J M Lee**, *Introduction to smooth manifolds*, Graduate Texts in Mathematics 218, Springer, New York (2003) MR1930091

[39] **P Libermann**, **C-M Marle**, *Symplectic geometry and analytical mechanics*, Mathematics and its Applications 35, D. Reidel Publishing Co., Dordrecht (1987) MR882548 Translated from the French by Bertram Eugene Schwarzbach

[40] **S Lie**, *Theorie der Transformationsgruppen. Zweiter Abschnitt*, Teubner, Leipzig (1890)

[41] **S V Manakov**, *A remark on the integration of the Eulerian equations of the dynamics of an n–dimensional rigid body*, Funkcional. Anal. i Priložen. 10 (1976) 93–94 MR0455031

[42] **J E Marsden**, *Lectures on mechanics*, London Mathematical Society Lecture Note Series 174, Cambridge University Press, Cambridge (1992) MR1171218

[43] **J E Marsden**, **T J R Hughes**, *Mathematical foundations of elasticity*, Dover Publications, New York (1994) MR1262126 Corrected reprint of the 1983 original

[44] **J E Marsden**, **T S Ratiu**, *Introduction to mechanics and symmetry*, Texts in Applied Mathematics 17, Springer, New York (1994) MR1304682 A basic exposition of classical mechanical systems

[45] **J E Marsden**, **T S Ratiu**, *Mechanics and Symmetry: Reduction Theory*, in preparation (2003)

[46] **J E Marsden**, **T Ratiu**, **A Weinstein**, *Reduction and Hamiltonian structures on duals of semidirect product Lie algebras*, from: "Fluids and plasmas: geometry and dynamics (Boulder, Colo., 1983)", Contemp. Math. 28, Amer. Math. Soc., Providence, RI (1984) 55–100 MR751975

[47] **J E Marsden**, **T Raţiu**, **A Weinstein**, *Semidirect products and reduction in mechanics*, Trans. Amer. Math. Soc. 281 (1984) 147–177 MR719663

[48] **J Marsden**, **A Weinstein**, *Reduction of symplectic manifolds with symmetry*, Rep. Mathematical Phys. 5 (1974) 121–130 MR0402819

[49] **J Marsden**, **A Weinstein**, *Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids*, Phys. D 7 (1983) 305–323 MR719058 Order in chaos (Los Alamos, N.M., 1982)

[50] **D McDuff**, **D Salamon**, *Introduction to symplectic topology*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York (1995) MR1373431

[51] **A S Miščenko**, **A T Fomenko**, *Euler equation on finite-dimensional Lie groups*, Izv. Akad. Nauk SSSR Ser. Mat. 42 (1978) 396–415, 471 MR0482832

[52] **K Ohkitani**, *Eigenvalue problems in three-dimensional Euler flows*, Phys. Fluids A 5 (1993) 2570–2572

[53] **P J Olver**, *Applications of Lie groups to differential equations*, second edition, Graduate Texts in Mathematics 107, Springer, New York (1993) MR1240056

[54] **J-P Ortega**, **T S Ratiu**, *Momentum maps and Hamiltonian reduction*, Progress in Mathematics 222, Birkhäuser, Boston (2004) MR2021152

[55] **H Poincaré**, *Sur une forme nouvelle des équations de la méchanique*, C.R. Acad. Sci. 132 (1901) 369–371

[56] **T Raţiu**, *The motion of the free n–dimensional rigid body*, Indiana Univ. Math. J. 29 (1980) 609–629 MR578210

[57] **T S Ratiu**, **R Tudoran**, **L Sbano**, **E Sousa Dias**, **G Terra**, *A crash course in geometric mechanics*, from: "Geometric mechanics and symmetry", London Math. Soc. Lecture Note Ser. 306, Cambridge Univ. Press, Cambridge (2005) 23–156 MR2172978 Notes of the courses given by Ratiu

[58] **B Riemann**, *Untersuchungen über die Bewegungen eines flüssigen gleichartigen Ellipsoides*, Abh. d. Königl. Gesell. der Wis. zu Göttingen 9 (1860) 3–36

[59] **S Smale**, *Topology and mechanics I*, Invent. Math. 10 (1970) 305–331 MR0309153

[60] **S Smale**, *Topology and mechanics II: The planar n–body problem*, Invent. Math. 11 (1970) 45–64 MR0321138

[61] **J-M Souriau**, *Structure des systèmes dynamiques*, Maîtrises de mathématiques, Dunod, Paris (1970) MR0260238

[62] **A Weinstein**, *Sophus Lie and symplectic geometry*, Exposition. Math. 1 (1983) 95–96 MR693812

[63] **A Weinstein**, *Groupoids: unifying internal and external symmetry. A tour through some examples*, from: "Groupoids in analysis, geometry, and physics (Boulder, CO, 1999)", Contemp. Math. 282, Amer. Math. Soc., Providence, RI (2001) 1–19 MR1855239

Department of Mathematics and Institute for Mathematical Sciences, Imperial College London, SW7 2AZ, United Kingdom

d.holm@imperial.ac.uk

Received: 18 May 2010