M3-4-5 A34 Handout: Geometric Mechanics II

Professor Darryl D Holm Spring Term 2013
Imperial College London d.holm@ic.ac.uk
http://www.ma.ic.ac.uk/~dholm/

Course meets T, W, Th noon @ Hux 6M27

Text for the course M3-4-5 A34:

Geometric Mechanics II: Rotating, Translating & Rolling (aka GM2)
by Darryl D Holm, World Scientific: Imperial College Press, Singapore, Second edition (2011).
ISBN 978-1-84816-777-3

Contents

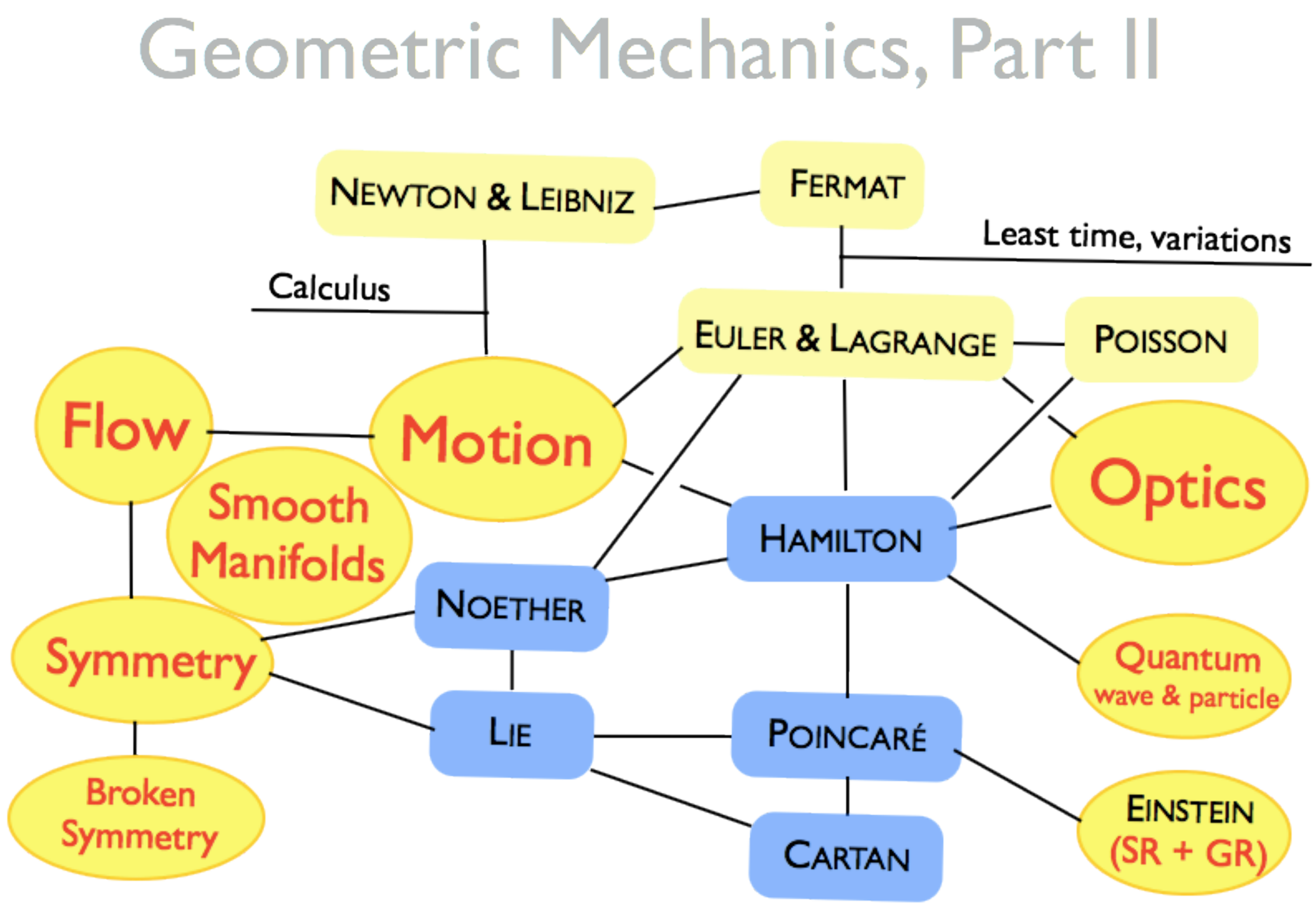

Figure 1: Geometric Mechanics has involved many great mathematicians!

1 Space, Time, Motion, ..., Symmetry, Dynamics!

Background reading: Chapter 2, [Ho2011GM1].

Space

Space is taken to be a manifold Q with points $q \in Q$ (Positions, States, Configurations). The manifold Q will sometime be taken to be a Lie group G. We will do this when we consider rotation and translation, for example. In this case the group is $G = SE(3)$ the special Euclidean group in three dimensions.

Time

Time is taken to be a manifold T with points $t \in T$. Usually $T = \mathbb{R}$ (for real 1D time), but we will also consider $T = \mathbb{R}^2$ and maybe let T and Q both be complex manifolds

Motion

Motion is a map $\phi_t : T \to Q$, where subscript t denotes dependence on time t. For example, when $T = \mathbb{R}$, the motion is a curve $q_t = \phi_t \circ q_0$ obtained by composition of functions. The motion is called a *flow* if $\phi_{t+s} = \phi_t \circ \phi_s$, for $s, t \in \mathbb{R}$, and $\phi_0 = \mathrm{Id}$, so that $\phi_t^{-1} = \phi_{-t}$. Note that the composition of functions is associative, $(\phi_t \circ \phi_s) \circ \phi_r = \phi_t \circ (\phi_s \circ \phi_r) = \phi_t \circ \phi_s \circ \phi_r = \phi_{t+s+r}$, but it is not commutative, in general. Thus, we should anticipate Lie group actions on manifolds.

Velocity

Velocity is an element of the tangent bundle TQ of the manifold Q. For example, $\dot{q}_t \in T_{q_t}Q$ along a flow q_t that describes a smooth curve in Q.

Motion equation

The motion equation that determines $q_t \in Q$ takes the form

$$\dot{q}_t = f(q_t)$$

where $f(q)$ is a prescribed *vector field* over Q. For example, if the curve $q_t = \phi_t \circ q_0$ is a flow, then

$$\dot{q}_t = \dot{\phi}_t \phi_t^{-1} \circ q_t = f(q_t)$$

so that

$$\dot{\phi}_t = f \circ \phi_t =: \phi_t^* f$$

which defines the pullback of f by ϕ_t.

Optimal motion equation – Hamilton's principle

An *optimal* motion equation arises from Hamilton's principle,

$$\delta S[q_t] = 0 \quad \text{for} \quad S[q_t] = \int L(q_t, \dot{q}_t)\, dt\,,$$

in which variational derivatives are given by

$$\delta S[q_t] = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} S[q_{t,\epsilon}]\,.$$

The introduction of a variational principle summons T^*Q, the cotangent bundle of Q. The cotangent bundle T^*Q is the dual space of the tangent bundle TQ, with respect to a pairing. That is, T^*Q is the space of real linear functionals on TQ with respect to the (real nondegenerate) pairing $\langle \cdot\,,\,\cdot \rangle$, induced by taking the variational derivative.

For example,

$$\text{if} \quad S = \int L(q, \dot{q})\, dt\,, \quad \text{then} \quad \delta S = \int \Big\langle \frac{\partial L}{\partial \dot{q}_t}\,,\, \delta \dot{q}_t) \Big\rangle + \Big\langle \frac{\partial L}{\partial q_t}\,,\, \delta q_t) \Big\rangle\, dt = 0$$

leads to the *Euler-Lagrange equations*

$$-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_t} + \frac{\partial L}{\partial q_t} = 0\,.$$

The map $p := \frac{\partial L}{\partial \dot{q}_t}$ is called the *fibre derivative* of the Lagrangian $L : TQ \to \mathbb{R}$. The Lagrangian is called *hyperregular* if the velocity can be solved from the fibre derivative, as $\dot{q}_t = v(q,p)$. Hyperregularity of the Lagrangian is sufficient for invertibility of the *Legendre transformation*

$$H(q,p) := \langle\, p\,,\, \dot{q}\,\rangle - L(q, \dot{q})$$

In this case, Hamilton's principle

$$0 = \delta \int \langle\, p\,,\, \dot{q}\,\rangle - H(q,p)\, dt\,,$$

gives *Hamilton's canonical equations*

$$\dot{q} = H_p \quad \text{and} \quad \dot{p} = -\,H_q\,,$$

whose solutions are equivalent to those of the Euler-Lagrange equations.

Symmetry

Lie group symmetries of the Lagrangian will be particularly important, both in reducing the number of independent degrees of freedom in Hamilton's principle and in finding conservation laws by Noether's theorem.

Dynamics!

Dynamics is the science of deriving, analysing, solving and interpreting the solutions of motion equations. GM2 will concentrate on dynamics in the case that the configuration space Q is a Lie group itself G and the Lagrangian $TG \to \mathbb{R}$ is transforms simply (e.g., is invariant) under the action of G. When the Lagrangian $TG \to \mathbb{R}$ is invariant under G, the problem reduces to a formulation on $TG/G \simeq \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of the Lie group G. With an emphasis on applications in mechanics, we will discuss a variety of interesting properties and results that are inherited from this formulation of dynamics on Lie groups.

What shall we study?

Figure 1 illustrates some of the relationships among the various accomplishments of the founders of geometric mechanics. We shall study these accomplishments and the relationships among them.

Hamilton: Quaternions, AD, Ad, ad, Ad*, ad* actions, variational principles
Lie: Groups of transformations that depend smoothly on parameters
Poincaré: Mechanics on Lie groups, $SO(3)$, $SU(2)$, $Sp(2)$, $SE(3) \simeq SO(3)\circledS\mathbb{R}^3$
Noether: Implications of symmetry in variational principles

These accomplishments lead to a new view of dynamics. In particular, we will study mechanics on Lie groups.

2 Quaternions

Quaternions are defined by their multiplication rules.

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

$$\mathbb{J}_1\mathbb{J}_2\mathbb{J}_3 = -\mathbb{J}_0\,, \tag{2.1}$$

and the squares

$$\mathbb{J}_1^2 = \mathbb{J}_2^2 = \mathbb{J}_3^2 = -\mathbb{J}_0\,, \tag{2.2}$$

where $\mathbb{J}_0$ is the identity element. Thus, $\mathbb{J}_1\mathbb{J}_2 = \mathbb{J}_3$ holds, with cyclic permutations of $(1, 2, 3)$.

Quaternions combine a real scalar $q_0 \in \mathbb{R}$ and a real three-vector $\mathbf{q} \in \mathbb{R}^3$ with components $q_a,\ a = 1, 2, 3$, into a **tetrad**

$$\mathfrak{q} = [q_0,\, \mathbf{q}\,] = q_0\mathbb{J}_0 + q_1\mathbb{J}_1 + q_2\mathbb{J}_2 + q_3\mathbb{J}_3 \in \mathbb{H}\,. \tag{2.3}$$

Definition

2.1 (Multiplication of quaternions in vector notation)**.** The ***multiplication rule*** for two quaternions,

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

may be defined in vector notation as

$$\mathfrak{q}\mathfrak{r} = [q_0,\, \boldsymbol{q}][r_0,\, \boldsymbol{r}] = [\, q_0r_0 - \boldsymbol{q}\cdot\boldsymbol{r}\,,\, q_0\boldsymbol{r} + r_0\boldsymbol{q} + \boldsymbol{q}\times\boldsymbol{r}\,]\ . \tag{2.4}$$

Remark

2.2 (Quaternionic product is associative)**.** The quaternionic product is ***associative***:

$$\mathfrak{p}(\mathfrak{q}\mathfrak{r}) = (\mathfrak{p}\mathfrak{q})\mathfrak{r}\,. \tag{2.5}$$

However, the quaternionic product is not commutative, since

$$[\mathfrak{q},\, \mathfrak{r}] := \mathfrak{q}\mathfrak{r} - \mathfrak{r}\mathfrak{q} = [0,\, 2\boldsymbol{q}\times\mathbf{r}\,]\,. \tag{2.6}$$

Quaternions are interesting for studying alignment and rotations of vectors.

2.1 Alignment of vectors

Remark

2.3. Given two vectors $\mathbf{v}$ and $\hat{\boldsymbol{\omega}}$ with $|\hat{\boldsymbol{\omega}}|^2 = 1$, the decomposition

$$\mathbf{v}_{\parallel} = (\hat{\boldsymbol{\omega}} \cdot \mathbf{v})\hat{\boldsymbol{\omega}} \quad \text{and} \quad \mathbf{v}_{\perp} = (\hat{\boldsymbol{\omega}} \times \mathbf{v}) \times \hat{\boldsymbol{\omega}} \quad \text{with} \quad \mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} \tag{2.7}$$

is found by a quaternionic multiplication. Namely,

$$\mathfrak{v} = [0, \boldsymbol{v}] = [\hat{\boldsymbol{\omega}} \cdot \mathbf{v}, \hat{\boldsymbol{\omega}} \times \mathbf{v}][0, \hat{\boldsymbol{\omega}}] = [0, (\hat{\boldsymbol{\omega}} \cdot \mathbf{v})\hat{\boldsymbol{\omega}} + (\hat{\boldsymbol{\omega}} \times \mathbf{v}) \times \hat{\boldsymbol{\omega}}] = [0, \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}]$$

The vector decomposition (2.7) is precisely the quaternionic product (2.4), in which the vectors $[0, \mathbf{v}]$ and $[0, \hat{\boldsymbol{\omega}}]$ are treated as quaternions.

This remark may be summarised, as follows.

Summary

2.4 (Vector decomposition). *Quaternionic left multiplication of* $[0, \hat{\boldsymbol{\omega}}]$ *by* $[\hat{\boldsymbol{\omega}} \cdot \mathbf{v}, \, \hat{\boldsymbol{\omega}} \times \mathbf{v}]$ *decomposes the pure quaternion* $[0\,, \mathbf{v}]$ *into components that are* $\parallel$ *and* $\perp$ *to the pure unit quaternion* $[0\,, \hat{\boldsymbol{\omega}}]$. *Note that* $\mathbf{v} \cdot \mathbf{v} = (\hat{\boldsymbol{\omega}} \cdot \mathbf{v})^2 + |\hat{\boldsymbol{\omega}} \times \mathbf{v}|^2$.

2.2 Rotations of vectors

To study rotations of vectors by quaternionic multiplication, we need to define the dot-product of quaternions. For this, we first need the quaternionic conjugate.

Definition

2.5 (Quaternionic conjugate). One defines the ***conjugate*** of $\mathfrak{q} := [q_0\,,\ \boldsymbol{q}]$ in analogy to complex variables as

$$\mathfrak{q}^* = [q_0\,, -\boldsymbol{q}]\ . \tag{2.8}$$

Following this analogy, the scalar and vector parts of a quaternion are defined as

$$\mathrm{Re}\,\mathfrak{q} \;:=\; \frac{1}{2}(\mathfrak{q} + \mathfrak{q}^*) = [\,q_0,\ 0\,]\,, \tag{2.9}$$

$$\mathrm{Im}\,\mathfrak{q} \;:=\; \frac{1}{2}(\mathfrak{q} - \mathfrak{q}^*) = [\,0\,,\ \boldsymbol{q}\,]\,. \tag{2.10}$$

Definition

2.6 (Properties of quaternionic conjugation)**.** Two important properties of quaternionic conjugation are easily demonstrated. Namely,

$$
\begin{aligned}
(\mathfrak{p}\mathfrak{q})^* &= \mathfrak{q}^*\mathfrak{p}^* \quad \text{(note reversed order)}, &(2.11)\\
\mathrm{Re}(\mathfrak{p}\mathfrak{q}^*) &:= \frac{1}{2}(\mathfrak{p}\mathfrak{q}^* + \mathfrak{q}\mathfrak{p}^*)\\
&= [p_0q_0 + \boldsymbol{p}\cdot\boldsymbol{q}\,,\, 0\,] \quad \text{(yields real part)}. &(2.12)
\end{aligned}
$$

Note that conjugation reverses the order in the product of two quaternions.

Definition

2.7 (Dot product of quaternions)**.** The quaternionic ***dot product***, or ***inner product***, is defined as

$$
\begin{aligned}
\mathfrak{p}\cdot\mathfrak{q} &= [p_0\,,\,\boldsymbol{p}\,]\cdot[q_0\,,\,\boldsymbol{q}\,]\\
&:= [p_0q_0 + \boldsymbol{p}\cdot\boldsymbol{q}\,,\,0\,] = \mathrm{Re}(\mathfrak{p}\mathfrak{q}^*)\,.
\end{aligned}
\tag{2.13}
$$

Definition

2.8 (Pairing of quaternions)**.** The quaternionic dot product (2.13) defines a real symmetric pairing $\langle\,\cdot\,,\,\cdot\,\rangle : \mathbb{H}\times\mathbb{H}\mapsto\mathbb{R}$, denoted as

$$
\langle\,\mathfrak{p}\,,\,\mathfrak{q}\,\rangle = \mathrm{Re}(\mathfrak{p}\mathfrak{q}^*) := \mathrm{Re}(\mathfrak{q}\mathfrak{p}^*) = \langle\,\mathfrak{q}\,,\,\mathfrak{p}\,\rangle\,. \tag{2.14}
$$

In particular, $\langle\,\mathfrak{q}\,,\,\mathfrak{q}\,\rangle = \mathrm{Re}(\mathfrak{q}\mathfrak{q}^*) =: |\mathfrak{q}|^2$ is a positive real number.

Definition

2.9 (Magnitude of a quaternion)**.** *The **magnitude** of a quaternion* $\mathfrak{q}$ *is defined by*

$$
|\mathfrak{q}| := (\mathfrak{q}\cdot\mathfrak{q})^{1/2} = ({q_0}^2 + \mathbf{q}\cdot\mathbf{q})^{1/2}\,. \tag{2.15}
$$

Definition

2.10 (Quaternionic inverse)**.** *We have the product*

$$
|\mathfrak{q}|^2 := \mathfrak{q}\mathfrak{q}^* = (\mathfrak{q}\cdot\mathfrak{q})\mathfrak{e}\,, \tag{2.16}
$$

where $\mathfrak{e} = [1, 0]$ *is the* ***identity quaternion****. Hence, one may define*

$$\mathfrak{q}^{-1} := \mathfrak{q}^*/|\mathfrak{q}|^2 \tag{2.17}$$

to be the ***inverse*** *of quaternion* $\mathfrak{q}$*.*

Lemma

2.11 (Rotation of vectors)**.** If $\mathfrak{r} = [0, \boldsymbol{r}]$ and $\hat{\mathfrak{q}} = [q_0, \boldsymbol{q}]$ is a unit quaternion (i.e., $\hat{\mathfrak{q}} \cdot \hat{\mathfrak{q}} = [q_0^2 + |\boldsymbol{q}|^2, 0] = [1, 0]$) then the conjugation $\mathfrak{r}' = \hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^* = [0, O\mathbf{r}]$ where $O \in SO(3)$ is a rotation.

Proof. Write

$$\begin{aligned} \mathfrak{r}' &= \hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^* = [0, \mathbf{r}'] \\ &= \big[0, \boldsymbol{r} + 2q_0(\boldsymbol{q} \times \boldsymbol{r}) + 2\boldsymbol{q} \times (\boldsymbol{q} \times \boldsymbol{r})\big] , \end{aligned} \tag{2.18}$$

for which

$$|\mathfrak{r}'|^2 = \mathfrak{r}' \cdot \mathfrak{r}'^* = [0, \mathbf{r}'] \cdot [0, -\mathbf{r}'] = [|\mathbf{r}'|^2, 0] = [|\mathbf{r}|^2, 0] = |\mathfrak{r}|^2 . \tag{2.19}$$

Since $|\mathbf{r}'|^2 = |\mathbf{r}|^2$, the transformation $\mathfrak{r}' \to \mathfrak{r}$ induces a rotation of the vector $\mathbf{r}$. □

Exercise. Check the calculation in the proof of Lemma 2.11. ★

Remark

2.12. The transformation property $|\mathbf{r}'|^2 = |\mathbf{r}|^2$ explains why quaternions were useful in Hamilton's astronomical observations. Namely, he needed an efficient way of composing rotations from one direction to another for his telescope.

Theorem

2.13 (Rodriques formula)**.** The rotation of a vector $\mathbf{r}$ by an angle θ about a unit vector $\hat{\boldsymbol{n}}$ is given by

$$\mathbf{r}' = \mathbf{r} + \sin\theta(\hat{\boldsymbol{n}} \times \mathbf{r}) + (1 - \cos\theta)\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{n}} \times \mathbf{r}) \tag{2.20}$$

Proof. Consider $\mathfrak{r}' = \hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^* = [0, \mathbf{r}']$ with unit quaternion $\mathfrak{q} = \pm[\cos(\theta/2), \sin(\theta/2)\hat{\boldsymbol{n}}]$ and substitute into formula (2.18) for the quaterion conjugate. □

Exercise. Check the calculation in the proof of Theorem 2.13. ★

Definition

2.14 (Euler parameters)**.** In the Euler–Rodrigues formula (2.20) for the rotation of vector $\boldsymbol{r}$ by angle θ about $\hat{\boldsymbol{n}}$, the quantities θ, $\hat{\boldsymbol{n}}$ are called the ***Euler parameters***.

Definition

2.15 (Cayley–Klein parameters)**.** The unit quaternion $\hat{\mathfrak{q}} = [q_0\,,\,\boldsymbol{q}]$ corresponding to the rotation of a pure quaternion $\mathfrak{r} = [0\,,\,\boldsymbol{r}]$ by angle θ about $\hat{\boldsymbol{n}}$ using quaternionic conjugation $\mathfrak{r}' = \hat{\mathfrak{q}}\mathfrak{r}\hat{\mathfrak{q}}^* = [0, O\mathbf{r}]$ was taken as

$$\hat{\mathfrak{q}} := \pm\Big[\cos\frac{\theta}{2}\,,\,\sin\frac{\theta}{2}\,\hat{\boldsymbol{n}}\Big]\,. \tag{2.21}$$

The quantities $q_0 = \pm\cos\frac{\theta}{2}$ and $\boldsymbol{q} = \pm\sin\frac{\theta}{2}\,\hat{\boldsymbol{n}}$ in (2.21) are called its ***Cayley–Klein parameters***.

Remark

2.16 (Cayley–Klein coordinates of a quaternion)**.** *An arbitrary quaternion may be written in terms of its magnitude and its Cayley–Klein parameters as*

$$\mathfrak{q} = |\mathfrak{q}|\hat{\mathfrak{q}} = |\mathfrak{q}|\Big[\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\,\hat{\mathbf{n}}\Big]\,. \tag{2.22}$$

The calculation of the Euler–Rodrigues formula (2.20) shows the equivalence of quaternionic conjugation and rotations of vectors. Moreover, compositions of quaternionic products imply the following.

Corollary

2.17. *The composition of rotations*

$$O^{\theta'}_{\hat{\mathbf{n}}'}O^{\theta}_{\hat{\mathbf{n}}}\,\mathbf{r} = \hat{\mathfrak{q}}'(\hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^*)\hat{\mathfrak{q}}'^*$$

is equivalent to multiplication of $(\pm)$ *unit quaternions.*

Exercise. Show directly by quaternionic multiplication that

$$O_{\hat{\mathbf{y}}}^{\pi/2} O_{\hat{\mathbf{x}}}^{\pi/2} = O_{\hat{\mathbf{n}}}^{2\pi/3} \quad \text{with} \quad \hat{\mathbf{n}} = (\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}})/\sqrt{3}\,.$$

★

2.3 Cayley–Klein parameters, rigid-body dynamics

Consider a time-dependent rotation by angle $\theta(t)$ about direction $\hat{\mathbf{n}}(t)$ is given by the unit quaternion (2.21) in Cayley–Klein parameters,

$$\hat{\mathfrak{q}}(t) = [q_0(t), \mathbf{q}(t)] = \pm\Big[\cos\frac{\theta(t)}{2}\,,\, \sin\frac{\theta(t)}{2}\,\hat{\mathbf{n}}(t)\Big]\,. \tag{2.23}$$

The **operation** of the unit quaternion $\hat{\mathfrak{q}}(t)$ on a vector $\mathbf{X} = [0,\, \mathbf{X}]$ is given by quaternionic multiplication as

$$\begin{aligned} \hat{\mathfrak{q}}(t)\mathbf{X} &= [q_0, \mathbf{q}][0,\, \mathbf{X}] = [-\mathbf{q}\cdot\mathbf{X}\,,\, q_0\mathbf{X} + \mathbf{q}\times\mathbf{X}] \\ &= \pm\Big[-\sin\frac{\theta(t)}{2}\,\hat{\mathbf{n}}(t)\cdot\mathbf{X},\ \cos\frac{\theta(t)}{2}\,\mathbf{X} + \sin\frac{\theta(t)}{2}\,\hat{\mathbf{n}}(t)\times\mathbf{X}\Big]\,. \end{aligned} \tag{2.24}$$

The corresponding time-dependent rotation is given by the Euler–Rodrigues formula (2.20) as

$$\mathbf{x}(t) = \hat{\mathfrak{q}}(t)\mathbf{X}\hat{\mathfrak{q}}^*(t) \quad \text{so that} \quad \mathbf{X} = \hat{\mathfrak{q}}^*(t)\mathbf{x}(t)\hat{\mathfrak{q}}(t) \tag{2.25}$$

in terms of the unit quaternion $\hat{\mathfrak{q}}(t)$. Its time derivative is given by

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*\mathbf{x}\hat{\mathfrak{q}}\hat{\mathfrak{q}}^* + \hat{\mathfrak{q}}\hat{\mathfrak{q}}^*\mathbf{x}\hat{\mathfrak{q}}\dot{\hat{\mathfrak{q}}}^* = \dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*\mathbf{x} + \mathbf{x}\hat{\mathfrak{q}}\dot{\hat{\mathfrak{q}}}^* \\ &= \dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*\mathbf{x} + \mathbf{x}(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)^* = \dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*\mathbf{x} - \mathbf{x}(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*) \\ &= \dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*\mathbf{x} - \big((\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)\mathbf{x}\big)^* \\ &= 2\mathrm{Im}\big((\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)\mathbf{x}\big) \\ &= 2(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)\mathbf{x}\,. \end{aligned} \tag{2.26}$$

In quaternion components, this equation may be rewritten using the Cayley-Klein representation of a unit quaternion (2.23) as

$$2\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^* = [\,0\,,\; \dot\theta\hat{\mathbf{n}} + \sin\theta\,\dot{\hat{\mathbf{n}}} + (1-\cos\theta)\hat{\mathbf{n}}\times\dot{\hat{\mathbf{n}}}\,]\,. \tag{2.27}$$

The quantity $2\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^* = [0,\, 2(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)]$ is a pure quaternion whose vector component is denoted, for the moment, by enclosing it in parentheses as

$$2(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*) = \dot\theta\hat{\mathbf{n}} + \sin\theta\,\dot{\hat{\mathbf{n}}} + (1-\cos\theta)\hat{\mathbf{n}}\times\dot{\hat{\mathbf{n}}}\,.$$

As a consequence, the vector component of the quaternion equation (2.26) becomes

$$[\,0,\, \dot{\mathbf{x}}(t)\,] = 2[0,\, (\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)]\,[0,\, \mathbf{x}\,] = 2[\,0\,,\, (\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)\times\mathbf{x}\,]\,. \tag{2.28}$$

2.3.1 Spatial angular frequency

Upon recalling the isomorphism provided by the Euler–Rodrigues formula (2.20) for finite rotations,

$$\mathbf{x}(t) = O(t)\mathbf{X} = \hat{\mathfrak{q}}(t)\mathbf{X}\hat{\mathfrak{q}}^*(t)\,, \tag{2.29}$$

the vector component of (2.28) yields a series of isomorphisms for the angular frequency,

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \dot{O}O^{-1}(t)\mathbf{x}\\ &= \hat\omega(t)\mathbf{x}\\ &= \boldsymbol{\omega}(t)\times\mathbf{x}\\ &= 2(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)(t)\times\mathbf{x}\,.\end{aligned} \tag{2.30}$$

Since the quaternion $2\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*(t) = [\,0\,,\, 2(\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*)(t)\,]$ is equivalent to a vector in $\mathbb{R}^3$, we may simply use vector notation for it and equate the spatial angular frequencies as **vectors**. That is, we shall write

$$\boldsymbol{\omega}(t) = 2\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*(t)\,, \tag{2.31}$$

and, henceforth, we drop the parentheses $(\,\cdot\,)$ when identifying purely imaginary quaternions with vectors.

Remark

2.18. Pure quaternions of the form $\dot{\hat{\mathfrak{q}}}\hat{\mathfrak{q}}^*(t)$ may be identified with the tangent space of the *unit* quaternions at the identity.

2.4 Body angular frequency

The quaternion for the body angular frequency will have the corresponding vector expression,

$$\boldsymbol{\Omega}(t) = 2\hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}(t)\,. \tag{2.32}$$

Thus, only vector parts enter the quaternionic descriptions of the spatial and body angular frequencies. The resulting isomorphisms are entirely sufficient to express the quaternionic versions of the rigid-body equations of motion in their Newtonian, Lagrangian and Hamiltonian forms, including the Lie–Poisson brackets. In particular, the kinetic energy for the rigid body is given by

$$K = \frac{1}{2}\boldsymbol{\Omega}(t)\,\cdot\,\mathbb{I}\,\boldsymbol{\Omega}(t) = 2\Big\langle\,\hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}(t)\,,\,\mathbb{I}\,\hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}(t)\,\Big\rangle\,. \tag{2.33}$$

So the quaternionic description of rigid-body dynamics reduces to the equivalent description in $\mathbb{R}^3$.

This equivalence in the two descriptions of rigid-body dynamics means that the relations for angular momentum, Hamilton's principle and the Lie–Poisson brackets in terms of vector quantities all have identical expressions in the quaternionic picture. Likewise, the reconstruction of the Cayley–Klein parameters from the solution for the body angular velocity vector may be accomplished by integrating the linear quaternionic equation

$$\dot{\hat{\mathfrak{q}}}(t) = \hat{\mathfrak{q}}\boldsymbol{\Omega}(t)/2\,, \tag{2.34}$$

or explicitly,

$$\frac{d}{dt}\Big[\cos\frac{\theta}{2}\,,\,\sin\frac{\theta}{2}\,\hat{\mathbf{n}}\Big] = \Big[\cos\frac{\theta}{2}\,,\,\sin\frac{\theta}{2}\,\hat{\mathbf{n}}\Big]\big[\,0,\,\boldsymbol{\Omega}(t)/2\,\big]\,. \tag{2.35}$$

This is the linear reconstruction formula for the Cayley–Klein parameters.

Remark

2.19. Expanding this linear equation for the Cayley–Klein parameters leads to a quaternionic equation for the Euler parameters that is linear in $\boldsymbol{\Omega}(t)$, but is nonlinear in θ and $\hat{\boldsymbol{n}}$, namely

$$\big[\,\dot\theta,\,\dot{\hat{\boldsymbol{n}}}\,\big] = \big[\,0,\,\hat{\boldsymbol{n}}\,\big]\Big[\,0,\,\boldsymbol{\Omega}(t) + \boldsymbol{\Omega}(t)\times\hat{\boldsymbol{n}}\cot\frac{\theta}{2}\,\Big] = \Big[-\,\hat{\boldsymbol{n}}\cdot\boldsymbol{\Omega}(t)\,,\,\hat{\boldsymbol{n}}\times\Big(\boldsymbol{\Omega}(t)\times\hat{\boldsymbol{n}}\cot\frac{\theta}{2}\Big)\Big]\,. \tag{2.36}$$

2.5 Cayley–Klein parameters

The series of isomorphisms in Equation (2.30) holds the key for writing Hamilton's principle for Euler's rigid-body equations in quaternionic form using Cayley–Klein parameters. The key step will be deriving a formula for the variation of the body angular velocity. For this, one invokes equality of cross derivatives with respect to time t and variational parameter s. This equality in in terms of Cayley–Klein parameters will produce the formula needed for the quaternionic form of Hamilton's principle.

Proposition

2.20 (Cayley–Klein variational formula)**.** The variation of the pure quaternion $\boldsymbol{\Omega} = 2\hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}$ corresponding to body angular velocity in Cayley–Klein parameters satisfies the identity

$$\boldsymbol{\Omega}' - \dot{\boldsymbol{\Xi}} = (\boldsymbol{\Omega}\,\boldsymbol{\Xi} - \boldsymbol{\Xi}\,\boldsymbol{\Omega})/2 = \mathrm{Im}(\boldsymbol{\Omega}\,\boldsymbol{\Xi})\,, \tag{2.37}$$

where $\boldsymbol{\Xi} := 2\hat{\mathfrak{q}}^*\hat{\mathfrak{q}}'$ and $(\,\cdot\,)'$ denotes variation.

Proof. The body angular velocity is defined as $\boldsymbol{\Omega} = 2\hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}$ in (2.32). Its variational derivative is found to be

$$\delta\boldsymbol{\Omega} := \frac{d}{ds}\boldsymbol{\Omega}(s)\Big|_{s=0} =: \boldsymbol{\Omega}'\,. \tag{2.38}$$

Thus, the variation of $\boldsymbol{\Omega}$ may be expressed as

$$\boldsymbol{\Omega}'/2 = (\hat{\mathfrak{q}}^*)'\dot{\hat{\mathfrak{q}}} + \hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}'\,. \tag{2.39}$$

Now $\mathfrak{e} = \hat{\mathfrak{q}}^*\hat{\mathfrak{q}}$ so that

$$\mathfrak{e}' = 0 = (\hat{\mathfrak{q}}^*)'\hat{\mathfrak{q}} + \hat{\mathfrak{q}}^*\hat{\mathfrak{q}}' \quad \text{and} \quad (\hat{\mathfrak{q}}^*)' = -\,\hat{\mathfrak{q}}^*\hat{\mathfrak{q}}'\hat{\mathfrak{q}}^*\,. \tag{2.40}$$

Hence, the variation of the angular frequency becomes

$$\begin{aligned} \delta\boldsymbol{\Omega}/2 = \boldsymbol{\Omega}'/2 &= -\,\hat{\mathfrak{q}}^*\hat{\mathfrak{q}}'\hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}} + \hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}' \\ &= -\,\boldsymbol{\Xi}\,\boldsymbol{\Omega}/4 + \hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}'\,, \end{aligned} \tag{2.41}$$

where we have defined $\boldsymbol{\Xi} := 2\hat{\mathfrak{q}}^*\hat{\mathfrak{q}}'$, which satisfies a similar relation,

$$\dot{\boldsymbol{\Xi}}/2 \;=\; -\,\boldsymbol{\Omega}\,\boldsymbol{\Xi}/4 + \hat{\mathfrak{q}}^*\dot{\hat{\mathfrak{q}}}'\,. \tag{2.42}$$

Taking the difference of (2.41) and (2.42) yields

$$\boldsymbol{\Omega}' - \dot{\boldsymbol{\Xi}} = (\boldsymbol{\Omega}\,\boldsymbol{\Xi} - \boldsymbol{\Xi}\,\boldsymbol{\Omega})/2 = \mathrm{Im}(\boldsymbol{\Omega}\,\boldsymbol{\Xi})\,. \tag{2.43}$$

In quaternion components this formula becomes

$$\begin{aligned} \big[\,0\,,\,\boldsymbol{\Omega}'\,\big] - \big[\,0\,,\,\dot{\boldsymbol{\Xi}}\,\big] &= \frac{1}{2}\Big([\,0\,,\,\boldsymbol{\Omega}\,][\,0\,,\,\boldsymbol{\Xi}\,] - [\,0\,,\,\boldsymbol{\Xi}\,][\,0\,,\,\boldsymbol{\Omega}\,]\Big) \\ &= \big[\,0\,,\,\boldsymbol{\Omega}\times\boldsymbol{\Xi}\,\big]\,, \end{aligned} \tag{2.44}$$

or, in the equivalent vector form,

$$\boldsymbol{\Omega}' - \dot{\boldsymbol{\Xi}} = \boldsymbol{\Omega} \times \boldsymbol{\Xi}\,. \tag{2.45}$$

This was the key formula needed for writing Hamilton's principle in vector form, now reproduced in its pure quaternionic form for the Cayley–Klein parameters. □

Remark

2.21. *Having expressed the key vector variational formula (2.45) in quaternionic form (2.43), the path for deriving Hamilton's principle for the rigid body in the quaternionic picture proceeds in parallel with the vector case, which was treated in M345A16 (Geometric Mechanics I).*

Exercise. State and prove Hamilton's principle for the rigid body in quaternionic form. ★

2.6 Pauli matrix representation of quaternionic product

Theorem

2.22 (Isomorphism with Pauli matrix product)**.**
The quaternionic multiplication rule (2.4) may be represented in a 2×2 *matrix basis as*

$$\mathfrak{q} = [q_0,\, \mathbf{q}] = q_0\sigma_0 - i\mathbf{q}\cdot\boldsymbol{\sigma}\,, \quad \textit{with} \;\; \mathbf{q}\cdot\boldsymbol{\sigma} := \sum_{a=1}^{3} q_a\sigma_a\,, \tag{2.46}$$

where σ_0 *is the* 2×2 *identity matrix and* σ_a*, with* $a = 1, 2, 3$*, are the* ***Pauli spin matrices****,*

$$\begin{aligned} \sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \sigma_2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \end{aligned} \tag{2.47}$$

which are Hermitian: $\sigma^\dagger = \sigma$*.*

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

$$\sigma_a\sigma_b = \delta_{ab}\,\sigma_0 + i\epsilon_{abc}\sigma_c \quad \text{for} \quad a, b, c = 1, 2, 3, \tag{2.48}$$

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

Theorem

2.23. *The unit quaternions form a representation of the matrix Lie group* $SU(2)$*.*

Proof. The matrix representation of a unit quaternion is given in (2.46). Let $\hat{\mathfrak{q}} = [q_0,\, \boldsymbol{q}]$ be a unit quaternion ($|\hat{\mathfrak{q}}|^2 = q_0^2 + \boldsymbol{q}\cdot\boldsymbol{q} = 1$) and define the matrix Q by

$$\begin{aligned} Q &= q_0\sigma_0 - i\boldsymbol{q}\cdot\boldsymbol{\sigma} \\ &= \begin{bmatrix} q_0 - iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & q_0 + iq_3 \end{bmatrix}. \end{aligned} \tag{2.49}$$

The matrix Q is a unitary 2×2 matrix ($QQ^\dagger = Id$) with unit determinant ($\det Q = 1$). That is, $Q \in SU(2)$.

We may rewrite the map (2.18) for quaternionic conjugation of a vector $\mathfrak{r} = [0, \boldsymbol{r}] \to -i\mathbf{r} \cdot \boldsymbol{\sigma}$ by a unit quaternion equivalently in terms of ***unitary conjugation of the Hermitian Pauli spin matrices*** as

$$\mathfrak{r}' = \hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^* \iff \mathbf{r}' \cdot \boldsymbol{\sigma} = Q\,\mathbf{r} \cdot \boldsymbol{\sigma} Q^\dagger \,, \tag{2.50}$$

with

$$\mathbf{r} \cdot \boldsymbol{\sigma} = \begin{bmatrix} r_3 & r_1 - ir_2 \\ r_1 + ir_2 & -r_3 \end{bmatrix} . \tag{2.51}$$

This is the standard representation of $SO(3)$ rotations as a double covering ($\pm Q$) by $SU(2)$ matrices, which is now seen to be equivalent to quaternionic multiplication. □

Remark

2.24. Composition of $SU(2)$ matrices by matrix multiplication forms a Lie subgroup of the Lie group of 2×2 complex matrices $GL(2, \mathbb{C})$, see, e.g., [MaRa1994].

Exercise. Check that the matrix Q in (2.49) is a special unitary matrix so that $Q \in SU(2)$. That is, show that Q is unitary and has unit determinant. ★

Exercise. Verify the conjugacy formula (2.50) arising from the isomorphism between unit quaternions and $SU(2)$. ★

Remark

2.25. The ($\pm$) in the Cayley–Klein parameters reflects the 2:1 covering of the map $SU(2) \to SO(3)$.

2.7 Tilde map: $\mathbb{R}^3 \simeq su(2) \simeq so(3)$

The following ***tilde map*** may be defined by considering the isomorphism (2.46) for a pure quaternion $[0,\, \boldsymbol{q}]$. Namely,

$$
\begin{aligned}
\boldsymbol{q} \in \mathbb{R}^3 \to \tilde{\boldsymbol{q}} \ &:= \ -i\,\boldsymbol{q}\cdot\boldsymbol{\sigma} = -\,i\sum_{j=1}^{3} q_j\sigma_j \\
&= \ \begin{bmatrix} -iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & iq_3 \end{bmatrix} \in su(2)\,.
\end{aligned}
\tag{2.52}
$$

The tilde map (2.52) is a Lie algebra isomorphism between the cross product of vectors in $\mathbb{R}^3$ and the Lie algebra $su(2)$ of 2×2 skew-Hermitian traceless matrices. Just as in the hat map one writes

$$
JJ^\dagger(t) = Id \quad \Longrightarrow \quad \dot{J}J^\dagger + (\dot{J}J^\dagger)^\dagger = 0\,,
$$

so the tangent space at the identity for the $SU(2)$ matrices comprises 2×2 skew-Hermitian traceless matrices, whose basis is $-i\,\boldsymbol{\sigma}$, the imaginary number $(-i)$ times the three Pauli matrices. This completes the ***circle of the isomorphisms*** between Pauli matrices and quaternions, and between *pure* quaternions and vectors in $\mathbb{R}^3$. In particular, their Lie products are all isomorphic. That is,

$$
\mathrm{Im}(\mathfrak{pq}) = \frac{1}{2}\Big(\mathfrak{pq} - \mathfrak{qp}\Big) \to [\tilde{\boldsymbol{p}},\, \tilde{\boldsymbol{q}}] = (\boldsymbol{p}\times\boldsymbol{q})\tilde{}\;. \tag{2.53}
$$

In addition, recalling that $\mathrm{Re}(\mathfrak{pq}^*) = [\,\boldsymbol{p}\cdot\boldsymbol{q},\, 0\,]$ helps prove the following identities:

$$
\det(\boldsymbol{q}\cdot\boldsymbol{\sigma}) = |\boldsymbol{q}|^2\,, \quad \mathrm{tr}(\tilde{\boldsymbol{p}}\tilde{\boldsymbol{q}}) = -\,2\boldsymbol{p}\cdot\boldsymbol{q}\,.
$$

Exercise. Verify these formulas. ★

2.8 Dual of the tilde map: $\mathbb{R}^{3^*} \simeq su(2)^* \simeq so(3)^*$

One may identify $su(2)^*$ with $\mathbb{R}^3$ via the map $\mu \in su(2)^* \to \breve{\boldsymbol{\mu}} \in \mathbb{R}^3$ defined by a pairing $\langle\,\cdot\,,\,\cdot\,\rangle : su(2)^* \times su(2) \to \mathbb{R}$,

$$
\breve{\boldsymbol{\mu}}\cdot\mathbf{q} := \big\langle \mu, \tilde{\mathbf{q}} \big\rangle
$$

for any $\mathbf{q} \in \mathbb{R}^3$.

Then, for example,

$$
\breve{\boldsymbol{\mu}}\cdot(\boldsymbol{p}\times\boldsymbol{q}) := \big\langle \mu, [\tilde{\boldsymbol{p}},\, \tilde{\boldsymbol{q}}]\big\rangle,
$$

which foreshadows the adjoint and coadjoint actions of $SU(2)$ to appear in our discussions of rigid-body dynamics later.

2.9 Pauli matrices and Poincaré's sphere $\mathbb{C}^2 \to S^2$

The Lie algebra isomorphisms given by the Pauli matrix representation of the quaternions (2.46) and the tilde map (2.52) are related to a map $\mathbb{C}^2 \mapsto S^2$ first introduced by Poincaré [Po1892] and later studied by Hopf [Ho1931]. Consider for $a_k \in \mathbb{C}^2$, with $k = 1, 2$ the four real combinations written in terms of the Pauli matrices

$$n_\alpha = \sum_{k,l=1}^{2} a_k^* \{\sigma_\alpha\}_{kl}\, a_l \quad \text{with} \quad \alpha = 0, 1, 2, 3\,. \tag{2.54}$$

The $n_\alpha \in \mathbb{R}^4$ have components

$$\begin{aligned} n_0 &= |a_1|^2 + |a_2|^2\,, \\ n_3 &= |a_1|^2 - |a_2|^2\,, \\ n_1 + i\,n_2 &= 2a_1^* a_2\,. \end{aligned} \tag{2.55}$$

Remark

2.26. One may motivate the definition of $n_\alpha \in \mathbb{R}^4$ in (2.54) by introducing the following Hermitian matrix,

$$\rho = \mathbf{a} \otimes \mathbf{a}^* = \frac{1}{2}\Big(n_0 \sigma_0 + \mathbf{n} \cdot \boldsymbol{\sigma} \Big)\,, \tag{2.56}$$

in which the vector $\mathbf{n}$ is defined as

$$\mathbf{n} = \operatorname{tr} \rho\, \boldsymbol{\sigma} = a_l a_k^* \boldsymbol{\sigma}_{kl}\,. \tag{2.57}$$

The last equation recovers (2.54). We will return to the interpretation of this map when we discuss momentum maps later in the course. For now, we simply observe that the components of the *singular Hermitian matrix*,

$$\rho = \mathbf{a} \otimes \mathbf{a}^* = \frac{1}{2} \begin{bmatrix} n_0 + n_3 & n_1 - in_2 \\ n_1 + in_2 & n_0 - n_3 \end{bmatrix}, \quad \det\, \rho = 0,$$

are all invariant under the diagonal action

$$S^1 :\ \mathbf{a} \to e^{i\phi}\mathbf{a},\ \mathbf{a}^* \to e^{-i\phi}\mathbf{a}^*.$$

A fixed value $n_0 = const$ defines a three-sphere $S^3 \in \mathbb{R}^4$. Moreover, because $\det\rho = 0$ the remaining three components satisfy an additional relation which defines the ***Poincaré sphere*** $S^2 \in S^3$ as

$$n_0^2 = n_1^2 + n_2^2 + n_3^2 = |\mathbf{n}|^2\,. \tag{2.58}$$

Each point on this sphere defines a direction introduced by Poincaré to represent polarised light. The north (resp. south) pole represents right (resp. left) circular polarisation and the equator represents the various inclinations of linear polarisation. Off the equator and the poles the remaining directions in the upper and lower hemispheres represent right- and left-handed elliptical polarisations, respectively. Opposing directions $\pm\mathbf{n}$ correspond to orthogonal polarisations. See [BoWo1965] for details of the physical interpretation of the Poincaré sphere for polarised ray optics.

2.10 Poincaré's sphere and Hopf's fibration

The same map $S^3 \mapsto S^2$ given by (2.54) from the $n_0 = const$ S^3 to the Poincaré sphere S^2 was later studied by Hopf [Ho1931], who found it to be a ***fibration*** of S^3 over S^2. That is, $S^3 \simeq S^2 \times S^1$ locally, where S^1 is the fibre. A fibre bundle structure may be defined descriptively, as follows.

Definition

2.27 (Fibre bundle)**.** In topology, a ***fibre bundle*** is a space which locally looks like a product of two spaces but may possess a different global structure. Every fibre bundle consists of a continuous surjective map $\pi : E \mapsto B$, where small regions in the total space E look like small regions in the product space $B \times F$, of the ***base space*** B with the ***fibre space*** F (Figure 3.1). Fibre bundles comprise a rich mathematical subject that is explained more completely in other courses. We shall confine our attention here to the one particular case leading to the Poincaré sphere.

Remark

2.28. The ***Hopf fibration***, or fibre bundle, has spheres $S^3 \simeq S^2 \times S^1$ as its total space, base space and fibre, respectively. In terms of the Poincaré sphere one may think of the Hopf fibration locally as a sphere S^2 which has a great circle S^1 attached at every point. The phase on the great circles at opposite points are orthogonal (rotated by $\pi/2$, not π); so passing once around the Poincaré sphere along a great circle rotates the S^1 phase only by π, not 2π. One must pass *twice* around a great circle on the Poincaré sphere to return to the original phase. Thus, the relation $S^3 \simeq S^2 \times S^1$ only holds locally, not globally.

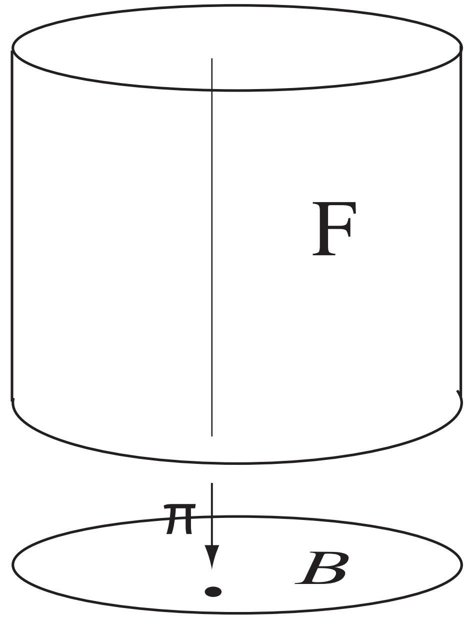

Figure 2: A fibre bundle E looks locally like the product space $B \times F$, of the base space B with the fibre space F. The map $\pi : E \approx B \times F \mapsto B$ projects E onto the base space B.

Remark

2.29. The conjugacy classes of S^3 by unit quaternions yield a version of the Hopf fibration $S^3 \simeq S^2 \times S^1$, obtained by identifying the Poincaré sphere (3.22) from the definitions (2.55).

Remark

2.30 (Hopf fibration)**.** Conjugating the Lie algebra $\mathfrak{su}(2)$ by the Lie group $SU(2)$ yields the entire unit two-sphere S^2. In particular,

$$\begin{bmatrix} a_1 & -a_2^* \\ a_2 & a_1^* \end{bmatrix}\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}\begin{bmatrix} a_1^* & a_2^* \\ -a_2 & a_1 \end{bmatrix} = \begin{bmatrix} -in_3 & -in_1+n_2 \\ -in_1-n_2 & in_3 \end{bmatrix} \quad \text{for} \quad |a_1|^2+|a_2|^2=1. \tag{2.59}$$

In other words,

$$\mathrm{Ad}_U(-i\mathbf{e}_3\cdot\boldsymbol{\sigma}) := U\,(-i\mathbf{e}_3\cdot\boldsymbol{\sigma})\,U^\dagger = -i\mathbf{n}\cdot\boldsymbol{\sigma}, \tag{2.60}$$

in which $U^\dagger = U^{-1}\in SU(2)$ and $|\mathbf{n}|^2=1$.

This is the tilde map (2.52) once again and (n_1,n_2,n_3) are the components of the Hopf fibration in (2.55).

Exercise. Express the Hopf fibration in terms of quaternionic multiplication.

Hint: Use the Cayley map $su(2)\to SU(2)$ given by

$$U = \big(q_0\sigma_0 + i\mathbf{q}\cdot\boldsymbol{\sigma}\big)\big(q_0\sigma_0 - i\mathbf{q}\cdot\boldsymbol{\sigma}\big)^{-1}, \quad \text{for} \quad q_0^2+|\mathbf{q}|^2 = 1\,. \tag{2.61}$$

★

3 $SU(2)$, quaternions and $1:1$ resonant coupled oscillators on $\mathbb{C}^2$

Resonant harmonic oscillators play a central role in physics. This is largely because the linearised dynamics of small excitations always leads to an eigenvalue problem.

Excitations oscillate. Nonlinear oscillations resonate. Under changes of parameters, resonant oscillations bifurcate.

The application of these ideas in physics is immense in scope, ranging from springs to swings, to molecules, to lasers, to coherent states in nuclear physics, to Bose–Einstein condensed (BEC) systems, to nonlinear optics, to telecommunication, to qubits in quantum computing.

This section treats the resonance of two coupled nonlinear oscillators as a physical application of the $SU(2)$ representation of quaternions.

3.1 Oscillator variables on $\mathbb{C}^2$

The linear transformation from $T^*\mathbb{R}^2$ with phase-space coordinates $(\mathbf{q},\mathbf{p})$ for two degrees of freedom to its ***oscillator variables*** $(\mathbf{a},\mathbf{a}^*)\in\mathbb{C}^2$ is defined by

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} q_1+ip_1 \\ q_2+ip_2 \end{bmatrix} = \mathbf{q}+i\mathbf{p}\in\mathbb{C}^2\,. \tag{3.1}$$

This linear transformation is canonical: its symplectic two-form is given by

$$\begin{aligned} dq_j \wedge dp_j &= \frac{1}{(-2i)}(dq_j + i dp_j) \wedge (dq_j - i dp_j) \\ &= \frac{1}{(-2i)} da_j \wedge da_j^* \,. \end{aligned} \tag{3.2}$$

Likewise, the Poisson bracket transforms by the chain rule as

$$\begin{aligned} \{a_j, a_k^*\} &= \{q_j + ip_j \,,\, q_k - ip_k\} \\ &= -2i \,\{q_j \,,\, p_k\} = -2i\,\delta_{jk} \,. \end{aligned} \tag{3.3}$$

Thus, in oscillator variables Hamilton's canonical equations become

$$\dot{a}_j = \{a_j,\, H\} = -2i\,\frac{\partial H}{\partial a_j^*}\,, \tag{3.4}$$

$$\text{and} \quad \dot{a}_j^* = \{a_j^*,\, H\} = 2i\,\frac{\partial H}{\partial a_j}\,.$$

The corresponding Hamiltonian vector field is

$$X_H = \{\cdot\,, H\} = -2i\left(\frac{\partial H}{\partial a_j^*}\frac{\partial}{\partial a_j} - \frac{\partial H}{\partial a_j}\frac{\partial}{\partial a_j^*}\right)\,, \tag{3.5}$$

satisfying the defining property of a Hamiltonian system,

$$X_H \,\lrcorner\, \frac{1}{(-2i)} da_j \wedge da_j^* = dH\,. \tag{3.6}$$

The norm $|\mathbf{a}|$ of a complex number $\mathbf{a} \in \mathbb{C}^2$ is defined via the real-valued ***pairing***

$$\Big\langle\!\Big\langle \cdot\,,\,\cdot \Big\rangle\!\Big\rangle : \mathbb{C}^2 \otimes \mathbb{C}^2 \mapsto \mathbb{R}\,. \tag{3.7}$$

For $\mathbf{a}\,,\,\mathbf{b} \in \mathbb{C}^2$, this pairing takes the value

$$\Big\langle\!\Big\langle \mathbf{a}\,,\,\mathbf{b} \Big\rangle\!\Big\rangle = \mathbf{a}^* \cdot \mathbf{b} = \left[a_1^*,\, a_2^*\right] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1^* b_1 + a_2^* b_2\,. \tag{3.8}$$

Here the dot $(\,\cdot\,)$ denotes the usual inner product of vectors. The pairing $\langle\,\cdot\,,\,\cdot\,\rangle$ defines a norm $|\mathbf{a}|$ on $\mathbb{C}^2$ by setting

$$|\mathbf{a}|^2 = \langle \mathbf{a}\,,\,\mathbf{a} \rangle = \mathbf{a}^* \cdot \mathbf{a} = a_1^* a_1 + a_2^* a_2\,. \tag{3.9}$$

3.2 The $1\!:\!1$ resonant action of S^1 on $\mathbb{C}^2$

- The norm $|\mathbf{a}|$ on $\mathbb{C}^2$ is *invariant* under a ***unitary change of basis*** $\mathbf{a} \to U\mathbf{a}$.
- The group of 2×2 ***unitary matrix transformations*** is denoted $U(2)$ and satisfies the condition

$$U^\dagger U = Id = UU^\dagger\,,$$

where the dagger in $U^\dagger$ denotes the ***Hermitian adjoint*** (conjugate transpose) of U. That is,

$$\begin{aligned}|U\mathbf{a}|^2 &= \langle U\mathbf{a}\,,\,U\mathbf{a}\rangle = (U\mathbf{a})^\dagger\cdot(U\mathbf{a})\\ &= (\mathbf{a}^*U^\dagger)\cdot(U\mathbf{a}) = \mathbf{a}^*\cdot\mathbf{a} = \langle\mathbf{a}\,,\,\mathbf{a}\rangle = |\mathbf{a}|^2\,.\end{aligned}$$

- The ***determinant*** $\det U$ of a unitary matrix U is a complex number of unit modulus; that is, $\det U$ belongs to $U(1)$.
- A unitary matrix U is called ***special unitary*** if its determinant satisfies $\det U = 1$. The special unitary 2×2 matrices are denoted $SU(2)$. Since determinants satisfy a product rule, one sees that $U(2)$ factors into $U(1)_{diag}\times SU(2)$ where $U(1)_{diag}$ is an overall phase times the identity matrix. For example,

$$\begin{bmatrix} e^{i(r+s)} & 0\\ 0 & e^{i(r-s)}\end{bmatrix} = \begin{bmatrix} e^{ir} & 0\\ 0 & e^{ir}\end{bmatrix}\begin{bmatrix} e^{is} & 0\\ 0 & e^{-is}\end{bmatrix}, \tag{3.10}$$

where $r, s\in\mathbb{R}$. The first of these matrices is the $U(1)_{diag}$ phase shift, while the second is a possible $SU(2)$ phase shift. Together they make up the individual phase shifts by $r+s$ and $r-s$. The $U(1)_{diag}$ phase shift is also called a ***diagonal*** S^1 ***action***.

Definition

3.1 (Diagonal S^1 action).

- The ***action*** of a $U(1)_{diag}$ phase shift $S^1:\mathbb{C}^2\mapsto\mathbb{C}^2$ on a complex two-vector with real angle parameter r is given by

$$\mathbf{a}(s) = e^{-2ir}\mathbf{a}(0) \quad\text{and}\quad \mathbf{a}^*(s) = e^{2ir}\mathbf{a}^*(0)\,, \tag{3.11}$$

which solves the evolution equations,

$$\frac{d}{dr}\mathbf{a}(r) = -2i\mathbf{a} \quad\text{and}\quad \frac{d}{dr}\mathbf{a}^*(r) = 2i\mathbf{a}^*\,. \tag{3.12}$$

- This operation is generated canonically by the ***Hamiltonian vector field***

$$X_R = \{\,\cdot\,, R\} = -2i\,\mathbf{a}\cdot\frac{\partial}{\partial \mathbf{a}} + 2i\,\mathbf{a}^*\cdot\frac{\partial}{\partial \mathbf{a}^*} =: \frac{d}{dr}\,, \tag{3.13}$$

  with canonically conjugate variables $(\mathbf{a},\mathbf{a}^*)\in\mathbb{C}^2$ satisfying the Poisson bracket relations in (3.3)

$$\{a_j, a_k^*\} = -2i\,\delta_{jk}.$$

- The variable R canonically conjugate to the ***angle*** r is called its corresponding ***action*** and is given by

$$R = |a_1|^2 + |a_2|^2 = |\mathbf{a}|^2 \quad \text{for which} \quad \{r, R\} = 1\,. \tag{3.14}$$

Definition

3.2 (1 : 1 resonance dynamics)**.**

- The ***flow*** ϕ_t^R of the Hamiltonian vector field $X_R = \{\,\cdot\,,\, R\}$ is the 1 : 1 resonant phase shift. In this flow, the phases of both simple harmonic oscillations (a_1, a_2) precess counterclockwise at the same constant rate.

- In 1 : 1 resonant motion on $\mathbb{C}^2 = \mathbb{C}\times\mathbb{C}$, the two complex amplitudes a_1 and a_2 stay in phase, because they both oscillate at the same rate.

Remark

3.3 (Restricting to S^1-invariant variables)**.** Choosing a particular value of R restricts $\mathbb{C}^2 \to S^3$. As we shall see, the further restriction to S^1-invariant variables will *locally* map

$$\mathbb{C}^2/S^1 \to S^3/S^1 = S^2, \quad \text{(Hopf fibration)}.$$

3.3 The S^1-invariant Hermitian coherence matrix

The bilinear product $\rho = \mathbf{a}\otimes\mathbf{a}^*$ of the vector amplitudes $(\mathbf{a},\mathbf{a}^*)\in\mathbb{C}^2$ comprises an S^1-invariant 2×2 Hermitian matrix for the 1 : 1 resonance, whose components are

$$\rho = \mathbf{a}\otimes\mathbf{a}^* = \begin{bmatrix} a_1 a_1^* & a_1 a_2^* \\ a_2 a_1^* & a_2 a_2^* \end{bmatrix}\,. \tag{3.15}$$

Its matrix properties are

$$\rho^\dagger = \rho\,,\quad \operatorname{tr}\rho = R\,,\quad \det\rho = 0 \quad\text{and}\quad \rho\,\mathbf{a} = R\,\mathbf{a}\,.$$

The vector $\mathbf{a}\in\mathbb{C}^2$ is a complex eigenvector of the Hermitian matrix ρ that belongs to the real eigenvalue R. In addition, ρ projects out the components of $\mathbf{a}$ in any complex vector $\mathbf{b}$ as

$$\rho\,\mathbf{b} = \mathbf{a}\,(\mathbf{a}^*\cdot\mathbf{b})\,.$$

Because the determinant of ρ vanishes, we may rescale $\mathbf{a}$ to set $\operatorname{tr}\rho = R = 1$. Under this rescaling, the complex amplitude $\mathbf{a}$ becomes a unit vector and ρ becomes a projection matrix with unit trace, since $\operatorname{tr}\rho = R$, $\rho\,\mathbf{a} = R\,\mathbf{a}$, $\rho^2 = R\rho$ and $R = 1$.[1] In what follows, however, we shall explicitly keep track of the value of the real quantity R, so it will be available for use later as a bifurcation parameter in our studies of S^1-invariant Hamiltonian dynamics.

3.4 The Poincaré sphere $S^2\subset S^3$

We expand the Hermitian matrix $\rho = \mathbf{a}\otimes\mathbf{a}^*$ in (3.15) as a linear combination of the four 2×2 Pauli spin matrices $(\sigma_0\,,\boldsymbol{\sigma})$, with $\boldsymbol{\sigma} = (\sigma_1,\,\sigma_2,\,\sigma_3)$ given by

$$\begin{aligned}\sigma_0 &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\quad \sigma_1 = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},\\ \sigma_2 &= \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix},\quad \sigma_3 = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}.\end{aligned} \tag{3.16}$$

The result of the expansion is, in vector notation,

$$\rho = \mathbf{a}\otimes\mathbf{a}^* = \begin{bmatrix}a_1a_1^* & a_1a_2^*\\ a_2a_1^* & a_2a_2^*\end{bmatrix} = \frac{1}{2}\Big(R\sigma_0 + \mathbf{Y}\cdot\boldsymbol{\sigma}\Big)\,. \tag{3.17}$$

where the vector $\mathbf{Y}$ is defined by the trace formula,

$$\mathbf{Y} = \operatorname{tr}\left(\rho\,\boldsymbol{\sigma}\right) = a_k^*\boldsymbol{\sigma}_{kl}a_l\,. \tag{3.18}$$

[1] In optics, matrix ρ is called the ***coherency*** of the pulse and the scalar R is its intensity [Sh1984, Bl1965].

Definition

3.4 (Poincaré sphere). *The coefficients* $(R, \mathbf{Y})$ *in the expansion of the matrix* ρ *in (3.17) are the four quadratic* S^1 *invariants,*

$$\begin{aligned} R &= |a_1|^2 + |a_2|^2 \,, \\ Y_3 &= |a_1|^2 - |a_2|^2 \,, \quad \textit{and} \\ Y_1 + i\, Y_2 &= 2a_1^* a_2 \,. \end{aligned} \tag{3.19}$$

These satisfy the relation

$$\det \rho = R^2 - |\mathbf{Y}|^2 = 0 \,, \quad \textit{with} \quad |\mathbf{Y}|^2 \equiv Y_1^2 + Y_2^2 + Y_3^2 \,, \tag{3.20}$$

which defines the ***Poincaré sphere*** $S^2 \in S^3$ *of radius* R.

Definition

3.5 (Stokes vector). *The matrix decomposition (3.17) of the Hermitian matrix* ρ *in Equation (3.15) into the Pauli spin matrix basis (3.16) may be written explicitly as*

$$\begin{aligned} \rho = \mathbf{a} \otimes \mathbf{a}^* &= \begin{bmatrix} a_1 a_1^* & a_1 a_2^* \\ a_2 a_1^* & a_2 a_2^* \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} R + Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & R - Y_3 \end{bmatrix} , \end{aligned} \tag{3.21}$$

or, on factoring $R > 0$,

$$\begin{aligned} \rho &= \frac{R}{2}\left(\sigma_0 + \frac{\mathbf{Y} \cdot \boldsymbol{\sigma}}{R} \right) \\ &= \frac{R}{2} \begin{bmatrix} 1 + s_3 & s_1 - is_2 \\ s_1 + is_2 & 1 - s_3 \end{bmatrix} \\ &= \frac{R}{2} \begin{bmatrix} 1 + \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & 1 - \cos\theta \end{bmatrix} , \end{aligned} \tag{3.22}$$

where the vector with components $\mathbf{s} = (s_1, s_2, s_3)$ *is a unit vector defined by*

$$\begin{aligned} \mathbf{Y}/R = (Y_1, Y_2, Y_3)/R &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\ &=: (s_1, s_2, s_3) = \mathbf{s} \,. \end{aligned} \tag{3.23}$$

The unit vector $\mathbf{s}$ *with polar angles* θ *and* ϕ *is called the* ***Stokes vector*** *[St1852] (Figure 4.1).*

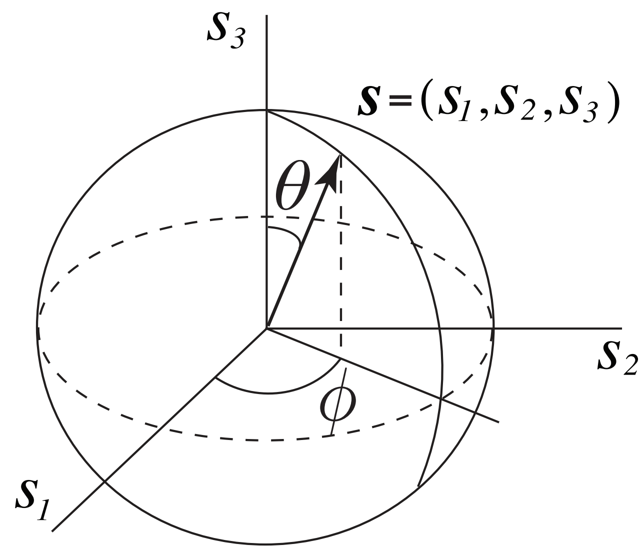

Figure 3: The Stokes vector on the unit Poincaré sphere.

3.5 Coherence matrix dynamics for the 1:1 resonance

For a given Hamiltonian $H : \mathbb{C}^2 \mapsto \mathbb{R}$, the matrix $\rho(t) = \mathbf{a} \otimes \mathbf{a}^*(t)$ in (3.15) evolves canonically according to

$$\dot{\rho}(t) = \{\rho,\, H\} = \dot{\mathbf{a}} \otimes \mathbf{a}^* + \mathbf{a} \otimes \dot{\mathbf{a}}^* = -4\,\mathrm{Im}\left(\mathbf{a} \otimes \frac{\partial H}{\partial \mathbf{a}}\right).$$

Thus, by the decomposition (3.17) of $\rho(t)$ in an $su(2)$ basis of Pauli matrices for an S^1-invariant Hamiltonian $H(\mathbf{Y}, R)$ we have

$$\dot{\rho} = \{\rho,\, H\} = \frac{1}{2}\{R\sigma_0 + \mathbf{Y}\cdot\boldsymbol{\sigma},\, H\} = \frac{1}{2}\{\mathbf{Y}\cdot\boldsymbol{\sigma},\, H\}\,,$$

where we have used $\dot{R} = \{R,\, H\} = 0$ which follows because of the S^1 symmetry of the Hamiltonian.

Proposition

3.6. *The Poisson brackets of the components $(Y_1, Y_2, Y_3) \in \mathbb{R}^3$ are computed by the product rule to close among themselves. This is expressed in tabular form as*

$$\{Y_i,\, Y_j\} = \begin{array}{|c|ccc|}\hline \{\,\cdot\,,\,\cdot\,\} & Y_1 & Y_2 & Y_3 \\ \hline Y_1 & 0 & 4Y_3 & -4Y_2 \\ Y_2 & -4Y_3 & 0 & 4Y_1 \\ Y_3 & 4Y_2 & -4Y_1 & 0 \\ \hline \end{array} \tag{3.24}$$

or, in index notation,

$$\{Y_k\,,\, Y_l\} = 4\epsilon_{klm} Y_m\,. \tag{3.25}$$

Proof. The proof of the result (3.24) is a direct verification using the chain rule for Poisson brackets,

$$\{Y_i,\, Y_j\} = \frac{\partial Y_i}{\partial z_A}\{z_A,\, z_B\}\frac{\partial Y_j}{\partial z_A}\,, \tag{3.26}$$

for the S^1-invariant bilinear functions $Y_i(z_A)$. In (3.26), one denotes $z_A = (a_A, a_A^*)$, with $A = 1, 2$ and $i, j = 1, 2, 3$. □

Corollary

3.7. *Thus, functions F, G of the S^1-invariant vector $\mathbf{Y} \in \mathbb{R}^3$ satisfy*

$$\{F\,,\, H\}(\mathbf{Y}) = 4\mathbf{Y} \cdot \frac{\partial F}{\partial \mathbf{Y}} \times \frac{\partial H}{\partial \mathbf{Y}}\,, \tag{3.27}$$

and one checks with $\dot{R} = 0$ that

$$\dot{\rho} = \{\rho,\, H\} = \Big\{\frac{1}{2}\mathbf{Y}\cdot\boldsymbol{\sigma},\, H\Big\} = -\,2\mathbf{Y} \times \frac{\partial H}{\partial \mathbf{Y}} \cdot \boldsymbol{\sigma} = \frac{1}{2}\dot{\mathbf{Y}}\cdot\boldsymbol{\sigma}\,.$$

Therefore, the motion equation for the S^1-invariant vector $\mathbf{Y} \in \mathbb{R}^3$ is

$$\dot{\mathbf{Y}} = -\,4\mathbf{Y} \times \frac{\partial H}{\partial \mathbf{Y}}\,, \tag{3.28}$$

which of course preserves the radius of the Poincaré sphere $|\mathbf{Y}| = R$.

Corollary

3.8. *Equation (3.27) for the Poisson bracket is contained in the set of Nambu brackets*

$$\{F\,,\,H\}(\mathbf{Y})\,d^3Y = 4\mathbf{Y}\cdot\frac{\partial F}{\partial \mathbf{Y}}\times\frac{\partial H}{\partial \mathbf{Y}}\,d^3Y = \nabla C(\mathbf{Y})\cdot\nabla F(\mathbf{Y})\times\nabla H(\mathbf{Y})\,d^3Y = dC\wedge dF\wedge dH\,,\quad \textit{when}\quad C(\mathbf{Y}) = 2|\mathbf{Y}|^2\,, \tag{3.29}$$

The Nambu brackets satisfy all the properties (bilinear, skew-symmetric, Leibniz, Jacobi) required to be a genuine Poisson bracket.

Thus, equation (3.28) proves the following theorem.

Theorem

3.9 (The map (3.20) to the ***Poincaré sphere*** $S^2 \in S^3$ of radius R is Poisson)**.** The map

$$J:\,\mathbb{C}^2/S^1 \to S^3/S^1 =_{\rm loc} S^2$$

in Equation (3.20) is a Poisson map. That is, it satisfies the Poisson property for smooth functions F and H, that

$$\big\{F\circ J\,,\,H\circ J\big\} = \big\{F\,,\,H\big\}\circ J\,. \tag{3.30}$$

Remark

3.10.

The Hamiltonian vector field (3.28)*in* $\mathbb{R}^3$ *for the evolution of the components of the coherence matrix* ρ *in (3.21) has the same form (modulo the factor of* -4*) as Euler's equations for a rigid body.*

As we will see later, this relation defines a ***Lie–Poisson bracket*** *on* $su(2)^*$ *that inherits the defining properties of a Poisson bracket from the canonical relations*

$$\{a_k,\,a_l^*\} = -2i\delta_{kl}\,,$$

for the canonical symplectic form, $\omega = \frac{1}{2}\,{\rm Im}\,(da_j\wedge da_j^*)$*.*

3.6 Poisson brackets on the surface of a sphere

The $\mathbb{R}^3$ bracket (3.27) for functions of the vector $\mathbf{Y}$ on the sphere $|\mathbf{Y}|^2 = const$ is expressible as

$$\begin{aligned}\{F,\,H\}(\mathbf{Y}) \;=\; 4Y_3\left(\frac{\partial F}{\partial Y_1}\frac{\partial H}{\partial Y_2}-\frac{\partial F}{\partial Y_2}\frac{\partial H}{\partial Y_1}\right)+4Y_1\left(\frac{\partial F}{\partial Y_2}\frac{\partial H}{\partial Y_3}-\frac{\partial F}{\partial Y_3}\frac{\partial H}{\partial Y_2}\right)\\ +\,4Y_2\left(\frac{\partial F}{\partial Y_3}\frac{\partial H}{\partial Y_1}-\frac{\partial F}{\partial Y_1}\frac{\partial H}{\partial Y_3}\right).\end{aligned} \tag{3.31}$$

This expression may be rewritten equivalently as

$$\{F,\,H\}\,d^{\,3}Y = dC\wedge dF\wedge dH\,,$$

with $C(\mathbf{Y}) = 2|\mathbf{Y}|^2$.

Exercise. In spherical coordinates,

$$Y_1 = r\cos\phi\sin\theta,\quad Y_2 = r\sin\phi\sin\theta,\quad Y_3 = r\cos\theta,$$

where $Y_1^2+Y_2^2+Y_3^2 = r^2$. The volume element is

$$d^{\,3}Y = dY_1\wedge dY_2\wedge dY_3 = \frac{1}{3}\,dr^3\wedge d\phi\wedge d\cos\theta\,.$$

Show that the area element on the sphere satisfies

$$r^2 d\phi\wedge d\cos\theta = \frac{r}{3}\left(\frac{dY_1\wedge dY_2}{Y_3}+\frac{dY_2\wedge dY_3}{Y_1}+\frac{dY_3\wedge dY_1}{Y_2}\right).$$

Explain why the canonical Poisson bracket

$$\{F,\,H\} = \left(\frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi}-\frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\right)$$

might be expected to be related to the $\mathbb{R}^3$ bracket in (3.31), up to a constant multiple. ★

On a constant level surface of r the functions (F, H) only depend on $(\cos\theta, \phi)$, so under the transformation to spherical coordinates we have

$$\begin{aligned}\{F,\,H\}\,d^3Y &= \frac{1}{3}\,dr^3 \wedge dF \wedge dH(\cos\theta, \phi) \\ &= \frac{1}{3}\,dr^3 \wedge \left(\frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi} - \frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\right) d\cos\theta \wedge d\phi\,.\end{aligned} \tag{3.32}$$

Consequently, on a level surface $r = const$, the Poisson bracket between two functions (F, H) depending only on $(\cos\theta, \phi)$ has symplectic form $\omega = d\cos\theta \wedge d\phi$, so that

$$\begin{aligned}dF \wedge dH(\cos\theta, \phi) &= \{F,\,H\}\,d\cos\theta \wedge d\phi \\ &= \left(\frac{\partial F}{\partial\cos\theta}\frac{\partial H}{\partial\phi} - \frac{\partial F}{\partial\phi}\frac{\partial H}{\partial\cos\theta}\right) d\cos\theta \wedge d\phi\,.\end{aligned} \tag{3.33}$$

Perhaps not surprisingly, the symplectic form on the surface of a unit sphere is its ***area element***.

3.7 Quotient map and orbit manifold of the 1:1 resonance

Definition

3.11 (Quotient map)**.** The $1:1$ ***quotient map*** $\pi : \mathbb{C}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{0\}$ is defined (for $R > 0$) by means of the formulas for the quadratic S^1 invariants in (3.19). This map may be expressed as

$$\mathbf{Y} := (Y_1 + iY_2,\, Y_3) = \pi(\mathbf{a})\,. \tag{3.34}$$

Explicitly, this is

$$\mathbf{Y} = \mathrm{tr}\,(\rho\,\boldsymbol{\sigma}) = a_k^*\boldsymbol{\sigma}_{kl}a_l\,, \tag{3.35}$$

also known as the ***Hopf map***.

Remark

3.12. Since $\{\mathbf{Y},\, R\} = 0$, the quotient map π in (3.34) collapses each $1:1$ orbit to a point. The converse also holds, namely that the inverse of the quotient map $\pi^{-1}\mathbf{Y}$ for $\mathbf{Y} \in \mathrm{Image}\,\pi$ consists of a 1:1 orbit (S^1).

Definition

3.13 (Orbit manifold)**.** The image in $\mathbb{R}^3$ of the quotient map $\pi : \mathbb{C}^2\backslash\{0\} \to \mathbb{R}^3\backslash\{0\}$ in (3.34) is the ***orbit manifold*** for the 1:1 resonance.

Remark

3.14 (1 : 1 orbit manifold is the Poincaré sphere)**.** *The image of the quotient map π in (3.34) may be conveniently displayed as the zero level set of the relation* $\det\rho = 0$ *using the S^1-invariant variables in Equation (3.21):*

$$
\begin{aligned}
\det\rho &= C(Y_1,\,Y_2,\,Y_3,\,R) \\
&:= (R+Y_3)(R-Y_3) - \big(Y_1^2+Y_2^2\big) = 0\,.
\end{aligned}
\tag{3.36}
$$

Consequently, each level set of R in the $1:1$ resonance map $\mathbb{C}^2 \to S^3$ obtained by transforming to S^1 invariants yields an orbit manifold defined by $C(Y_1,\,Y_2,\,Y_3,\,R) = 0$ in three dimensions.

For the 1:1 resonance, the image in $\mathbb{R}^3$ of the quotient map π in (3.34) is the ***Poincaré sphere*** *S^2.*

Remark

3.15 (Poincaré sphere $S^3 \simeq S^2 \times S^1$)**.** *Being invariant under the flow of the Hamiltonian vector field $X_R = \{\,\cdot\,,\,R\}$, each point on the Poincaré sphere S^2 consists of a resonant S^1 orbit under the 1:1 circle action:*

$$
\begin{aligned}
\phi_{1:1} : \mathbb{C}^2 \mapsto \mathbb{C}^2 \quad &\textit{as} \quad (a_1,a_2) \to (e^{i\phi}a_1, e^{i\phi}a_2) \\
&\textit{and} \quad (a_1^*,a_2^*) \to (e^{-i\phi}a_1^*, e^{-i\phi}a_2^*)\,.
\end{aligned}
\tag{3.37}
$$

3.8 The basic qubit: Quantum computing in the Bloch ball

In quantum mechanics, the Poincaré sphere is known as the Bloch sphere, and it corresponds to a two-level atomic system [FeVeHe1957]. In particular, the Hermitian coherence matrix ρ in (3.22) corresponds to the ***density matrix*** (also denoted as ρ) in quantum mechanics. The density matrix ρ is a key element in the applications of quantum mechanics, including quantum computing. In traditional computing, a ***bit*** is a scalar that can assume either of the values 0 or 1. In quantum computing, a ***qubit*** is a vector in a two-dimensional complex Hilbert space that can assume the states up and down, as well as all other states intermediate between them in a certain sense. The basic qubit in quantum computing is a two-level spin system whose density matrix ρ was first introduced in [FeVeHe1957] for describing maser dynamics.[2] This is the 2×2 positive Hermitian matrix with unit trace,

$$\rho = \frac{1}{2}\Big(\sigma_0 + \mathbf{r}\cdot\boldsymbol{\sigma}\Big) = \frac{1}{2}\left[\begin{array}{cc} 1+r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1-r_3 \end{array}\right], \tag{3.38}$$

so that $\mathbf{r} = \mathrm{tr}\,(\rho\,\boldsymbol{\sigma})$.

To be positive, the 2×2 Hermitian density matrix must have both positive trace and positive determinant. By construction, it has unit trace and its determinant

$$\det \rho = 1 - |\mathbf{r}|^2 \tag{3.39}$$

will also be positive, provided one requires

$$|\mathbf{r}|^2 = r_1^2 + r_2^2 + r_3^2 \le 1\,, \tag{3.40}$$

which defines the Bloch ball.

Definition

3.16 (Bloch ball)**.** *The* ***Bloch ball*** *is the locus of states for a given density matrix ρ in (3.38) satisfying $|\mathbf{r}| \le 1$.*

The determinant of the density matrix ρ vanishes for points on the surface of the ***Bloch sphere***, on which $|\mathbf{r}| = 1$. These are the ***pure states*** of the two-level system. For example, the north (respectively, south) pole of the Bloch sphere may be chosen to represent the pure up quantum state $(1,0)^T$ (respectively, down state $(0,1)^T$) for the two-state wave function [FeVeHe1957],

$$\psi(t) = a_1(t)\left[\begin{array}{c} 1 \\ 0 \end{array}\right] + a_2(t)\left[\begin{array}{c} 0 \\ 1 \end{array}\right] = \left[\begin{array}{c} a_1(t) \\ a_2(t) \end{array}\right].$$

The other points that are on the surface but away from the poles of the Bloch sphere represent a superposition of these pure states, with complex probability amplitude function $\psi(t)$, corresponding to probability

$$|\psi(t)|^2 = |a_1(t)|^2 + |a_2(t)|^2\,.$$

Restriction of the density matrix $\rho|_{r=1}$ to the surface of the Bloch sphere $|\mathbf{r}| = 1$ recovers the Hermitian matrix ρ in (3.15) whose vanishing determinant defines the unit Poincaré sphere, with $R = 1$.

Points for which $|\mathbf{r}| < 1$ are within the Bloch ball. These points represent ***impure states***, also known as ***mixed states***. The mixed states are distinguished from the pure states as follows.

---

[2]The acronym MASER stands for microwave amplification by stimulated emission of radiation.

Definition

3.17 (Pure vs mixed states). *A given state corresponding to density matrix* ρ *is* ***impure****, or* ***mixed****, if*

$$\operatorname{tr}\rho^2 = \frac{1}{2}(1+|\mathbf{r}|^2) < 1\,.$$

If $\operatorname{tr}\rho^2 = 1$*, then* $|\mathbf{r}| = 1$ *and the state is* ***pure****, or* ***unmixed****.*

Remark

3.18 (Optics vs quantum mechanics). *The quantum mechanical case deals with mixed and pure states inside and on the surface of the Bloch ball, respectively. The dynamics of optical polarisation directions on the Poincaré sphere deals exclusively with pure states.*

3.9 The action of $SU(2)$ on $\mathbb{C}^2$

The Lie group $SU(2)$ of complex 2×2 unitary matrices $U(s)$ with unit determinant acts on $\mathbf{a}\in\mathbb{C}^2$ by matrix multiplication as

$$\mathbf{a}(s) = U(s)\mathbf{a}(0) = \exp(is\xi)\mathbf{a}(0)\,.$$

Here, the quantity

$$i\xi = \Big[U'(s)U^{-1}(s)\Big]_{s=0} \in su(2)$$

is a 2×2 traceless skew-Hermitian matrix, $(i\xi)^\dagger = -(i\xi)$. This conclusion follows from unitarity,

$$UU^\dagger = Id\,,\quad \text{which implies}\quad U'U^\dagger + UU'^{\,\dagger} = 0 = U'U^\dagger + (U'U^\dagger)^\dagger\,.$$

Consequently, ξ alone (without multiple i) is a 2×2 traceless Hermitian matrix, $\xi^\dagger = \xi$, which may be written as a sum over the Pauli matrices in (3.16) as

$$\xi_{kl} = \sum_{j=1}^{3}\xi^j(\sigma_j)_{kl} \quad\text{with}\quad k,l = 1,2\,.$$

The corresponding vector field $\xi_U(\mathbf{a})\in T\mathbb{C}^2$ may be expressed as a Lie derivative,

$$\xi_U(\mathbf{a}) = \frac{d}{ds}\big[\exp(is\xi)\mathbf{a}\big]\Big|_{s=0} =: \pounds_{\xi_U}\mathbf{a} = i\xi\mathbf{a}\,,$$

in which the product $(\xi\mathbf{a})$ of the Hermitian matrix (ξ) and the two-component complex vector $(\mathbf{a})$ has components $\xi_{kl}a_l$, with $k,l = 1,2$.

Definition

3.19 (Momentum map $J:\, T^*\mathbb{C}^2 \mapsto su(2)^*$). *The **momentum map***

$$J(\mathbf{a}^*,\mathbf{a}) :\, T^*\mathbb{C}^2 \simeq \mathbb{C}^2\times\mathbb{C}^2 \mapsto su(2)^*$$

for the action of $SU(2)$ on $\mathbb{C}^2$ is defined by

$$\begin{aligned} J^\xi(\mathbf{a}^*,\mathbf{a}) &:= \Big\langle J(\mathbf{a}^*,\mathbf{a}),\, \xi \Big\rangle_{su(2)^*\times su(2)} = \Big\langle\!\Big\langle \mathbf{a}^*,\, \pounds_{\xi_U}\mathbf{a} \Big\rangle\!\Big\rangle_{\mathbb{C}^2\times\mathbb{C}^2} \\ &:= a_k^*\xi_{kl}a_l = a_l a_k^* \xi_{kl} \\ &:= \operatorname{tr}\big((\mathbf{a}\otimes\mathbf{a}^*)\xi\big) = \operatorname{tr}\big(\rho\xi\big)\,. \end{aligned} \tag{3.41}$$

Note: In these expressions we treat $\mathbf{a}^*\in\mathbb{C}^2$ as $\mathbf{a}^*\in T^*_{\mathbf{a}}\mathbb{C}^2$.

Remark

3.20. *This map may also be expressed using the canonical symplectic form, $\Omega(\mathbf{a},\mathbf{b}) = \operatorname{Im}(\mathbf{a}^*\cdot\mathbf{b})$ on $\mathbb{C}^2$, as in [MaRa1994],*

$$\begin{aligned} J^\xi(\mathbf{a}^*,\mathbf{a}) &:= \Omega(\mathbf{a},\xi_U(\mathbf{a})) = \Omega(\mathbf{a}, i\xi\mathbf{a}) \\ &= \operatorname{Im}(a_k^*(i\xi)_{kl}a_l) = a_k^*\xi_{kl}a_l = \operatorname{tr}\big(\rho\xi\big)\,. \end{aligned} \tag{3.42}$$

Remark

3.21 (Removing the trace). *Being traceless, ξ has zero pairing with any multiple of the identity; so one may remove the trace of ρ by subtracting $\operatorname{tr}\rho$ times a multiple of the identity. Thus, the momentum map*

$$J(\mathbf{a}^*,\mathbf{a}) = \tilde\rho = \rho - \frac{1}{2}Id\operatorname{tr}\rho \in su(2)^* \tag{3.43}$$

sends $(\mathbf{a}^*,\mathbf{a})\in\mathbb{C}^2\times\mathbb{C}^2$ to the traceless part $\tilde\rho$ of the Hermitian matrix $\rho = \mathbf{a}\otimes\mathbf{a}^*$, which is an element of $su(2)^*$. Here, $su(2)^*$ is the dual space to $su(2)$ under the pairing $\langle\,\cdot\,,\,\cdot\,\rangle :\, su(2)^*\times su(2)\mapsto\mathbb{R}$ given by the trace of the matrix product,

$$\Big\langle \tilde\rho,\,\xi \Big\rangle = \operatorname{tr}\big(\tilde\rho\xi\big) \quad \textit{for} \quad \tilde\rho\in su(2)^* \textit{ and } \xi\in su(2)\,. \tag{3.44}$$

Remark

3.22 (Momentum map and Poincaré sphere). *A glance at Equation (3.17) reveals that the momentum map $\mathbb{C}^2 \times \mathbb{C}^2 \mapsto su(2)^*$ for the action of $SU(2)$ acting on $\mathbb{C}^2$ in Equation (3.43) is none other than the map $\mathbb{C}^2 \mapsto S^2$ to the Poincaré sphere. To see this, one simply replaces $\xi \in su(2)$ by the vector of Pauli matrices $\boldsymbol{\sigma}$ in Equation (3.17) to recover the Hopf map,*

$$\tilde{\rho} = \rho - \frac{1}{2}R\sigma_0 = \frac{1}{2}\mathbf{Y}\cdot\boldsymbol{\sigma}\,, \tag{3.45}$$

in which

$$\mathbf{Y} = \operatorname{tr}\tilde{\rho}\,\boldsymbol{\sigma} = a_k\boldsymbol{\sigma}_{kl}a_l^*\,. \tag{3.46}$$

Thus, the traceless momentum map, $J : T^*\mathbb{C}^2 \simeq \mathbb{C}^2 \times \mathbb{C}^2 \mapsto su(2)^*$, in Equation (3.43) explicitly recovers the components of the vector $\mathbf{Y}$ on the Poincaré sphere $|\mathbf{Y}|^2 = R^2$. Namely,

$$\tilde{\rho} = \frac{1}{2}\begin{bmatrix} Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & -Y_3 \end{bmatrix}. \tag{3.47}$$

Remark

3.23 (Poincaré sphere and Hopf fibration of S^3). *The sphere $|\mathbf{Y}|^2 = R^2$ defined by the map $\mathbb{C}^2/S^1 \mapsto S^2 \simeq S^3/S^1$ for $1:1$ resonance was first introduced by Poincaré to describe the two transverse polarisation states of light [Po1892,BoWo1965]. It was later studied by Hopf [Ho1931], who showed that it has interesting topological properties. Namely, it is a* ***fibration*** *of $S^3 \simeq SU(2)$. That is, the Poincaré sphere $S^2 = S^3/S^1$ has an S^1 fibre sitting over every point on the sphere [Ho1931]. This reflects the $SU(2)$ group decomposition which locally factorises into $S^2 \times S^1$ at each point on the sphere S^2. For $R = 1$, this may be expressed as a matrix factorisation for any $A \in SU(2)$ with $s, \phi \in [0, 2\pi)$ and $\theta \in [0, \pi/2)$, as*

$$\begin{aligned} A &= \begin{pmatrix} a_1^* & -a_2 \\ a_2^* & a_1 \end{pmatrix} \\ &= \begin{pmatrix} -i\cos\theta & -\sin\theta e^{i\phi} \\ \sin\theta e^{-i\phi} & i\cos\theta \end{pmatrix}\begin{pmatrix} \exp(-is) & 0 \\ 0 & \exp(is) \end{pmatrix}. \end{aligned} \tag{3.48}$$

For further discussion of the ***Hopf fibration*** *in the context of geometric dynamics for other resonant coupled oscillators (for example, in resonant frequency ratios 1:2. 1:3, 2:3, etc.), see [Ho2011].*

4 The matrix Lie group $Sp(2)$

Introduce matrix Lie group $Sp(2)$ and its matrix Lie algebra $\mathfrak{sp}(2)$

Consider our goal: understanding coadjoint motion on $\mathfrak{sp}(2)^*$

Study the transformations in $Sp(2)$ and their infinitesimal generators, the Hamiltonian matrices $\{m_0, m_1, m_2, m_3\} \in \mathfrak{sp}(2)$. The latter are related to co-quaternions.

Write infinitesimal transformations of the matrix Lie algebra $\mathfrak{sp}(2)$ acting on $\mathbf{z} = (\boldsymbol{q}, \boldsymbol{p})^T \in T^*\mathbb{R}^2$. (Note that this action leaves $\boldsymbol{p} \times \boldsymbol{q}$ invariant.)

Study the Adjoint orbits $\mathrm{Ad} : Sp(2) \times \mathfrak{sp}(2) \to \mathfrak{sp}(2)$

Define the momentum map $\mathcal{J} : T^*\mathbb{R}^2 \to \mathfrak{sp}(2)^*$ and the Hamiltonian $\mathcal{J}^\xi = \langle \mathcal{J}, \xi \rangle$ for the infinitesimal transformations $\dot{\mathbf{z}} = \{\mathbf{z}, \mathcal{J}^\xi\}$.

Derive the momentum map $\mathcal{J}$ from Hamilton's principle $\delta S = 0$ with $S = \int \ell(\xi) dt$, constrained by the action of $T^*\mathbb{R}^2$ under $Sp(2)$, thereby finding with right invariant $\xi = \dot{M}M^{-1}$

$$\frac{\partial \ell}{\partial \xi} = \mathcal{J}(\mathbf{z}) \qquad \dot{\mathbf{z}} = \xi \mathbf{z} \quad \Longrightarrow \quad \frac{d\mathcal{J}}{dt} + \mathrm{ad}^*_\xi \mathcal{J} = 0 \quad \text{or} \quad \mathrm{Ad}^*_{M^{-1}(t)} \mathcal{J}(t) = \mathcal{J}(0)$$

Write the components of the momentum map $\mathcal{J}(\mathbf{z}) \in \mathfrak{sp}(2)^*$ in terms of $\mathbf{Y} \in \mathbb{R}^3$, with

$$\mathcal{J} = \mathbf{z} \otimes \mathbf{z}^T J = 2 \begin{bmatrix} Y_3 & -Y_1 \\ Y_2 & -Y_3 \end{bmatrix} \qquad \mathbf{Y} = (Y_1, Y_2, Y_3) = (|\boldsymbol{q}|^2, |\boldsymbol{p}|^2, \boldsymbol{p} \cdot \boldsymbol{q})$$

noting that $S^2(\mathbf{Y}) := Y_1 Y_2 - Y_3^2 = |\boldsymbol{p} \times \boldsymbol{q}|^2$ in which $\boldsymbol{p} \times \boldsymbol{q}$ is invariant under $Sp(2)$.

Explain how $\mathbb{R}^3$-bracket dynamics of $\dot{\mathbf{Y}} = \nabla S^2 \times \nabla H$ is related to coadjoint motion on $\mathfrak{sp}(2)^*$.

5 The matrix Lie group $Sp(2)$ and its matrix Lie algebra $\mathfrak{sp}(2)$

Symplectic 2×2 matrices $M(s) \in Sp(2)$ depending smoothly on a real parameter $s \in \mathbb{R}$ satisfy

$$M(s)JM(s)^T = J \tag{5.1}$$

where

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{5.2}$$

Set $\xi(s) = \dot{M}(s)M^{-1}(s)$, so that $\dot{M}(s) = \xi(s)M(s)$ is the reconstruction equation. Take d/ds of the defining relation

$$\begin{aligned}\frac{d}{ds}\left(M(s)JM(s)^T\right) &= \xi J + J\xi^T \\ &= \xi J + (\xi J^T)^T \\ &= \xi J - (\xi J)^T = 0\end{aligned}$$

Thus, $\xi J = (\xi J)^T \in sym$ is symmetric.

- Conjugation by J shows that $J\xi = (J\xi)^T \in sym$ is also symmetric.
- Replacing $M \leftrightarrow M^T$ gives the corresponding result for left invariant $a(s) := M^{-1}(s)\dot{M}(s)$.

5.1 Examples of matrices in $Sp(2)$

One easily checks that the following are $Sp(2)$ matrices

$$M_1(\tau_1) = \begin{bmatrix} 0 & -1 \\ -2\tau_1 & 1 \end{bmatrix}, \qquad M_3(\tau_3) = \begin{bmatrix} e^{\tau_3} & 0 \\ 0 & e^{-\tau_3} \end{bmatrix}, \qquad M(\omega) = \begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix}.$$

The product is also an $Sp(2)$ matrix

$$M(\tau_1, \tau_3, \omega) = M(\omega)M_3(\tau_3)M_1(\tau_1) = \begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} e^{\tau_3} & 0 \\ 0 & e^{-\tau_3} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -2\tau_1 & 1 \end{bmatrix} \tag{5.3}$$

This is a general result, called the ***Iwasawa decomposition*** of the symplectic matrix group, usually written as

$$Sp(2, \mathbb{R}) = \mathsf{KAN}\,.$$

The rightmost matrix factor represents the nilpotent subgroup N. The middle factor is the abelian subgroup A. The leftmost factor is the maximal compact subgroup K.

The KAN decomposition (5.3) also follows by exponentiation

$$M(\tau_1, \tau_3, \omega) = M(\omega)M_3(\tau_3)M_1(\tau_1) = e^{\frac{1}{2}\omega(m_1+m_2)}e^{\tau_3 m_3}e^{\tau_1 m_1} \tag{5.4}$$

of the 2×2 Hamiltonian matrices

$$m_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad m_1 = \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix}, \quad m_2 = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad m_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5.5}$$

These are related to the Pauli matrices and thus to the quaternions by

$$m_0 = \sigma_0, \quad m_1 = -\sigma_1 + i\sigma_2, \quad m_2 = \sigma_1 + i\sigma_2, \quad m_3 = \sigma_3, \tag{5.6}$$

so that $\frac{1}{2}(m_1 + m_2) = i\sigma_2$.

5.2 Hamiltonian matrices and co-quaternions

One may define a co-quaternion C and its conjugate C^* as

$$C = wm_0 + xm_1 + ym_2 + zm_3 = \begin{bmatrix} w+z & 2y \\ -2x & w-z \end{bmatrix} \qquad C^* = \begin{bmatrix} w-z & -2y \\ 2x & w+z \end{bmatrix}$$

Then $\operatorname{tr} CC^* = 2(w^2 - z^2) + 8xy$ is a hyperbolic relation that leads to co-quaternions whose multiplication law is obtained from the relation (5.6) between the Hamiltonian basis matrices and the Pauli matrices.

5.3 Infinitesimal transformations

Definition

5.1. The infinitesimal transformation of the $Sp(2)$ matrix Lie group acting on the manifold $T^*\mathbb{R}^2$ is a vector field $\xi_M(\mathbf{z}) \in T\mathbb{R}^2$ that may be expressed as the derivative of the group transformation, evaluated at the identity,

$$\xi_M(\mathbf{z}) = \frac{d}{ds}\Big[\exp(s\xi)\mathbf{z}\Big]\Big|_{s=0} = \xi\mathbf{z}\,. \tag{5.7}$$

Here, the diagonal action of the Hamiltonian matrix (ξ) and the two-component real multi-vector $\mathbf{z} = (\mathbf{q},\, \mathbf{p})^T$ (denoted as $\xi\mathbf{z}$) has components given by $(\xi_{kl}q_l, \xi_{kl}p_l)^T$, with $k, l = 1, 2$. The matrix ξ is any linear combination of the traceless constant Hamiltonian matrices (5.5).

Examples The action of $Sp(2)$ on $T^*\mathbb{R}^2$ is obtained from the infinitesimal actions, expressed in terms of the Hamiltonian matrices, as

$$m_1: \quad \frac{d}{d\tau_1}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}0 & 0\\ -2 & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}0\\ -2\boldsymbol{q}\end{bmatrix} \quad\Longrightarrow\quad \begin{bmatrix}\boldsymbol{q}(\tau_1)\\ \boldsymbol{p}(\tau_1)\end{bmatrix} = \begin{bmatrix}1 & 0\\ -2\tau_1 & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{q}(0)\\ \boldsymbol{p}(0)\end{bmatrix}$$

$$m_2: \quad \frac{d}{d\tau_2}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}2\boldsymbol{p}\\ 0\end{bmatrix} \quad\Longrightarrow\quad \begin{bmatrix}\boldsymbol{q}(\tau_2)\\ \boldsymbol{p}(\tau_2)\end{bmatrix} = \begin{bmatrix}1 & 2\tau_2\\ 0 & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{q}(0)\\ \boldsymbol{p}(0)\end{bmatrix}$$

$$m_3: \quad \frac{d}{d\tau_3}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}\boldsymbol{q}\\ -\boldsymbol{p}\end{bmatrix} \quad\Longrightarrow\quad \begin{bmatrix}\boldsymbol{q}(\tau_3)\\ \boldsymbol{p}(\tau_3)\end{bmatrix} = \begin{bmatrix}e^{\tau_3} & 0\\ 0 & e^{-\tau_3}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}(0)\\ \boldsymbol{p}(0)\end{bmatrix}$$

$$\tfrac{1}{2}(m_1+m_2): \quad \frac{d}{d\omega}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\\ \boldsymbol{p}\end{bmatrix} = \begin{bmatrix}\boldsymbol{p}\\ -\boldsymbol{q}\end{bmatrix} \quad\Longrightarrow\quad \begin{bmatrix}\boldsymbol{q}(\tau_3)\\ \boldsymbol{p}(\tau_3)\end{bmatrix} = \begin{bmatrix}\cos\omega & \sin\omega\\ -\sin\omega & \cos\omega\end{bmatrix}\begin{bmatrix}\boldsymbol{q}(0)\\ \boldsymbol{p}(0)\end{bmatrix}$$

The last is obtained easily by writing $\frac{d}{d\omega}(\boldsymbol{q}+i\boldsymbol{p}) = -i(\boldsymbol{q}+i\boldsymbol{p})$.

Thus, as claimed, the KAN decomposition of a symplectic matrix M may be written in terms of the Hamiltonian matrices as

$$M = e^{\frac{1}{2}\omega(m_1+m_2)}e^{\tau_3 m_3}e^{\tau_1 m_1}\,. \tag{5.8}$$

Remark

5.2. Notice that all of these $Sp(2)$ transformations leave invariant the cross-product $\boldsymbol{p}\times\boldsymbol{q}$.

The quantity $\boldsymbol{p}\times\boldsymbol{q}$ is called Lagrange's invariant in the study of geometric optics, for which the linear symplectic transformations play a key role.

Remark

5.3. Under the matrix commutator $[m_i, m_j] := m_i m_j - m_j m_i$, the Hamiltonian matrices m_i with $i = 1,2,3$ close among themselves, as

$$[m_1, m_2] = 4m_3\,, \quad [m_2, m_3] = -\,2m_2\,, \quad [m_3, m_1] = -\,2m_1\,. \tag{5.9}$$

6 Adjoint orbits of the action of $Sp(2)$ on $\mathfrak{sp}(2)$

We consider the Hamiltonian matrix

$$\begin{aligned} m(\omega,\gamma,\tau) &= \frac{\omega}{2}(m_1+m_2) + \frac{\gamma}{2}(m_2-m_1) + \tau m_3 \\ &= \begin{bmatrix}\tau & \gamma+\omega\\ \gamma-\omega & -\tau\end{bmatrix}, \end{aligned} \tag{6.1}$$

which may be regarded as a pure co-quaternion.

- The eigenvalues of the Hamiltonian matrix (6.1) are determined from

$$\lambda^2 + \Delta = 0\,, \quad \text{with} \quad \Delta = \det m = \omega^2 - \gamma^2 - \tau^2\,. \tag{6.2}$$

Consequently, the eigenvalues come in pairs, given by

$$\lambda^\pm = \pm\sqrt{-\Delta} = \pm\sqrt{\tau^2 + \gamma^2 - \omega^2}\,. \tag{6.3}$$

- Adjoint orbits in the space $(\gamma + \omega, \gamma - \omega, \tau) \in \mathbb{R}^3$ are obtained from the action of a symplectic matrix $M(\tau_i)$ on a Hamiltonian matrix $m(\omega, \gamma, \tau)$ by matrix conjugation

$$m \to m' = M(\tau_i) m M^{-1}(\tau_i) \quad \text{(no sum on } i = 1, 2, 3)$$

The values of the parameters (ω, γ, τ) in (6.1) may change along the Adjoint orbits, but the eigenvalues will not change.

- Since Adjoint action preserves eigenvalues, it preserves the value of the determinant Δ. This means the Adjoint orbits of the Hamiltonian matrices lie on the level sets of the determinant Δ.

- The Adjoint orbits of the Hamiltonian matrices corresponding to these eigenvalues are of different type, depending on whether $\Delta < 0$ (hyperbolic), $\Delta = 0$ (parabolic), or $\Delta > 0$ (elliptic), as illustrated in Figure 4 and summarised in the table below.

Harmonic (elliptic) orbit	*Trajectories: Ellipses*
$\Delta = 1\,, \quad \lambda^\pm = \pm\, i$	$m_H = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$
Free (parabolic) orbit	*Trajectories: Straight lines*
$\Delta = 0\,, \quad \lambda^\pm = 0$	$m_H = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$
Repulsive (hyperbolic) orbit	*Trajectories: Hyperbolas*
$\Delta = -1\,, \quad \lambda^\pm = \pm\, 1$	$m_H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$

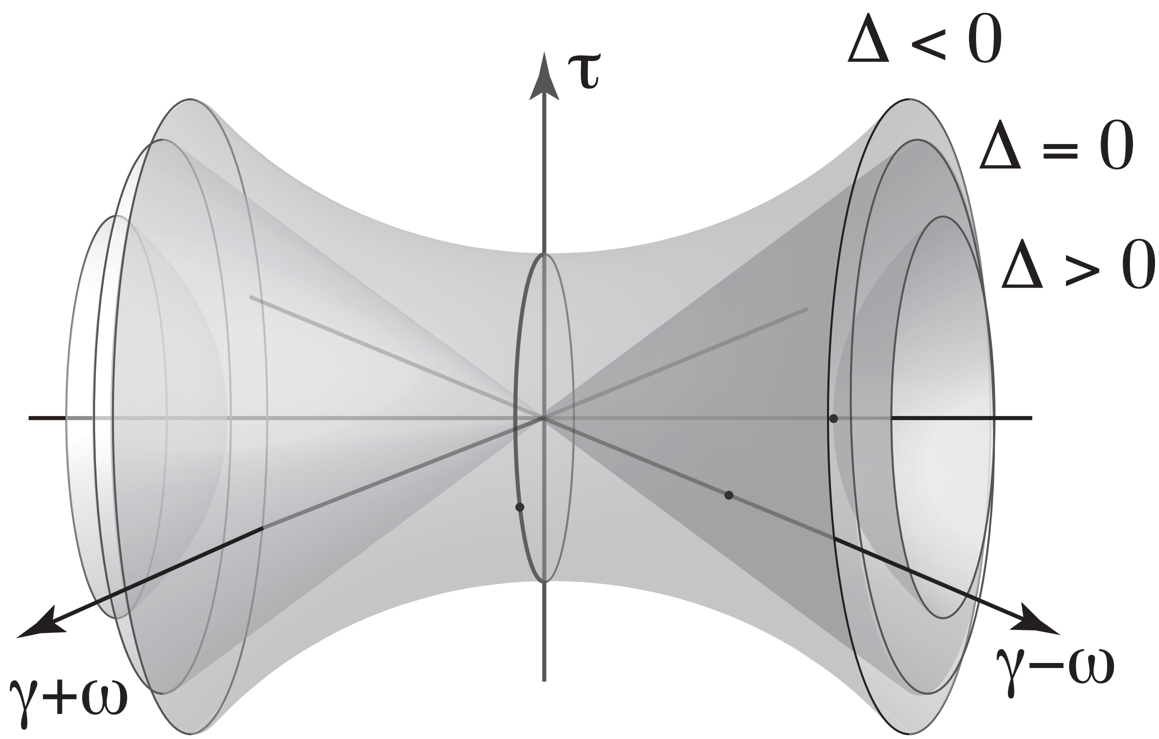

Figure 4: The action by matrix conjugation of a symplectic matrix on a Hamiltonian matrix changes its parameters $(\omega, \gamma, \tau) \in \mathbb{R}^3$, while preserving the value of the discriminant $\Delta = \omega^2 - \gamma^2 - \tau^2$. Three families of orbits emerge, that are hyperbolic ($\Delta < 0$), parabolic ($\Delta = 0$) and elliptic ($\Delta > 0$).

7 The momentum map and coadjoint orbits

Definition

7.1. *The momentum map* $\mathcal{J}: T^*\mathbb{R}^2 \simeq \mathbb{R}^2\times\mathbb{R}^2 \to \mathfrak{sp}(2)^*$ *is defined by*

$$\begin{aligned}\mathcal{J}^\xi(\mathbf{z}) &:= \Big\langle \mathcal{J}(\mathbf{z}),\, \xi \Big\rangle_{sp(2,\mathbb{R})^*\times sp(2,\mathbb{R})}\\ &= \Big(\mathbf{z},\, J\xi\mathbf{z}\Big)_{\mathbb{R}^2\times\mathbb{R}^2}\\ &:= z_k(J\xi)_{kl}z_l\\ &= \mathbf{z}^T\cdot J\xi\mathbf{z}\\ &= \operatorname{tr}\Big((\mathbf{z}\otimes\mathbf{z}^TJ)\xi\Big)\,,\end{aligned} \tag{7.1}$$

where $\mathbf{z} = (\mathbf{q},\,\mathbf{p})^T \in \mathbb{R}^2\times\mathbb{R}^2$.

Remark

7.2. The map $\mathcal{J}(\mathbf{z})$ given in (7.1) by

$$\mathcal{J}(\mathbf{z}) = (\mathbf{z}\otimes\mathbf{z}^TJ) \in \mathfrak{sp}(2)^* \tag{7.2}$$

sends $\mathbf{z} = (\mathbf{q},\,\mathbf{p})^T \in \mathbb{R}^2\times\mathbb{R}^2$ to $\mathcal{J}(\mathbf{z}) = (\mathbf{z}\otimes\mathbf{z}^TJ) \in \mathfrak{sp}(2)^*$.

7.1 Deriving the momentum map (7.1) from Hamilton's principle

Proposition

7.3. *The momentum map (7.1) may be derived from Hamilton's principle.*

Proof. Let $\xi\in\mathfrak{sp}(2)$ act on $\mathbf{z}\in T^*\mathbb{R}^2$ as in formula (5.7), namely,

$$\frac{d\mathbf{z}}{dt} = \xi\mathbf{z}\,,$$

and let $\ell(\xi)$ be a reduced Lagrangian in Hamilton's principle $\delta S = 0$ with $S = \int \ell(\xi)dt$, constrained by the action of $T^*\mathbb{R}^2$ under $Sp(2)$. Taking the variations yields

$$\begin{aligned}0 = \delta S &= \delta\int \ell(\xi) + \mathbf{z}^T\cdot J(\dot{\mathbf{z}} - \xi\mathbf{z})\,dt\\ &= \int \Big\langle \frac{\partial\ell}{\partial\xi} - \mathbf{z}\otimes\mathbf{z}^TJ\,,\,\delta\xi\Big\rangle + \delta\mathbf{z}^T\cdot J(\dot{\mathbf{z}}-\xi\mathbf{z}) - \delta\mathbf{z}\cdot J(\dot{\mathbf{z}}-\xi\mathbf{z})^T\,dt\end{aligned}$$

thereby finding $\frac{\partial \ell}{\partial \xi} = \mathbf{z} \otimes \mathbf{z}^T J = \mathcal{J}(\mathbf{z})$. □

Proposition

7.4. *Suppose the Lagrangian $\ell(\xi)$ is hyper-regular, so that one may solve $\frac{\partial \ell}{\partial \xi}$ for ξ. Then, applying the time derivative to $\mathcal{J}(\mathbf{z})$ and using the Lie algebra action shows that the momentum map $\mathcal{J}(\mathbf{z})$ undergoes coadjoint motion,*

$$\frac{d\mathbf{z}}{dt} = \xi \mathbf{z} \quad \Longrightarrow \quad \frac{d\mathcal{J}}{dt} + \mathrm{ad}^*_\xi \mathcal{J} = 0 \quad \textit{or, equivalently,} \quad \mathrm{Ad}^*_{M^{-1}(t)} \mathcal{J}(t) = \mathcal{J}(0).$$

Proof. The proof of this proposition is not given in these lecture notes. Instead, we shall show that each component of the momentum map undergoes coadjoint motion. □

7.2 Components of the momentum map

The map $\mathcal{J} : T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2 \to \mathfrak{sp}(2)^*$ in (7.1) for $Sp(2,\mathbb{R})$ acting diagonally on $\mathbb{R}^2 \times \mathbb{R}^2$ in Equation (7.2) may be expressed in matrix form as

$$\begin{aligned} \mathcal{J} &= (\mathbf{z} \otimes \mathbf{z}^T J) \\ &= 2 \begin{pmatrix} \mathbf{p} \cdot \mathbf{q} & -|\mathbf{q}|^2 \\ |\mathbf{p}|^2 & -\mathbf{p} \cdot \mathbf{q} \end{pmatrix} \\ &= 2 \begin{pmatrix} Y_3 & -Y_1 \\ Y_2 & -Y_3 \end{pmatrix}, \end{aligned} \tag{7.3}$$

consisting of rotation-invariant components,

$$T^*\mathbb{R}^2 \to \mathbb{R}^3 : \, (\mathbf{q}, \mathbf{p})^T \to \mathbf{Y} = (Y_1, Y_2, Y_3)\,,$$

defined as

$$Y_1 = |\mathbf{q}|^2 \geq 0\,, \quad Y_2 = |\mathbf{p}|^2 \geq 0\,, \quad Y_3 = \mathbf{p} \cdot \mathbf{q}\,. \tag{7.4}$$

Remark

7.5. Under the pairing $\langle \cdot\,,\,\cdot \rangle : \mathfrak{sp}(2)^* \times \mathfrak{sp}(2) \to \mathbb{R}$ given by the trace of the matrix product, one finds the Hamiltonian, or phase-space function,

$$\Big\langle \mathcal{J}(\mathbf{z}),\, \xi \Big\rangle = \mathrm{tr}\,\big(\mathcal{J}(\mathbf{z})\, \xi\big)\,, \tag{7.5}$$

for $\mathcal{J}(\mathbf{z}) = (\mathbf{z} \otimes \mathbf{z}^T J) \in \mathfrak{sp}(2)^*$ and $\xi \in \mathfrak{sp}(2)$.

Applying the momentum map $\mathcal{J}$ to the vector of Hamiltonian matrices $\mathbf{m} = (m_1,\, m_2,\, m_3)$ in equation (5.5) yields the individual components in (7.3), as

$$\mathcal{J} \cdot \mathbf{m} = 2\mathbf{Y} \quad \Longleftrightarrow \quad \mathbf{Y} = \frac{1}{2}\, z_k (J\mathbf{m})_{kl} z_l \,. \tag{7.6}$$

Remark

7.6.

(a) The components of the vector $\mathbf{Y} = (Y_1,\, Y_2,\, Y_3)$ satisfy the relation $S^2 = Y_1Y_2 - Y_3^2 = |\boldsymbol{p} \times \boldsymbol{q}|^2$ in which $\boldsymbol{p} \times \boldsymbol{q}$ is invariant under $Sp(2)$. Thus, coadjoint motion lies on level sets of S^2.

(b) The canonical Poisson brackets among the components $(Y_1,\, Y_2,\, Y_3)$ reflect the matrix commutation relations in (5.9),

$$[m_i,\, m_j] = \begin{array}{|c|ccc|} \hline [\,\cdot\,,\,\cdot\,] & m_1 & m_2 & m_3 \\ \hline m_1 & 0 & 4m_3 & 2m_1 \\ m_2 & -4m_3 & 0 & -2m_2 \\ m_3 & -2m_1 & 2m_2 & 0 \\ \hline \end{array} = c_{ij}^k m_k \,. \tag{7.7}$$

In tabular form, these Poisson brackets are

$$\{Y_i,\, Y_j\} = \begin{array}{|c|ccc|} \hline \{\,\cdot\,,\,\cdot\,\} & Y_1 & Y_2 & Y_3 \\ \hline Y_1 & 0 & 4Y_3 & 2Y_1 \\ Y_2 & -4Y_3 & 0 & -2Y_2 \\ Y_3 & -2Y_1 & 2Y_2 & 0 \\ \hline \end{array} = -\,\epsilon_{ijk}\frac{\partial S^2}{\partial Y_k} \,. \tag{7.8}$$

That is, the canonical Poisson brackets of the components of the momentum map close among themselves.

$$\{Y_1, Y_2\} = 4Y_3\,, \quad \{Y_2, Y_3\} = -\,2Y_2\,, \quad \{Y_3, Y_1\} = -\,2Y_1 \quad \text{and} \quad \{Y_i, S^2\} = 0 \quad \text{for} \quad i = 1, 2, 3\,.$$

In a moment we will show that these Poisson brackets among the components $(Y_1,\, Y_2,\, Y_3)$ satisfy the Jacobi identity.

(c) The Poisson bracket relations (7.8) among the components of the momentum map $\mathcal{J}$ imply the following for smooth functions F and H of the vector $\mathbf{Y} = (Y_1,\, Y_2,\, Y_3)$

$$\frac{dF}{dt} = \{F, H\} = -\nabla S^2 \cdot \nabla F \times \nabla H(\mathbf{Y}) = Y_k\, c^k_{ij}\, \frac{\partial F}{\partial Y_i} \frac{\partial H}{\partial Y_i} = -\left\langle \mathbf{Y}\,,\, \mathrm{ad}_{\frac{\partial H}{\partial \mathbf{Y}}} \frac{\partial F}{\partial \mathbf{Y}} \right\rangle = -\left\langle \mathrm{ad}^*_{\frac{\partial H}{\partial \mathbf{Y}}} \mathbf{Y}\,,\, \frac{\partial F}{\partial \mathbf{Y}} \right\rangle$$

which is the Lie-Poisson bracket on $\mathfrak{sp}(2)^*$. (The Lie-Poisson bracket satisfies the Jacobi identity by virtue of being dual to a Lie algebra.)

(d) This means that the momentum map (7.3) is Poisson.

(e) In particular, the components of the momentum map in vector form $\mathbf{Y}$ satisfy the equation for coadjoint motion

$$\frac{d\mathbf{Y}}{dt} = \nabla S^2 \times \nabla H(\mathbf{Y}) = -\,\mathrm{ad}^*_{\frac{\partial H}{\partial \mathbf{Y}}} \mathbf{Y}$$

for any choice of Hamiltonian $H(\mathbf{Y})$ that depends on the momentum map components.

(f) Thus, each component of the momentum map in (7.3) undergoes coadjoint motion.

Exercise. What corresponds to the quotient map, orbit manifold (image of the quotient map) and Poincaré sphere for the transmission of optical rays by Fermat's principle in an axisymmetric, translation-invariant medium?

Hint: This problem is discussed in the first chapter of [Ho2011GM1]. Use co-quaternions as representations of the symplectic group $Sp(2)$, as discussed in the third chapter of [Ho2011GM2]. ★

8 AD, Ad, and ad operations for Lie algebras and groups

The notation for the conjugacy relations among the quaternions in Section 2.6 follows the standard notation for the corresponding actions of a Lie group on itself, on its Lie algebra (its tangent space at the identity), the action of the Lie algebra on itself, and their dual actions. By the isomorphism between the quaternions and the matrix Lie group $G = SU(2)$, one may define these corresponding operations in general, for other ***matrix Lie groups***.

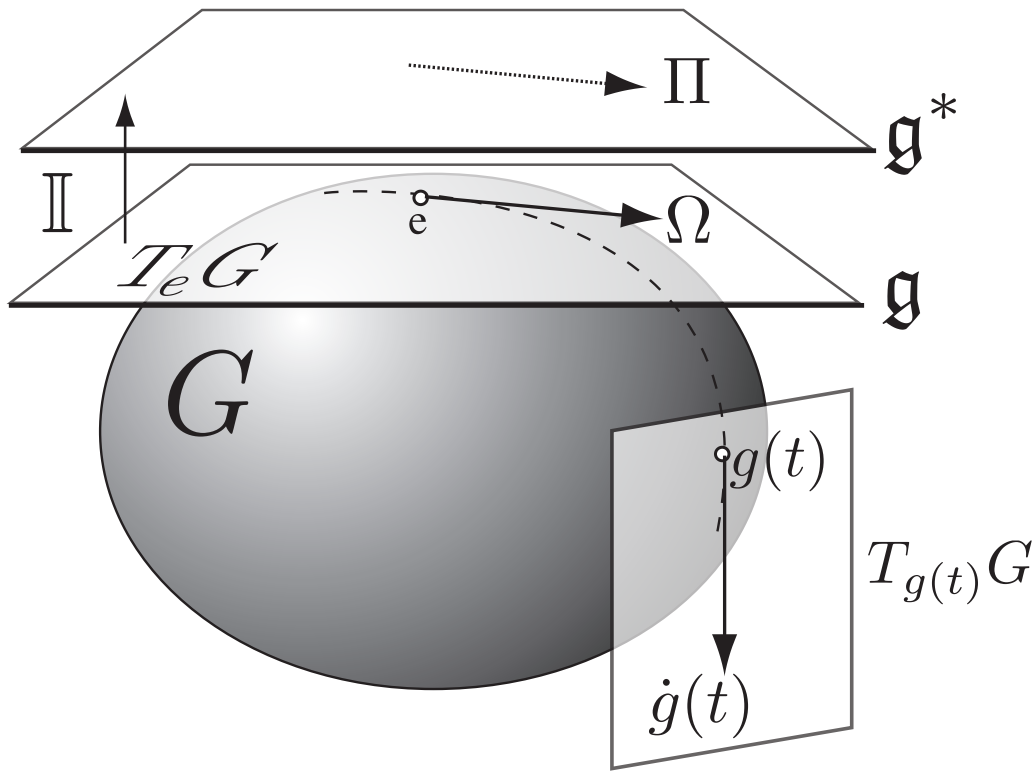

Figure 5: The tangent space at the identity e of the group G is its Lie algebra $\mathfrak{g}$, a vector space represented here as a plane. The moment of inertia $\mathbb{I}$ maps the vector $\Omega \in \mathfrak{g}$ into the dual vector $\Pi = \mathbb{I}\Omega \in \mathfrak{g}^*$. The dual Lie algebra $\mathfrak{g}^*$ is another vector space, also represented as a plane in the figure. A group orbit in G has tangent vector $\dot{g}(t)$ at point $g(t)$ which may be transported back to the identity by acting with $g^{-1}(t) \in G$ from either the left as $\Omega = g^{-1}(t)\dot{g}(t)$ or the right as $\omega = \dot{g}(t)g^{-1}(t)$.

8.1 ADjoint, Adjoint and adjoint operations for matrix Lie groups

- AD (conjugacy classes of a matrix Lie group): The map $I_g : G \to G$ given by $I_g(h) \to ghg^{-1}$ for matrix Lie group elements $g, h \in G$ is the ***inner automorphism*** associated with g. Orbits of this action are called ***conjugacy classes***.

$$\mathrm{AD} : G \times G \to G : \quad \mathrm{AD}_g h := ghg^{-1} .$$

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

$$\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g} : \quad \mathrm{Ad}_g \, \eta = T_e I_g \, \eta =: g\eta g^{-1} ,$$

  with $\eta = h'(0)$.

- Differentiate $\mathrm{Ad}_g \, \eta$ with respect to g at $g = e$ in the direction ξ to produce the ***adjoint operation***,

$$\mathrm{ad} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} : \quad T_e(\mathrm{Ad}_g \, \eta) \, \xi = [\xi, \eta] = \mathrm{ad}_\xi \, \eta .$$

  Explicitly, one computes the ad operation by differentiating the Ad operation directly as

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{g(t)} \, \eta &= \frac{d}{dt}\Big|_{t=0} \Big(g(t)\eta g^{-1}(t) \Big) \\ &= \dot{g}(0)\eta g^{-1}(0) - g(0)\eta g^{-1}(0)\dot{g}(0)g^{-1}(0) \\ &= \xi \, \eta - \eta \, \xi = [\xi, \eta] = \mathrm{ad}_\xi \, \eta , \end{aligned} \tag{8.1}$$

  where $g(0) = Id$, $\xi = \dot{g}(0)$ and the ***Lie bracket***

$$[\xi, \eta] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} ,$$

  is the matrix commutator for a matrix Lie algebra.

Remark

8.1 (Adjoint action). *Composition of the Adjoint action of $G \times \mathfrak{g} \to \mathfrak{g}$ of a Lie group on its Lie algebra represents the group composition law as*

$$\mathrm{Ad}_g \mathrm{Ad}_h \eta = g(h\eta h^{-1})g^{-1} = (gh)\eta(gh)^{-1} = \mathrm{Ad}_{gh}\eta ,$$

for any $\eta \in \mathfrak{g}$.

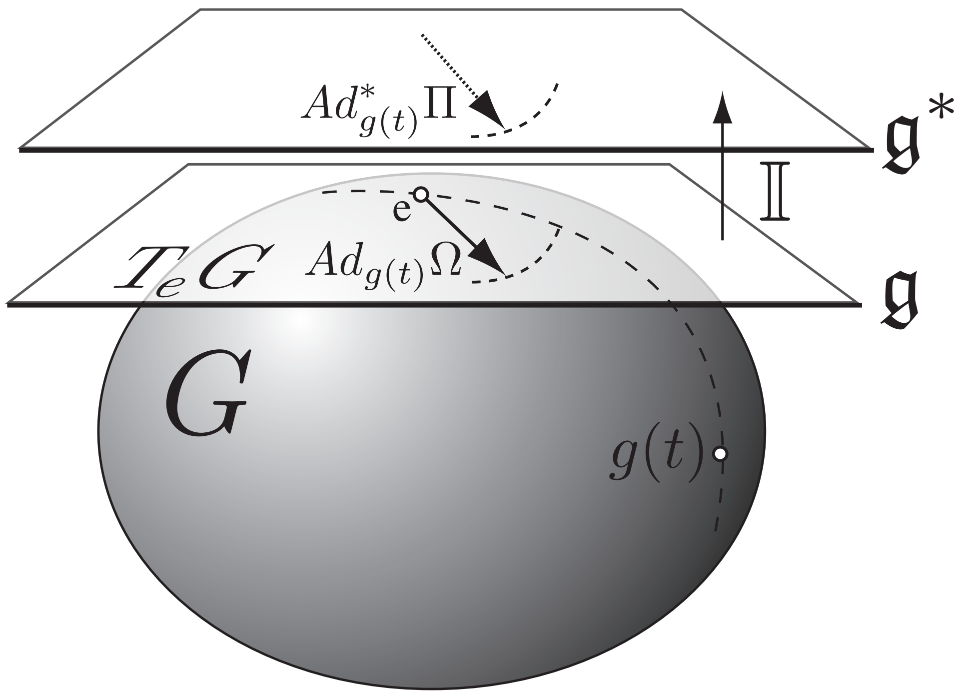

Figure 6: The Ad and Ad* operations of $g(t)$ act, respectively, on the Lie algebra $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$ and on its dual $\mathrm{Ad}^* : G \times \mathfrak{g}^* \to \mathfrak{g}^*$.

Exercise. Verify that (note the minus sign)

$$\frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{g^{-1}(t)}\, \eta = -\, \mathrm{ad}_\xi\, \eta \,,$$

for any fixed $\eta \in \mathfrak{g}$. ★

Proposition

8.2 (Adjoint motion equation)**.** *Let $g(t)$ be a path in a Lie group G and $\eta(t)$ be a path in its Lie algebra $\mathfrak{g}$. Then*

$$\frac{d}{dt} Ad_{g(t)}\eta(t) = Ad_{g(t)}\left[\frac{d\eta}{dt} + ad_{\xi(t)}\eta(t)\right],$$

where $\xi(t) = g(t)^{-1}\dot{g}(t)$.

Proof. By Equation (8.1), for a curve $\eta(t) \in \mathfrak{g}$,

$$\begin{aligned} \frac{d}{dt}\Big|_{t=t_0} \mathrm{Ad}_{g(t)}\,\eta(t) &= \frac{d}{dt}\Big|_{t=t_0} \Big(g(t)\eta(t)g^{-1}(t)\Big) \\ &= g(t_0)\Big(\dot{\eta}(t_0) + g^{-1}(t_0)\dot{g}(t_0)\eta(t_0) \\ &\quad - \eta(t_0)g^{-1}(t_0)\dot{g}(t_0)\Big)g^{-1}(t_0) \\ &= \left[\mathrm{Ad}_{g(t)}\left(\frac{d\eta}{dt} + \mathrm{ad}_\xi \eta\right)\right]_{t=t_0}. \end{aligned} \tag{8.2}$$

□

Exercise. (Inverse Adjoint motion relation) Verify that

$$\frac{d}{dt}\mathrm{Ad}_{g(t)^{-1}}\eta = -\mathrm{ad}_\xi \mathrm{Ad}_{g(t)^{-1}}\eta\,, \tag{8.3}$$

for any fixed $\eta \in \mathfrak{g}$. Note the placement of $\mathrm{Ad}_{g(t)^{-1}}$ and compare with Exercise on page 53. ★

8.1.1 Compute the coAdjoint and coadjoint operations by taking duals

The pairing

$$\big\langle \cdot\,,\,\cdot \big\rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R} \tag{8.4}$$

(which is assumed to be nondegenerate) between a Lie algebra $\mathfrak{g}$ and its dual vector space $\mathfrak{g}^*$ allows one to define the following dual operations:

- The ***coAdjoint operation*** of a Lie group on the dual of its Lie algebra is defined by the pairing with the Ad operation,

$$\mathrm{Ad}^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* : \quad \langle \mathrm{Ad}^*_g \mu \,, \eta \rangle := \langle \mu \,, \mathrm{Ad}_g \eta \rangle \,, \tag{8.5}$$

  for $g \in G$, $\mu \in \mathfrak{g}^*$ and $\xi \in \mathfrak{g}$.

- Likewise, the ***coadjoint operation*** is defined by the pairing with the ad operation,

$$\mathrm{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g}^* : \quad \langle \mathrm{ad}^*_\xi \mu \,, \eta \rangle := \langle \mu \,, \mathrm{ad}_\xi \eta \rangle \,, \tag{8.6}$$

  for $\mu \in \mathfrak{g}^*$ and $\xi, \eta \in \mathfrak{g}$.

Definition

8.3 (CoAdjoint action)**.** *The map*

$$\Phi^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* \quad \textit{given by} \quad (g, \mu) \mapsto \mathrm{Ad}^*_{g^{-1}} \mu \tag{8.7}$$

defines the ***coAdjoint action*** *of the Lie group G on its dual Lie algebra $\mathfrak{g}^*$.*

Remark

8.4 (Coadjoint group action with g^{-1})**.** *Composition of coAdjoint operations with Φ^* reverses the order in the group composition law as*

$$\mathrm{Ad}^*_g \mathrm{Ad}^*_h = \mathrm{Ad}^*_{hg} \,.$$

However, taking the inverse g^{-1} in Definition 8.3 of the coAdjoint action Φ^* restores the order and thereby allows it to represent the group composition law when acting on the dual Lie algebra, for then

$$\mathrm{Ad}^*_{g^{-1}} \mathrm{Ad}^*_{h^{-1}} = \mathrm{Ad}^*_{h^{-1}g^{-1}} = \mathrm{Ad}^*_{(gh)^{-1}} \,. \tag{8.8}$$

(See [MaRa1994] for further discussion of this point.)

The following proposition will be used later in the context of Euler–Poincaré reduction.

Proposition

8.5 (Coadjoint motion relation)**.** *Let $g(t)$ be a path in a Lie group G and $\mu(t)$ be a path in $\mathfrak{g}^*$. The corresponding* Ad^* *operation satisfies*

$$\frac{d}{dt} Ad^*_{g(t)^{-1}}\mu(t) = Ad^*_{g(t)^{-1}}\left[\frac{d\mu}{dt} - ad^*_{\xi(t)}\mu(t)\right] , \tag{8.9}$$

where $\xi(t) = g(t)^{-1}\dot{g}(t)$.

Proof. The exercise on page 54 introduces the inverse Adjoint motion relation (8.3) for any fixed $\eta \in \mathfrak{g}$, repeated as

$$\frac{d}{dt}\mathrm{Ad}_{g(t)^{-1}}\eta = -\mathrm{ad}_{\xi(t)}\left(\mathrm{Ad}_{g(t)^{-1}}\eta\right) .$$

Relation (8.3) may be proven by the following computation,

$$\begin{aligned} \frac{d}{dt}\bigg|_{t=t_0} \mathrm{Ad}_{g(t)^{-1}}\eta &= \frac{d}{dt}\bigg|_{t=t_0} \mathrm{Ad}_{g(t)^{-1}g(t_0)}\left(\mathrm{Ad}_{g(t_0)^{-1}}\eta\right) \\ &= -\mathrm{ad}_{\xi(t_0)}\left(\mathrm{Ad}_{g(t_0)^{-1}}\eta\right) , \end{aligned}$$

in which for the last step one recalls

$$\frac{d}{dt}\bigg|_{t=t_0} g(t)^{-1}g(t_0) = \left(-g(t_0)^{-1}\dot{g}(t_0)g(t_0)^{-1}\right) g(t_0) = -\xi(t_0) .$$

Relation (8.3) plays a key role in demonstrating relation (8.9) in the theorem, as follows. Using the pairing $\langle\,\cdot\,,\,\cdot\,\rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$

between the Lie algebra and its dual, one computes

$$
\begin{aligned}
\left\langle \frac{d}{dt}\mathrm{Ad}^*_{g(t)^{-1}}\mu(t), \eta \right\rangle &= \frac{d}{dt}\left\langle \mathrm{Ad}^*_{g(t)^{-1}}\mu(t), \eta \right\rangle \\
\text{by (8.5)} \quad &= \frac{d}{dt}\left\langle \mu(t), \mathrm{Ad}_{g(t)^{-1}}\eta \right\rangle \\
&= \left\langle \frac{d\mu}{dt}, \mathrm{Ad}_{g(t)^{-1}}\eta \right\rangle + \left\langle \mu(t), \frac{d}{dt}\mathrm{Ad}_{g(t)^{-1}}\eta \right\rangle \\
\text{by (8.3)} \quad &= \left\langle \frac{d\mu}{dt}, \mathrm{Ad}_{g(t)^{-1}}\eta \right\rangle + \left\langle \mu(t), -\mathrm{ad}_{\xi(t)}\left(\mathrm{Ad}_{g(t)^{-1}}\eta\right) \right\rangle \\
\text{by (8.6)} \quad &= \left\langle \frac{d\mu}{dt}, \mathrm{Ad}_{g(t)^{-1}}\eta \right\rangle - \left\langle \mathrm{ad}^*_{\xi(t)}\mu(t), \mathrm{Ad}_{g(t)^{-1}}\eta \right\rangle \\
\text{by (8.5)} \quad &= \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\frac{d\mu}{dt}, \eta \right\rangle - \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\mathrm{ad}^*_{\xi(t)}\mu(t), \eta \right\rangle \\
&= \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\left[\frac{d\mu}{dt} - \mathrm{ad}^*_{\xi(t)}\mu(t)\right], \eta \right\rangle .
\end{aligned}
$$

This concludes the proof. □

Corollary

8.6. *The* ***coadjoint orbit relation***

$$\mu(t) = Ad^*_{g(t)}\mu(0) \tag{8.10}$$

is the solution of the ***coadjoint motion equation*** *for* $\mu(t)$*,*

$$\frac{d\mu}{dt} - ad^*_{\xi(t)}\mu(t) = 0\,. \tag{8.11}$$

Proof. Substituting Equation (8.11) into Equation (8.9) yields

$$\mathrm{Ad}^*_{g(t)^{-1}}\mu(t) = \mu(0)\,.$$

Operating on this equation with $\mathrm{Ad}^*_{g(t)}$ and recalling the composition rule for Ad^* from Remark 8.4 yields the result (8.10). □

9 Example: The Heisenberg Lie group

9.1 Definitions for the Heisenberg group

The subset of the 3×3 real matrices $SL(3, \mathbb{R})$ given by the ***upper triangular matrices***

$$\left\{ H = \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \quad a, b, c \in \mathbb{R} \right\} \tag{9.1}$$

defines a noncommutative group under matrix multiplication.

The 3×3 matrix representation of this group acts on the *extended* planar vector $(x, y, 1)^T$ as

$$\begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + ay + c \\ y + b \\ 1 \end{pmatrix}.$$

The group H is called the Heisenberg group and it has three parameters. To begin studying its properties, consider the matrices in H given by

$$A = \begin{bmatrix} 1 & a_1 & a_3 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & b_1 & b_3 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9.2}$$

The matrix product gives another element of H,

$$AB = \begin{bmatrix} 1 & a_1 + b_1 & a_3 + b_3 + a_1 b_2 \\ 0 & 1 & a_2 + b_2 \\ 0 & 0 & 1 \end{bmatrix}, \tag{9.3}$$

and the inverses are

$$A^{-1} = \begin{bmatrix} 1 & -a_1 & a_1 a_2 - a_3 \\ 0 & 1 & -a_2 \\ 0 & 0 & 1 \end{bmatrix}, \qquad B^{-1} = \begin{bmatrix} 1 & -b_1 & b_1 b_2 - b_3 \\ 0 & 1 & -b_2 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9.4}$$

We are dealing with a matrix (Lie) group. The ***group commutator*** is defined by

$$[A,\, B] := ABA^{-1}B^{-1} = \begin{bmatrix} 1 & 0 & a_1 b_2 - b_1 a_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9.5}$$

Hence, the ***commutator subgroup*** $\Gamma_1(H) = [H,\, H]$ has the form

$$\Gamma_1(H) = \big\{[A,\, B] : A, B \in H\big\} \left\{ \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; \quad k \in \mathbb{R} \right\}. \tag{9.6}$$

An element C of the commutator subgroup $\Gamma_1(H)$ is of the form

$$C = \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \Gamma_1(H), \tag{9.7}$$

and we have the products

$$AC = \begin{bmatrix} 1 & a_1 & a_3 + k \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix} = CA. \tag{9.8}$$

Consequently, $[A, C] = AC(CA)^{-1} = AC(AC)^{-1} = I_3$. Hence, the subgroup of second commutators $\Gamma_2(H) = [\Gamma_1(H),\, H]$ commutes with the rest of the group, which is thus ***nilpotent of second order***.

9.2 Adjoint actions: AD, Ad and ad

Using the inverses in Equation (9.4) we compute the ***group automorphism***

$$\mathrm{AD}_B A = BAB^{-1} = \begin{bmatrix} 1 & a_1 & a_3 - a_1 b_2 + b_1 a_2 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9.9}$$

Linearising the group automorphism $\mathrm{AD}_B A$ in A at the identity yields the Ad operation,

$$\begin{aligned} \mathrm{Ad}_B \xi = B\, \xi|_{\mathrm{Id}}\, B^{-1} \quad &= \begin{bmatrix} 1 & b_1 & b_3 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -b_1 & b_1 b_2 - b_3 \\ 0 & 1 & -b_2 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & \xi_1 & \xi_3 + b_1 \xi_2 - b_2 \xi_1 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned} \tag{9.10}$$

where $\xi|_{\rm Id} = \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix}$. Equation (9.10) expresses the Ad operation of the Heisenberg group H on its Lie algebra $\mathfrak{h}(\mathbb{R}) \simeq \mathbb{R}^3$:

$$\mathrm{Ad} : H(\mathbb{R}) \times \mathfrak{h}(\mathbb{R}) \to \mathfrak{h}(\mathbb{R})\,. \tag{9.11}$$

One defines the right-invariant tangent vector,

$$\xi = \dot{A}A^{-1} = \begin{bmatrix} 0 & \dot{a}_1 & \dot{a}_3 - a_2\dot{a}_1 \\ 0 & 0 & \dot{a}_2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{h}, \tag{9.12}$$

and the left-invariant tangent vector,

$$\Xi = A^{-1}\dot{A} = \begin{bmatrix} 0 & \dot{a}_1 & \dot{a}_3 - a_1\dot{a}_2 \\ 0 & 0 & \dot{a}_2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \Xi_1 & \Xi_3 \\ 0 & 0 & \Xi_2 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{h}. \tag{9.13}$$

Next, we linearise $\mathrm{Ad}_B\xi$ in B around the identity to find the ad operation of the Heisenberg Lie algebra $\mathfrak{h}$ on itself,

$$\mathrm{ad} : \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}\,. \tag{9.14}$$

This is given explicitly by

$$\mathrm{ad}_\eta\xi = [\eta,\,\xi] := \eta\xi - \xi\eta = \begin{bmatrix} 0 & 0 & \eta_1\xi_2 - \xi_1\eta_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\,. \tag{9.15}$$

Under the equivalence $\mathfrak{h} \simeq \mathbb{R}^3$ provided by

$$\begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \mapsto \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} := \boldsymbol{\xi} \tag{9.16}$$

we may identify the Lie bracket with the projection onto the third component of the vector cross product:

$$[\eta,\,\xi] \mapsto \begin{bmatrix} 0 \\ 0 \\ \hat{\mathbf{3}} \cdot \boldsymbol{\eta} \times \boldsymbol{\xi} \end{bmatrix}\,. \tag{9.17}$$

9.3 Coadjoint actions: Ad^* and ad^*

The inner product on the Heisenberg Lie algebra $\mathfrak{h}\times\mathfrak{h}\to\mathbb{R}$ is defined by the matrix trace pairing

$$\langle \eta,\, \xi\rangle = \mathrm{Tr}(\eta^T\,\xi) = \boldsymbol{\eta}\cdot\boldsymbol{\xi}\,. \tag{9.18}$$

Thus, elements of the dual Lie algebra $\mathfrak{h}^*(\mathbb{R})$ may be represented as ***lower triangular matrices***,

$$\mu = \begin{bmatrix} 0 & 0 & 0\\ \mu_1 & 0 & 0\\ \mu_3 & \mu_2 & 0\end{bmatrix} \in \mathfrak{h}^*(\mathbb{R})\,. \tag{9.19}$$

The Ad^* operation of the Heisenberg group $H(\mathbb{R})$ on its dual Lie algebra $\mathfrak{h}^*\simeq\mathbb{R}^3$ is defined in terms of the matrix pairing by

$$\langle \mathrm{Ad}^*_B\mu,\, \xi\rangle := \langle \mu,\, \mathrm{Ad}_B\xi\rangle. \tag{9.20}$$

Explicitly, one may compute

$$\begin{aligned}
\langle \mu,\, \mathrm{Ad}_B\xi\rangle &= \mathrm{Tr}\left(\begin{bmatrix} 0 & 0 & 0\\ \mu_1 & 0 & 0\\ \mu_3 & \mu_2 & 0\end{bmatrix}\begin{bmatrix} 0 & \xi_1 & \xi_3 + b_1\xi_2 - b_2\xi_1\\ 0 & 0 & \xi_2\\ 0 & 0 & 0\end{bmatrix}\right)\\
&= \boldsymbol{\mu}\cdot\boldsymbol{\xi} + \mu_3(b_1\xi_2 - b_2\xi_1) \qquad (9.21)\\
&= \mathrm{Tr}\left(\begin{bmatrix} 0 & 0 & 0\\ \mu_1 - b_2\mu_3 & 0 & 0\\ \mu_3 & \mu_2 + b_1\mu_3 & 0\end{bmatrix}\begin{bmatrix} 0 & \xi_1 & \xi_3\\ 0 & 0 & \xi_2\\ 0 & 0 & 0\end{bmatrix}\right)\\
&= \langle \mathrm{Ad}^*_B\mu,\, \xi\rangle\,. \qquad (9.22)
\end{aligned}$$

Thus, we have the formula for $\mathrm{Ad}^*_B\mu$:

$$\mathrm{Ad}^*_B\mu = \begin{bmatrix} 0 & 0 & 0\\ \mu_1 - b_2\mu_3 & 0 & 0\\ \mu_3 & \mu_2 + b_1\mu_3 & 0\end{bmatrix}\,. \tag{9.23}$$

Likewise, the ad^* operation of the Heisenberg Lie algebra $\mathfrak{h}$ on its dual $\mathfrak{h}^*$ is defined in terms of the matrix pairing by

$$\langle \mathrm{ad}^*_\eta\mu,\, \xi\rangle := \langle \mu,\, \mathrm{ad}_\eta\xi\rangle \tag{9.24}$$

$$
\begin{aligned}
\langle \mu,\, \mathrm{ad}_\eta \xi \rangle &= \mathrm{Tr}\left(\begin{bmatrix} 0 & 0 & 0 \\ \mu_1 & 0 & 0 \\ \mu_3 & \mu_2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \eta_1\xi_2 - \xi_1\eta_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
&= \mu_3(\eta_1\xi_2 - \eta_2\xi_1) \qquad (9.25) \\
&= \mathrm{Tr}\left(\begin{bmatrix} 0 & 0 & 0 \\ -\eta_2\mu_3 & 0 & 0 \\ 0 & \eta_1\mu_3 & 0 \end{bmatrix} \begin{bmatrix} 0 & \xi_1 & \xi_3 \\ 0 & 0 & \xi_2 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
&= \langle \mathrm{ad}^*_\eta \mu,\, \xi \rangle . \qquad (9.26)
\end{aligned}
$$

Thus, we have the formula for $\mathrm{ad}^*_\eta \mu$:

$$
\mathrm{ad}^*_\eta \mu = \begin{bmatrix} 0 & 0 & 0 \\ -\eta_2\mu_3 & 0 & 0 \\ 0 & \eta_1\mu_3 & 0 \end{bmatrix} . \qquad (9.27)
$$

9.4 Coadjoint motion and harmonic oscillations

According to Proposition 8.5, the coadjoint motion relation arises by differentiating along the coadjoint orbit. Let $A(t)$ be a path in the Heisenberg Lie group H and $\mu(t)$ be a path in $\mathfrak{h}^*$. Then we compute

$$
\frac{d}{dt}\Big(\mathrm{Ad}^*_{A(t)^{-1}} \mu(t) \Big) = \mathrm{Ad}^*_{A(t)^{-1}} \left[\frac{d\mu}{dt} - \mathrm{ad}^*_{\eta(t)} \mu(t) \right] , \qquad (9.28)
$$

where $\eta(t) = A(t)^{-1}\dot{A}(t)$.

With $\eta = A^{-1}\dot{A}$, Corollary 8.11 provides the differential equation for the coadjoint orbit,

$$
\mu(t) = \mathrm{Ad}^*_{A(t)} \mu(0) .
$$

The desired differential equation is the ***coadjoint motion equation***

$$
\dot{\mu} = \mathrm{ad}^*_\eta \mu ,
$$

which may be written for the Heisenberg Lie group H as

$$
\dot{\mu} \;=\; \begin{bmatrix} 0 & 0 & 0 \\ \dot{\mu}_1 & 0 & 0 \\ \dot{\mu}_3 & \dot{\mu}_2 & 0 \end{bmatrix} = \mathrm{ad}^*_\eta \mu = \begin{bmatrix} 0 & 0 & 0 \\ -\eta_2\mu_3 & 0 & 0 \\ 0 & \eta_1\mu_3 & 0 \end{bmatrix} . \qquad (9.29)
$$

That is,

$$\frac{d}{dt}(\mu_1, \mu_2, \mu_3) = (-\eta_2\mu_3, \eta_1\mu_3, 0). \tag{9.30}$$

Thus, the coadjoint motion equation for the Heisenberg group preserves the level sets of μ_3.

If we define the linear map $\mathfrak{h} \to \mathfrak{h}^* : (\mu_1, \mu_2) = (I_1\eta_1, I_2\eta_2)$ then the coadjoint motion equations become

$$\begin{aligned} \dot{\mu}_1 &= -\mu_3\mu_2/I_2\,, \\ \dot{\mu}_2 &= \mu_3\mu_1/I_1\,, \\ \dot{\mu}_3 &= 0\,. \end{aligned} \tag{9.31}$$

Upon taking another time derivative, this set reduces to the equations

$$\ddot{\mu}_k = -\frac{\mu_3^2}{I_1 I_2}\,\mu_k\,, \quad \text{for} \quad k = 1, 2. \tag{9.32}$$

These are the equations for a planar isotropic harmonic oscillator on a level set of μ_3.

This calculation has proved the following.

Proposition

9.1. *Planar isotropic harmonic oscillations describe coadjoint orbits on the Heisenberg Lie group. The coadjoint orbits are (μ_1, μ_2) ellipses on level sets of μ_3.*

Exercise. What is the momentum map for the action of the Heisenberg Lie group on phase space? ★

10 Action principles on Lie algebras

10.1 The Euler–Poincaré theorem

Hamilton's principle for stationary action was explained earlier for deriving Euler's equations for rigid-body rotations in either their vector or quaternion forms. In the notation for the AD, Ad and ad actions of Lie groups and Lie algebras, Hamilton's principle (that the equations of motion arise from stationarity of the action) for Lagrangians defined on Lie algebras may be expressed as follows. This is the Euler–Poincaré theorem [Po1901].

Theorem

10.1 (Euler–Poincaré theorem). *Stationarity*

$$\delta S(\xi) = \delta \int_a^b l(\xi)\, dt = 0 \tag{10.1}$$

of an action

$$S(\xi) = \int_a^b l(\xi)\, dt\,,$$

whose Lagrangian is defined on the (left-invariant) Lie algebra $\mathfrak{g}$ *of a Lie group* G *by* $l(\xi) : \mathfrak{g} \mapsto \mathbb{R}$*, yields the* ***Euler–Poincaré equation*** *on* $\mathfrak{g}^*$*,*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi}\,, \tag{10.2}$$

for variations of the left-invariant Lie algebra element

$$\xi = g^{-1}\dot{g}(t) \in \mathfrak{g}$$

that are restricted to the form

$$\delta\xi = \dot{\eta} + \mathrm{ad}_\xi\, \eta\,, \tag{10.3}$$

in which $\eta(t) \in \mathfrak{g}$ *is a curve in the Lie algebra* $\mathfrak{g}$ *that vanishes at the endpoints in time.*

Exercise. What is the solution to the Euler–Poincaré Equation (10.2) in terms of $\mathrm{Ad}^*_{g(t)}$?

Hint: Take a look at the earlier equation (8.10). ★

Remark

10.2. *The earlier forms (2.43) and (2.45) of the variational formula for vectors and quaternions are now seen to apply more generally. Namely, such variations are defined for any Lie algebra.*

Proof. A direct computation proves Theorem 10.1. Later, we will explain the source of the constraint (10.3) on the form of the variations on the Lie algebra. One verifies the statement of the theorem by computing with a nondegenerate pairing $\langle \cdot\,,\,\cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$,

$$
\begin{aligned}
0 = \delta \int_a^b l(\xi)\, dt &= \int_a^b \Big\langle \frac{\delta l}{\delta \xi}\,,\, \delta \xi \Big\rangle\, dt \\
&= \int_a^b \Big\langle \frac{\delta l}{\delta \xi}\,,\, \dot\eta + \mathrm{ad}_\xi\, \eta \Big\rangle\, dt \\
&= \int_a^b \Big\langle -\,\frac{d}{dt}\frac{\delta l}{\delta \xi} + \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi}\,,\, \eta \Big\rangle dt + \Big\langle \frac{\delta l}{\delta \xi}\,,\, \eta \Big\rangle \Big|_a^b\,,
\end{aligned}
$$

upon integrating by parts. The last term vanishes, by the endpoint conditions, $\eta(b) = \eta(a) = 0$.

Since $\eta(t) \in \mathfrak{g}$ is otherwise arbitrary, (10.1) is equivalent to

$$
-\,\frac{d}{dt}\frac{\delta l}{\delta \xi} + \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} = 0\,,
$$

which recovers the Euler–Poincaré Equation (10.2) in the statement of the theorem. □

Corollary

10.3 (Noether's theorem for Euler–Poincaré)**.**
If η is an infinitesimal symmetry of the Lagrangian, then $\langle \frac{\delta l}{\delta \xi}\,,\, \eta \rangle$ is its associated constant of the Euler–Poincaré motion.

Proof. Consider the endpoint terms $\langle \frac{\delta l}{\delta \xi}\,,\, \eta \rangle |_a^b$ arising in the variation δS in (10.1) and note that this implies for any time $t \in [a, b]$ that

$$
\Big\langle \frac{\delta l}{\delta \xi(t)}\,,\, \eta(t) \Big\rangle = \text{constant},
$$

when the Euler–Poincaré Equations (10.2) are satisfied. □

Corollary

10.4 (Interpretation of Noether's theorem)**.** *Noether's theorem for the Euler–Poincaré stationary principle may be interpreted as conservation of the* spatial momentum *quantity*

$$\left(\mathrm{Ad}^*_{g^{-1}(t)}\frac{\delta l}{\delta \xi(t)}\right) = \text{constant},$$

as a consequence of the Euler–Poincaré Equation (10.2).

Proof. Invoke left-invariance of the Lagrangian $l(\xi)$ under $g \to h_\epsilon g$ with $h_\epsilon \in G$. For this symmetry transformation, one has $\delta g = \zeta g$ with $\zeta = \frac{d}{d\epsilon}\big|_{\epsilon=0} h_\epsilon$, so that

$$\eta = g^{-1}\delta g = \mathrm{Ad}_{g^{-1}}\zeta \in \mathfrak{g}\,.$$

In particular, along a curve $\eta(t)$ we have

$$\eta(t) = \mathrm{Ad}_{g^{-1}(t)}\eta(0) \quad \text{on setting} \quad \zeta = \eta(0),$$

at any initial time $t = 0$ (assuming of course that $[0, t] \in [a, b]$). Consequently,

$$\left\langle \frac{\delta l}{\delta \xi(t)}\,,\,\eta(t)\right\rangle = \left\langle \frac{\delta l}{\delta \xi(0)}\,,\,\eta(0)\right\rangle = \left\langle \frac{\delta l}{\delta \xi(t)}\,,\,\mathrm{Ad}_{g^{-1}(t)}\eta(0)\right\rangle.$$

For the nondegenerate pairing $\langle\,\cdot\,,\,\cdot\,\rangle$, this means that

$$\frac{\delta l}{\delta \xi(0)} = \left(\mathrm{Ad}^*_{g^{-1}(t)}\frac{\delta l}{\delta \xi(t)}\right) = \text{constant}.$$

The constancy of this quantity under the Euler–Poincaré dynamics in (10.2) is verified, upon taking the time derivative and using the coadjoint motion relation (8.9) in Proposition 8.5. □

Remark

10.5. *The form of the variation in (10.3) arises directly by*

(i) computing the variations of the left-invariant Lie algebra element $\xi = g^{-1}\dot{g} \in \mathfrak{g}$ *induced by taking variations* δg *in the group;*

(ii) taking the time derivative of the variation $\eta = g^{-1}g' \in \mathfrak{g}$ *; and*

(iii) using the equality of cross derivatives ($g^{\cdot\,\prime} = d^2g/dtds = g^{\prime\,\cdot}$).

Namely, one computes,

$$\begin{aligned}\xi' &= (g^{-1}\dot g)' = -\,g^{-1}g'g^{-1}\dot g + g^{-1}g^{\cdot\,\prime} = -\,\eta\xi + g^{-1}g^{\cdot\,\prime}\,,\\ \dot\eta &= (g^{-1}g')^{\cdot} = -\,g^{-1}\dot g g^{-1}g' + g^{-1}g^{\prime\,\cdot} = -\,\xi\eta + g^{-1}g^{\prime\,\cdot}\,.\end{aligned}$$

On taking the difference, the terms with cross derivatives cancel and one finds the variational formula (10.3),

$$\xi' - \dot\eta = [\,\xi\,,\,\eta\,] \quad \text{with} \quad [\,\xi\,,\,\eta\,] := \xi\,\eta - \eta\,\xi = \mathrm{ad}_\xi\,\eta\,. \tag{10.4}$$

Thus, the same formal calculations as for vectors and quaternions also apply to Hamilton's principle on (matrix) Lie algebras.

Example

10.6 (Euler–Poincaré equation for $SE(3)$)**.** *The Euler–Poincaré Equation (10.2) for* $SE(3)$ *is equivalent to*

$$\left(\frac{d}{dt}\frac{\delta l}{\delta\xi}\,,\,\frac{d}{dt}\frac{\delta l}{\delta\alpha}\right) = \mathrm{ad}^*_{(\xi\,,\,\alpha)}\left(\frac{\delta l}{\delta\xi}\,,\,\frac{\delta l}{\delta\alpha}\right)\,. \tag{10.5}$$

This formula produces the Euler–Poincaré Equation for $SE(3)$ *upon using the definition of the* ad^* *operation for* $se(3)$.

10.2 Hamilton–Pontryagin principle

Formula (10.4) for the variation of the vector $\xi = g^{-1}\dot g \in \mathfrak{g}$ may be imposed as a constraint in Hamilton's principle and thereby provide an immediate derivation of the Euler–Poincaré Equation (10.2). This constraint is incorporated into the following theorem.

Theorem

10.7 (Hamilton–Pontryagin principle [BoMa2009])**.** *The Euler–Poincaré equation*

$$\frac{d}{dt}\frac{\delta l}{\delta\xi} = \mathrm{ad}^*_\xi\frac{\delta l}{\delta\xi} \tag{10.6}$$

on the dual Lie algebra $\mathfrak{g}^*$ is equivalent to the following implicit variational principle,

$$\delta S(\xi, g, \dot{g}) = \delta \int_a^b l(\xi, g, \dot{g})\, dt = 0, \tag{10.7}$$

for a constrained action

$$\begin{aligned} S(\xi, g, \dot{g}) &= \int_a^b l(\xi, g, \dot{g})\, dt \\ &= \int_a^b \Big[l(\xi) + \langle\, \mu \,,\, (g^{-1}\dot{g} - \xi)\,\rangle \Big]\, dt\,. \end{aligned} \tag{10.8}$$

Proof. The variations of S in formula (10.8) are given by

$$\delta S = \int_a^b \Big\langle\, \frac{\delta l}{\delta \xi} - \mu \,,\, \delta\xi \,\Big\rangle + \Big\langle\, \delta\mu \,,\, (g^{-1}\dot{g} - \xi) \,\Big\rangle + \Big\langle\, \mu \,,\, \delta(g^{-1}\dot{g}) \,\Big\rangle\, dt\,.$$

Substituting $\delta(g^{-1}\dot{g})$ from (10.4) into the last term produces

$$\begin{aligned} \int_a^b \Big\langle\, \mu \,,\, \delta(g^{-1}\dot{g}) \,\Big\rangle\, dt &= \int_a^b \Big\langle\, \mu \,,\, \dot{\eta} + \mathrm{ad}_\xi\, \eta \,\Big\rangle\, dt \\ &= \int_a^b \Big\langle\, -\dot{\mu} + \mathrm{ad}^*_\xi\, \mu \,,\, \eta \,\Big\rangle\, dt + \Big\langle\, \mu \,,\, \eta \,\Big\rangle\Big|_a^b\,, \end{aligned}$$

where $\eta = g^{-1}\delta g$ vanishes at the endpoints in time. Thus, stationarity $\delta S = 0$ of the Hamilton–Pontryagin variational principle yields the following set of equations:

$$\frac{\delta l}{\delta \xi} = \mu\,, \quad g^{-1}\dot{g} = \xi\,, \quad \dot{\mu} = \mathrm{ad}^*_\xi\, \mu\,. \tag{10.9}$$

□

Remark

10.8 (Interpreting the variational formulas (10.9)).

The first formula in (10.9) is the fibre derivative needed in the Legendre transformation $\mathfrak{g} \mapsto \mathfrak{g}^*$, for passing to the Hamiltonian formulation. The second is the reconstruction formula for obtaining the solution curve $g(t) \in G$ on the Lie group G given the solution $\xi(t) = g^{-1}\dot{g} \in \mathfrak{g}$. The third formula in (10.9) is the Euler–Poincaré equation on $\mathfrak{g}^*$. The interpretation of Noether's theorem in Corollary 10.4 transfers to the Hamilton–Pontryagin variational principle as preservation of the quantity

$$\left(\mathrm{Ad}^*_{g^{-1}(t)}\mu(t)\right) = \mu(0) = constant,$$

under the Euler–Poincaré dynamics.

This Hamilton's principle is said to be ***implicit*** *because the definitions of the quantities describing the motion emerge only* after *the variations have been taken.*

Exercise. Compute the Euler–Poincaré equation on $\mathfrak{g}^*$ when $\xi(t) = \dot{g}g^{-1} \in \mathfrak{g}$ is *right-invariant.* ★

10.3 Clebsch approach to Euler–Poincaré

The Hamilton–Pontryagin (HP) Theorem 10.7 elegantly delivers the three key formulas in (10.9) needed for deriving the Lie–Poisson Hamiltonian formulation of the Euler–Poincaré equation. Perhaps surprisingly, the HP theorem accomplishes this without invoking any properties of how the invariance group of the Lagrangian G acts on the configuration space M.

An alternative derivation of these formulas exists that uses the Clebsch approach and does invoke the action $G \times M \to M$ of the Lie group on the configuration space, M, which is assumed to be a manifold. This alternative derivation is a bit more elaborate than the HP theorem. However, invoking the Lie group action on the configuration space provides additional valuable information. In particular, the alternative approach will yield information about the momentum map $T^*M \mapsto \mathfrak{g}^*$ which explains precisely how the canonical phase space T^*M maps to the Poisson manifold of the dual Lie algebra $\mathfrak{g}^*$.

Proposition

10.9 (Clebsch version of the Euler–Poincaré principle). *The Euler–Poincaré equation*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} \tag{10.10}$$

on the dual Lie algebra $\mathfrak{g}^*$ *is equivalent to the following implicit variational principle,*

$$\delta S(\xi, q, \dot{q}, p) = \delta \int_a^b l(\xi, q, \dot{q}, p)\, dt = 0, \tag{10.11}$$

for an action constrained by the reconstruction formula

$$\begin{aligned} S(\xi, q, \dot{q}, p) &= \int_a^b l(\xi, q, \dot{q}, p)\, dt \\ &= \int_a^b \Big[l(\xi) + \Big\langle\!\!\Big\langle\, p\,,\, \dot{q} + \pounds_\xi q \,\Big\rangle\!\!\Big\rangle \Big]\, dt\,, \end{aligned} \tag{10.12}$$

in which the pairing $\langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle : T^*M \times TM \mapsto \mathbb{R}$ *maps an element of the cotangent space (a momentum covector) and an element from the tangent space (a velocity vector) to a real number. This is the natural pairing for an action integrand and it also occurs in the Legendre transformation.*

Remark

10.10. *The Lagrange multiplier p in the second term of (10.12) imposes the constraint*

$$\dot{q} + \pounds_\xi q = 0\,. \tag{10.13}$$

This is the formula for the evolution of the quantity $q(t) = g^{-1}(t)q(0)$ *under the left action of the Lie algebra element* $\xi \in \mathfrak{g}$ *on it by the Lie derivative* $\pounds_\xi$ *along* ξ*. (For right action by* g *so that* $q(t) = q(0)g(t)$*, the formula is* $\dot{q} - \pounds_\xi q = 0$*.)*

10.4 Recalling the definition of the Lie derivative

One assumes the motion follows a trajectory $q(t) \in M$ in the configuration space M given by $q(t) = g(t)q(0)$, where $g(t) \in G$ is a time-dependent curve in the Lie group G which operates on the configuration space M by a flow $\phi_t : G \times M \mapsto M$. The flow property of the map $\phi_t \circ \phi_s = \phi_{s+t}$ is guaranteed by the group composition law.

Just as for the free rotations, one defines the left-invariant and right-invariant velocity vectors. Namely, as for the body angular velocity,

$$\xi_L(t) = g^{-1}\dot{g}(t) \quad \text{is left-invariant under } g(t) \to hg(t),$$

and as for the spatial angular velocity,

$$\xi_R(t) = \dot{g}g^{-1}(t) \quad \text{is right-invariant under } g(t) \to g(t)h,$$

for any choice of matrix $h \in G$. This means neither of these velocities depends on the initial configuration.

10.4.1 Right-invariant velocity vector

The Lie derivative $\pounds_\xi$ appearing in the reconstruction relation $\dot{q} = -\pounds_\xi q$ in (10.13) is defined via the Lie group operation on the configuration space exactly as for free rotation. For example, one computes the tangent vectors to the motion induced by the group operation acting from the left as $q(t) = g(t)q(0)$ by differentiating with respect to time t,

$$\dot{q}(t) = \dot{g}(t)q(0) = \dot{g}g^{-1}(t)q(t) =: \pounds_{\xi_R} q(t)\,,$$

where $\xi_R = \dot{g}g^{-1}(t)$ is right-invariant. This is the analogue of the spatial angular velocity of a freely rotating rigid body.

10.4.2 Left-invariant velocity vector

Likewise, differentiating the right action $q(t) = q(0)g(t)$ of the group on the configuration manifold yields

$$\dot{q}(t) = q(t)g^{-1}\dot{g}(t) =: \pounds_{\xi_L} q(t)\,,$$

in which the quantity

$$\xi_L(t) = g^{-1}\dot{g}(t) = \mathrm{Ad}_{g^{-1}(t)}\xi_R(t)$$

is the left-invariant tangent vector.

This analogy with free rotation dynamics should be a good guide for understanding the following manipulations, at least until we have a chance to illustrate the ideas with further examples.

Exercise. Compute the time derivatives and thus the forms of the right- and left-invariant velocity vectors for the group operations by the inverse $q(t) = q(0)g^{-1}(t)$ and $q(t) = g^{-1}(t)q(0)$. Observe the equivalence (up to a sign) of these velocity vectors with the vectors ξ_R and ξ_L, respectively. Note that the reconstruction formula (10.13) arises from the latter choice. ★

10.5 Clebsch Euler–Poincaré principle

Let us first define the concepts and notation that will arise in the course of the proof of Proposition 10.9.

Definition

10.11 (The diamond operation $\diamond$). *The diamond operation ($\diamond$) in Equation (10.17) is defined as minus the dual of the Lie derivative with respect to the pairing induced by the variational derivative in q, namely,*

$$\Big\langle\, p\diamond q\,,\,\xi\,\Big\rangle = \Big\langle\!\!\Big\langle\, p\,,\,-\,\pounds_\xi q\,\Big\rangle\!\!\Big\rangle\,. \tag{10.14}$$

Definition

10.12 (Transpose of the Lie derivative). *The* ***transpose*** *of the Lie derivative $\pounds_\xi^T p$ is defined via the pairing $\langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle$ between $(q,p)\in T^*M$ and $(q,\dot q)\in TM$ as*

$$\Big\langle\!\!\Big\langle\, \pounds_\xi^T p\,,\,q\,\Big\rangle\!\!\Big\rangle = \Big\langle\!\!\Big\langle\, p\,,\,\pounds_\xi q\,\Big\rangle\!\!\Big\rangle\,. \tag{10.15}$$

Proof. The variations of the action integral

$$S(\xi,q,\dot q,p) = \int_a^b \Big[l(\xi) + \Big\langle\!\!\Big\langle\, p\,,\,\dot q + \pounds_\xi q\,\Big\rangle\!\!\Big\rangle\Big]\,dt \tag{10.16}$$

from formula (10.12) are given by

$$\begin{aligned}\delta S &= \int_a^b \Big\langle\,\frac{\delta l}{\delta\xi}\,,\,\delta\xi\,\Big\rangle + \Big\langle\!\!\Big\langle\,\frac{\delta l}{\delta p}\,,\,\delta p\,\Big\rangle\!\!\Big\rangle + \Big\langle\!\!\Big\langle\,\frac{\delta l}{\delta q}\,,\,\delta q\,\Big\rangle\!\!\Big\rangle + \Big\langle\!\!\Big\langle\, p\,,\,\pounds_{\delta\xi} q\,\Big\rangle\!\!\Big\rangle\,dt\\ &= \int_a^b \Big\langle\,\frac{\delta l}{\delta\xi} - p\diamond q\,,\,\delta\xi\,\Big\rangle + \Big\langle\!\!\Big\langle\,\delta p\,,\,\dot q + \pounds_\xi q\,\Big\rangle\!\!\Big\rangle - \Big\langle\!\!\Big\langle\,\dot p - \pounds_\xi^T p\,,\,\delta q\,\Big\rangle\!\!\Big\rangle\,dt\,.\end{aligned}$$

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

$$\frac{\delta l}{\delta\xi} = p\diamond q\,,\quad \dot q = -\,\pounds_\xi q\,,\quad \dot p = \pounds_\xi^T p\,. \tag{10.17}$$

In these formulas, the notation distinguishes between the two types of pairings,

$$\langle\,\cdot\,,\,\cdot\,\rangle\,:\,\mathfrak{g}^*\times\mathfrak{g}\mapsto\mathbb{R}\quad\text{and}\quad \langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle\,:\,T^*M\times TM\mapsto\mathbb{R}\,. \tag{10.18}$$

(The third pairing in the formula for δS is not distinguished because it is equivalent to the second one under integration by parts in time.)

The Euler–Poincaré equation emerges from elimination of (q,p) using these formulas and the properties of the diamond operation that arise from its definition, as follows, for any vector $\eta\in\mathfrak{g}$:

$$
\begin{aligned}
\Big\langle \frac{d}{dt}\frac{\delta l}{\delta \xi}\,,\,\eta\Big\rangle &= \frac{d}{dt}\Big\langle \frac{\delta l}{\delta \xi}\,,\,\eta\Big\rangle\,,\\
[\text{Definition of }\diamond] &= \frac{d}{dt}\Big\langle p\diamond q\,,\,\eta\Big\rangle = \frac{d}{dt}\Big\langle\!\!\Big\langle p\,,\,-\pounds_\eta q\Big\rangle\!\!\Big\rangle\,,\\
[\text{Equations (10.17)}] &= \Big\langle\!\!\Big\langle \pounds_\xi^T p\,,\,-\pounds_\eta q\Big\rangle\!\!\Big\rangle + \Big\langle\!\!\Big\langle p\,,\,\pounds_\eta\pounds_\xi q\Big\rangle\!\!\Big\rangle\,,\\
[\text{Transpose, }\diamond\text{ and ad}\,] &= \Big\langle\!\!\Big\langle p\,,\,-\pounds_{[\xi,\eta]} q\Big\rangle\!\!\Big\rangle = \Big\langle p\diamond q\,,\,\mathrm{ad}_\xi\eta\Big\rangle\,,\\
[\text{Definition of ad}^*\,] &= \Big\langle\,\mathrm{ad}^*_\xi\frac{\delta l}{\delta \xi}\,,\,\eta\Big\rangle\,.
\end{aligned}
$$

This is the Euler–Poincaré Equation (10.26). □

Exercise. Show that the diamond operation defined in Equation (10.14) is antisymmetric,

$$\Big\langle p\diamond q\,,\,\xi\Big\rangle = -\Big\langle q\diamond p\,,\,\xi\Big\rangle\,. \tag{10.19}$$

★

Exercise. (Euler–Poincaré equation for right action) Compute the Euler–Poincaré equation for the Lie group action $G\times M\mapsto M: q(t)=q(0)g(t)$ in which the group acts from the right on a point $q(0)$ in the configuration manifold M along a time-dependent curve $g(t)\in G$. Explain why the result differs in sign from the case of left G-action on manifold M. ★

Exercise. (Clebsch approach for motion on $T^*(G\times V)$) Often the Lagrangian will contain a parameter taking values in a vector space V that represents a feature of the potential energy of the motion. We have encountered this situation already with the heavy top, in which the parameter is the vector in the body pointing from the contact point to the centre of mass. Since the potential energy will affect the motion we assume an action $G\times V\to V$ of the Lie group G on the vector space V. The Lagrangian then takes the form $L:TG\times V\to\mathbb{R}$.

Compute the variations of the action integral

$$S(\xi,q,\dot q,p)=\int_a^b\Big[\tilde l(\xi,q)+\Big\langle\!\Big\langle\,p\,,\,\dot q+\pounds_\xi q\,\Big\rangle\!\Big\rangle\Big]\,dt$$

and determine the effects in the Euler–Poincaré equation of having $q\in V$ appear in the Lagrangian $\tilde l(\xi,q)$.

Show first that stationarity of S implies the following set of equations:

$$\frac{\delta\tilde l}{\delta\xi}=p\diamond q\,,\quad \dot q=-\,\pounds_\xi q\,,\quad \dot p=\pounds_\xi^T p+\frac{\delta\tilde l}{\delta q}\,.$$

Then transform to the variable $\delta l/\delta\xi$ to find the associated Euler–Poincaré equations on the space $\mathfrak{g}^*\times V$,

$$\begin{aligned}\frac{d}{dt}\frac{\delta\tilde l}{\delta\xi} &= \mathrm{ad}^*_\xi\frac{\delta\tilde l}{\delta\xi}+\frac{\delta\tilde l}{\delta q}\diamond q\,,\\ \frac{dq}{dt} &= -\,\pounds_\xi q\,.\end{aligned}$$

Perform the Legendre transformation to derive the Lie–Poisson Hamiltonian formulation corresponding to $\tilde l(\xi,q)$. ★

10.6 Lie–Poisson Hamiltonian formulation

The Clebsch variational principle for the Euler–Poincaré equation provides a natural path to its canonical and Lie–Poisson Hamiltonian formulations. The Legendre transform takes the Lagrangian

$$l(p,q,\dot q,\xi)=l(\xi)+\Big\langle\!\Big\langle\,p\,,\,\dot q+\pounds_\xi q\,\Big\rangle\!\Big\rangle$$

in the action (10.16) to the Hamiltonian,

$$H(p,q) = \Big\langle\!\Big\langle\, p\,,\, \dot q \,\Big\rangle\!\Big\rangle - l(p,q,\dot q,\xi) = \Big\langle\!\Big\langle\, p\,,\, -\pounds_\xi q \,\Big\rangle\!\Big\rangle - l(\xi)\,,$$

whose variations are given by

$$\begin{aligned}\delta H(p,q) &= \Big\langle\!\Big\langle\, \delta p\,,\, -\pounds_\xi q \,\Big\rangle\!\Big\rangle + \Big\langle\!\Big\langle\, p\,,\, -\pounds_\xi \delta q \,\Big\rangle\!\Big\rangle \\ &\quad + \Big\langle\!\Big\langle\, p\,,\, -\pounds_{\delta\xi} q \,\Big\rangle\!\Big\rangle - \Big\langle\, \frac{\delta l}{\delta \xi}\,,\, \delta\xi \,\Big\rangle \\ &= \Big\langle\!\Big\langle\, \delta p\,,\, -\pounds_\xi q \,\Big\rangle\!\Big\rangle + \Big\langle\!\Big\langle\, -\pounds_\xi^T p\,,\, \delta q \,\Big\rangle\!\Big\rangle + \Big\langle\, p \diamond q - \frac{\delta l}{\delta \xi}\,,\, \delta\xi \,\Big\rangle\,.\end{aligned}$$

These variational derivatives recover Equations (10.17) in canonical Hamiltonian form,

$$\dot q = \delta H/\delta p = -\pounds_\xi q \quad \text{and} \quad \dot p = -\delta H/\delta q = \pounds_\xi^T p\,.$$

Moreover, independence of H from ξ yields the momentum relation,

$$\frac{\delta l}{\delta \xi} = p \diamond q\,. \tag{10.20}$$

The Legendre transformation of the Euler–Poincaré equations using the Clebsch canonical variables leads to the ***Lie–Poisson Hamiltonian form*** of these equations,

$$\frac{d\mu}{dt} = \{\mu, h\} = \mathrm{ad}^*_{\delta h/\delta\mu}\,\mu\,, \tag{10.21}$$

with

$$\mu = p \diamond q = \frac{\delta l}{\delta \xi}\,, \quad h(\mu) = \langle \mu, \xi\rangle - l(\xi)\,, \quad \xi = \frac{\delta h}{\delta \mu}\,. \tag{10.22}$$

By Equation (10.22), the evolution of a smooth real function $f : \mathfrak{g}^* \to \mathbb{R}$ is governed by

$$
\begin{aligned}
\frac{df}{dt} &= \left\langle \frac{\delta f}{\delta \mu}\,,\, \frac{d\mu}{dt} \right\rangle \\
&= \left\langle \frac{\delta f}{\delta \mu}\,,\, \mathrm{ad}^*_{\delta h/\delta \mu}\,\mu \right\rangle \\
&= \left\langle \mathrm{ad}_{\delta h/\delta \mu}\,\frac{\delta f}{\delta \mu}\,,\, \mu \right\rangle \\
&= -\left\langle \mu\,,\, \left[\frac{\delta f}{\delta \mu}\,,\, \frac{\delta h}{\delta \mu} \right] \right\rangle \\
&=: \big\{ f, h \big\}\,. \qquad (10.23)
\end{aligned}
$$

The last equality defines the ***Lie–Poisson bracket*** $\{f, h\}$ for smooth real functions f and h on the dual Lie algebra $\mathfrak{g}^*$. One may check directly that this bracket operation is a bilinear, skew-symmetric derivation that satisfies the Jacobi identity. Thus, it defines a proper Poisson bracket on $\mathfrak{g}^*$.

10.7 Generalised rigid body (grb)

Let the Hamiltonian H_{grb} for a generalised rigid body (grb) be defined as the pairing of the cotangent-lift momentum map J with its dual $J^\sharp = K^{-1}J \in \mathfrak{g}$,

$$
H_{grb} = \frac{1}{2}\Big\langle p \diamond q\,,\, (p \diamond q)^\sharp \Big\rangle = \frac{1}{2}\Big(p \diamond q\,,\, K^{-1}(p \diamond q) \Big)\,,
$$

for an appropriate inner product $(\,\cdot\,,\,\cdot\,) : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$.

Problem statement:

(a) Compute the canonical equations for the Hamiltonian H_{grb}.

(b) Use these equations to compute the evolution equation for $J = p \diamond q$.

(c) Identify the resulting equation and give a plausible argument why this was to be expected, by writing out its associated Hamilton's principle and Euler-Poincaré equations for left and right actions.

(d) Write the dynamical equations for q, p and J on $\mathbb{R}^3$ and explain why the name generalised rigid body might be appropriate.

Solution:

(a) By rearranging the Hamiltonian, we find

$$H_{grb} \;=\; \frac{1}{2}\Big\langle p\,,\,-\mathcal{L}_{(p\diamond q)^\sharp}\,q\Big\rangle_V = \frac{1}{2}\Big\langle -\mathcal{L}^T_{(p\diamond q)^\sharp}\,p\,,\,q\Big\rangle_V .$$

Consequently, the canonical equations for this Hamiltonian are

$$\dot q \;=\; \frac{\delta H_{grb}}{\delta p} = -\mathcal{L}_{(p\diamond q)^\sharp}\,q\,, \tag{10.24}$$

$$\dot p \;=\; -\frac{\delta H_{grb}}{\delta q} = \mathcal{L}^T_{(p\diamond q)^\sharp}\,p\,. \tag{10.25}$$

(b) These Hamiltonian equations allow us to compute the evolution equation for the cotangent-lift momentum map $J = p\diamond q$ as

$$\begin{aligned}
\Big\langle \dot J\,,\,\eta\Big\rangle &= \Big\langle \dot p\diamond q + p\diamond \dot q\,,\,\eta\Big\rangle\\
&= \Big\langle \mathcal{L}^T_\nu p\diamond q - p\diamond \mathcal{L}_\nu q\,,\,\eta\Big\rangle\,, \quad \text{where} \quad \nu = (p\diamond q)^\sharp = J^\sharp\\
&= -\Big\langle \mathcal{L}^T_\nu p\,,\,\mathcal{L}_\eta q\Big\rangle + \Big\langle p\,,\,\mathcal{L}_\eta\mathcal{L}_\nu q\Big\rangle = -\Big\langle p\,,\,\mathcal{L}_\nu\mathcal{L}_\eta q\Big\rangle + \Big\langle p\,,\,\mathcal{L}_\eta\mathcal{L}_\nu q\Big\rangle\\
&= \Big\langle p\,,\,-\mathcal{L}_{(\mathrm{ad}_\nu\eta)} q\Big\rangle = \Big\langle p\diamond q\,,\,\mathrm{ad}_\nu\eta\Big\rangle\\
&= \Big\langle \mathrm{ad}^*_\nu (p\diamond q)\,,\,\eta\Big\rangle = \Big\langle \mathrm{ad}^*_{J^\sharp} J\,,\,\eta\Big\rangle\,, \quad \text{for any} \quad \eta\in\mathfrak{g}.
\end{aligned}$$

Thus, we find that the equation of motion for a generalised rigid body is the same as the Euler-Poincaré equation for a left-invariant Lagrangian, namely,

$$\dot J = \mathrm{ad}^*_{J^\sharp} J\,. \tag{10.26}$$

(c) Equation (10.26) also results from Hamilton's principle $\delta S = 0$ given by

$$S(\xi; p, q) = \int \Big(l(\xi) + \langle p\,,\,\dot q + \mathcal{L}_\xi q\rangle \Big) dt$$

for the Clebsch-constrained reduced Lagrangian defined in terms of variables $(\xi; p, q) \in \mathfrak{g}\times T^*Q$ when we identify $\delta l/\delta\xi = J$.

(d) On $\mathbb{R}^3$ the EP equation (10.26) for grb becomes

$$\dot{J} = -\,J^\sharp \times J\,,$$

which recovers the rigid body when J is the body angular momentum and $J^\sharp = K^{-1}J$ is the body angular velocity.

The corresponding canonical Hamiltonian equations (10.24) and (10.25) for $q, p \in \mathbb{R}^3$ are

$$\dot{q} = -\,J^\sharp \times q \quad \text{and} \quad \dot{p} = -\,J^\sharp \times p\,.$$

These equations describe rigid rotations of vectors $q, p \in \mathbb{R}^3$ at angular velocity $J^\sharp$. These are in the same form as Euler's equations for rigid body motion.

In the rigid body case, $\pounds_\xi q = \xi \times q$ and, hence,

$$-p \cdot \pounds_\xi q = -p \cdot \xi \times q = p \times q \cdot \xi = p \diamond q \cdot \xi$$

so in this case $J = p \diamond q = p \times q$. That is, $(\diamond)$ in $\mathbb{R}^3$ is the cross product of vectors and $J^\sharp = K^{-1}J = K^{-1}(p \times q)$.

10.8 Cotangent-lift momentum maps

Although it is more elaborate than the Hamilton–Pontryagin principle and it requires input about the action of a Lie algebra on the configuration space, the Clebsch variational principle for the Euler–Poincaré equation reveals useful information.

As we shall see, the Clebsch approach provides a direct means of computing the ***momentum map*** for the specified Lie algebra action on a given configuration manifold M. In fact, the first equation in (10.22) is the standard example of the momentum map obtained by the ***cotangent lift*** of a Lie algebra action on a configuration manifold.

Momentum maps will be discussed later. For now, just notice that the formulas (10.22) and (**??**) involving the diamond operation have remarkable similarities. In particular, the term $q \diamond p$ in these formulas has the same meaning. Consequently, we may state the following proposition.

Proposition

10.13 (Momentum maps).

> *The Lie–Poisson form (10.21) of the Euler–Poincaré Equation (10.26) governs the evolution of the momentum map derived from the cotangent lift of the Lie algebra action on the configuration manifold.*

The remainder of the course should provide the means to fully understand this statement.

11 Formulating continuum spin chain equations

In this section we will begin thinking in terms of Hamiltonian partial differential equations in the specific example of G***-strands***, which are evolutionary maps into a Lie group $g(t,x) : \mathbb{R}\times\mathbb{R}\to G$ that follow from Hamilton's principle for a certain class of G-invariant Lagrangians. The case when $G = SO(3)$ may be regarded physically as a smooth distribution of $so(3)$-valued spins attached to a one-dimensional straight strand lying along the x-axis. We will investigate its three-dimensional orientation dynamics at each point along the strand. For no additional cost, we may begin with the Euler–Poincaré theorem for a left-invariant Lagrangian defined on the tangent space of an *arbitrary* Lie group G and later specialise to the case where G is the rotation group $SO(3)$.

The Lie–Poisson Hamiltonian formulation of the Euler–Poincaré Equation (10.26) for this problem will be derived via the Legendre Transformation by following calculations similar to those done previously for the rigid body. To emphasise the systematic nature of the Legendre transformation from the Euler–Poincaré picture to the Lie–Poisson picture, we will lay out the procedure in well-defined steps.

We shall consider Hamilton's principle $\delta S = 0$ for a left-invariant Lagrangian,

$$S = \int_a^b\int_{-\infty}^{\infty} \ell(\Omega,\Xi)\,dx\,dt\,, \tag{11.1}$$

with the following definitions of the tangent vectors Ω and Ξ,

$$\Omega(t,x) = g^{-1}\partial_t g(t,x) \quad\text{and}\quad \Xi(t,x) = g^{-1}\partial_x g(t,x)\,, \tag{11.2}$$

where $g(t,x)\in G$ is a real-valued map $g : \mathbb{R}\times\mathbb{R}\to G$ for a Lie group G. Later, we shall specialise to the case where G is the rotation group $SO(3)$. We shall apply the by now standard Euler–Poincaré procedure, modulo the partial spatial derivative in the definition of $\Xi(t,x) = g^{-1}\partial_x g(t,x)\in\mathfrak{g}$. This procedure takes the following steps:

(i) Write the auxiliary equation for the evolution of $\Xi : \mathbb{R}\times\mathbb{R}\to\mathfrak{g}$, obtained by differentiating its definition with respect to time and invoking equality of cross derivatives.

(ii) Use the Euler–Poincaré theorem for left-invariant Lagrangians to obtain the equation of motion for the momentum variable $\partial\ell/\partial\Omega : \mathbb{R}\times\mathbb{R}\to\mathfrak{g}^*$, where $\mathfrak{g}^*$ is the dual Lie algebra. Use the L^2 pairing defined by the spatial integration.

(These will be partial differential equations. Assume homogeneous boundary conditions on $\Omega(t,x)$, $\Xi(t,x)$ and vanishing endpoint conditions on the variation $\eta = g^{-1}\delta g(t,x)\in\mathfrak{g}$ when integrating by parts.)

(iii) Legendre-transform this Lagrangian to obtain the corresponding Hamiltonian. Differentiate the Hamiltonian and determine its partial derivatives. Write the Euler–Poincaré equation in terms of the new momentum variable $\Pi = \delta\ell/\delta\Omega\in\mathfrak{g}^*$.

(iv) Determine the Lie–Poisson bracket implied by the Euler–Poincaré equation in terms of the Legendre-transformed quantities $\Pi = \delta\ell/\delta\Omega$, by rearranging the time derivative of a smooth function $f(\Pi, \Xi) : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$.

(v) Specialise to $G = SO(3)$ and write the Lie–Poisson Hamiltonian form in terms of vector operations in $\mathbb{R}^3$.

(vi) For $G = SO(3)$ choose the Lagrangian

$$\begin{aligned} \ell &= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{Tr}\Big(\big[g^{-1}\partial_t g, g^{-1}\partial_x g\big]^2\Big)\,dx \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{Tr}\Big(\big[\Omega,\,\Xi\big]^2\Big)\,dx\,, \end{aligned} \tag{11.3}$$

where $[\Omega,\,\Xi] = \Omega\,\Xi - \Xi\,\Omega$ is the commutator in the Lie algebra $\mathfrak{g}$. Use the hat map to write the Euler–Poincaré equation and its Lie–Poisson Hamiltonian form in terms of vector operations in $\mathbb{R}^3$.

12 Euler–Poincaré equations

The Euler–Poincaré procedure systematically produces the following results.

Auxiliary equations By definition, $\Omega(t,x) = g^{-1}\partial_t g(t,x)$ and $\Xi(t,x) = g^{-1}\partial_x g(t,x)$ are Lie-algebra-valued functions over $\mathbb{R}\times\mathbb{R}$. The evolution of Ξ is obtained from these definitions by taking the difference of the two equations for the partial derivatives

$$\begin{aligned} \partial_t\Xi(t,x) &= -\big(g^{-1}\partial_t g\big)\big(g^{-1}\partial_x g\big) + g^{-1}\partial_t\partial_x g(t,x)\,, \\ \partial_x\Omega(t,x) &= -\big(g^{-1}\partial_x g\big)\big(g^{-1}\partial_t g\big) + g^{-1}\partial_x\partial_t g(t,x)\,, \end{aligned}$$

and invoking equality of cross derivatives. Hence, Ξ evolves by the adjoint operation, much like in the derivation of the variational derivative of Ω,

$$\partial_t\Xi(t,x) - \partial_x\Omega(t,x) = \Xi\,\Omega - \Omega\,\Xi = [\Xi,\,\Omega] =: -\,\mathrm{ad}_\Omega\Xi\,. \tag{12.1}$$

This is the auxiliary equation for $\Xi(t,x)$. In differential geometry, this relation is called a ***zero curvature relation***, because it implies that the curvature vanishes for the Lie-algebra-valued connection one-form $A = \Omega dt + \Xi dx$ [**?**].

Hamilton's principle For $\eta = g^{-1}\delta g(t,x) \in \mathfrak{g}$, Hamilton's principle $\delta S = 0$ for $S = \int_a^b \ell(\Omega, \Xi)\,dt$ leads to

$$\begin{aligned}
\delta S &= \int_a^b \Big\langle \frac{\delta \ell}{\delta \Omega}\,,\,\delta\Omega \Big\rangle + \Big\langle \frac{\delta \ell}{\delta \Xi}\,,\,\delta\Xi \Big\rangle\,dt \\
&= \int_a^b \Big\langle \frac{\delta \ell}{\delta \Omega}\,,\,\partial_t \eta + \mathrm{ad}_\Omega \eta \Big\rangle + \Big\langle \frac{\delta \ell}{\delta \Xi}\,,\,\partial_x \eta + \mathrm{ad}_\Xi \eta \Big\rangle\,dt \\
&= \int_a^b \Big\langle - \partial_t \frac{\delta \ell}{\delta \Omega} + \mathrm{ad}^*_\Omega \frac{\delta \ell}{\delta \Omega}\,,\,\eta \Big\rangle + \Big\langle - \partial_x \frac{\delta \ell}{\delta \Xi} + \mathrm{ad}^*_\Xi \frac{\delta \ell}{\delta \Xi}\,,\,\eta \Big\rangle\,dt \\
&= \int_a^b \Big\langle - \frac{\partial}{\partial t}\frac{\delta \ell}{\delta \Omega} + \mathrm{ad}^*_\Omega \frac{\delta \ell}{\delta \Omega} - \frac{\partial}{\partial x}\frac{\delta \ell}{\delta \Xi} + \mathrm{ad}^*_\Xi \frac{\delta \ell}{\delta \Xi}\,,\,\eta \Big\rangle\,dt\,,
\end{aligned}$$

where the formulas for the variations $\delta\Omega$ and $\delta\Xi$ are obtained by essentially the same calculation as in part (i). Hence, $\delta S = 0$ yields

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \Omega} = \mathrm{ad}^*_\Omega \frac{\delta \ell}{\delta \Omega} - \frac{\partial}{\partial x}\frac{\delta \ell}{\delta \Xi} + \mathrm{ad}^*_\Xi \frac{\delta \ell}{\delta \Xi}\,. \tag{12.2}$$

This is the Euler–Poincaré equation for $\delta\ell/\delta\Omega \in \mathfrak{g}^*$.

Exercise. Use Equation (8.9) in Proposition 8.5 to show that the Euler–Poincaré Equation (12.2) is a ***conservation law*** for spin angular momentum $\Pi = \delta\ell/\delta\Omega$,

$$\frac{\partial}{\partial t}\left(\mathrm{Ad}^*_{g(t,x)^{-1}}\frac{\delta l}{\delta \Omega}\right) = -\frac{\partial}{\partial x}\left(\mathrm{Ad}^*_{g(t,x)^{-1}}\frac{\delta l}{\delta \Xi}\right). \tag{12.3}$$

★

13 Hamiltonian formulation of the continuum spin chain

Legendre transform Legendre-transforming the Lagrangian $\ell(\Omega,\Xi)$: $\mathfrak{g} \times V \to \mathbb{R}$ yields the Hamiltonian $h(\Pi,\Xi) : \mathfrak{g}^* \times V \to \mathbb{R}$,

$$h(\Pi,\Xi) = \Big\langle \Pi\,,\,\Omega \Big\rangle - \ell(\Omega,\Xi)\,. \tag{13.1}$$

Differentiating the Hamiltonian determines its partial derivatives:

$$
\begin{aligned}
\delta h &= \Big\langle \delta\Pi\,,\, \frac{\delta h}{\delta \Pi} \Big\rangle + \Big\langle \frac{\delta h}{\delta \Xi}\,,\, \delta\Xi \Big\rangle \\
&= \Big\langle \delta\Pi\,,\, \Omega \Big\rangle + \Big\langle \Pi - \frac{\delta l}{\delta \Omega}\,,\, \delta\Omega \Big\rangle - \Big\langle \frac{\delta \ell}{\delta \Xi}\,,\, \delta\Xi \Big\rangle \\
\Rightarrow \quad & \frac{\delta l}{\delta \Omega} = \Pi\,, \quad \frac{\delta h}{\delta \Pi} = \Omega \quad \text{and} \quad \frac{\delta h}{\delta \Xi} = -\,\frac{\delta \ell}{\delta \Xi}\,.
\end{aligned}
$$

The middle term vanishes because $\Pi - \delta l/\delta\Omega = 0$ defines Π. These derivatives allow one to rewrite the Euler–Poincaré equation solely in terms of momentum Π as

$$
\begin{aligned}
\partial_t \Pi &= \mathrm{ad}^*_{\delta h/\delta \Pi}\,\Pi + \partial_x \frac{\delta h}{\delta \Xi} - \mathrm{ad}^*_{\Xi}\,\frac{\delta h}{\delta \Xi}\,, \\
\partial_t \Xi &= \partial_x \frac{\delta h}{\delta \Pi} - \mathrm{ad}_{\delta h/\delta \Pi}\,\Xi\,.
\end{aligned}
\tag{13.2}
$$

Hamiltonian equations The corresponding Hamiltonian equation for any functional of $f(\Pi,\Xi)$ is then

$$
\begin{aligned}
\frac{\partial}{\partial t} f(\Pi,\Xi) &= \Big\langle \partial_t\Pi\,,\, \frac{\delta f}{\delta \Pi} \Big\rangle + \Big\langle \partial_t\Xi\,,\, \frac{\delta f}{\delta \Xi} \Big\rangle \\
&= \Big\langle \mathrm{ad}^*_{\delta h/\delta \Pi}\Pi + \partial_x \frac{\delta h}{\delta \Xi} - \mathrm{ad}^*_{\Xi}\frac{\delta h}{\delta \Xi}\,,\, \frac{\delta f}{\delta \Pi} \Big\rangle \\
&\quad + \Big\langle \partial_x \frac{\delta h}{\delta \Pi} - \mathrm{ad}_{\delta h/\delta \Pi}\Xi\,,\, \frac{\delta f}{\delta \Xi} \Big\rangle \\
&= -\Big\langle \Pi\,,\, \Big[\frac{\delta f}{\delta \Pi}\,,\, \frac{\delta h}{\delta \Pi} \Big] \Big\rangle \\
&\quad + \Big\langle \partial_x \frac{\delta h}{\delta \Xi}\,,\, \frac{\delta f}{\delta \Pi} \Big\rangle - \Big\langle \partial_x \frac{\delta f}{\delta \Xi}\,,\, \frac{\delta h}{\delta \Pi} \Big\rangle \\
&\quad + \Big\langle \Xi\,,\, \mathrm{ad}^*_{\delta f/\delta \Pi}\,\frac{\delta h}{\delta \Xi} - \mathrm{ad}^*_{\delta h/\delta \Pi}\,\frac{\delta f}{\delta \Xi} \Big\rangle \\
&=: \{f\,,\, h\}(\Pi,\Xi)\,.
\end{aligned}
$$

Assembling these equations into Hamiltonian form gives, symbolically,

$$
\frac{\partial}{\partial t}\begin{bmatrix}\Pi \\ \Xi\end{bmatrix} = \begin{bmatrix} \mathrm{ad}^*_{\square}\Pi & (\mathrm{div} - \mathrm{ad}^*_{\Xi})\square \\ (\mathrm{grad} - \mathrm{ad}_{\square})\Xi & 0 \end{bmatrix} \begin{bmatrix} \delta h/\delta\Pi \\ \delta h/\delta\Xi \end{bmatrix}
\tag{13.3}
$$

The boxes □ in Equation (16.4) indicate how the ad and ad* operations are applied in the matrix multiplication. For example,

$$\mathrm{ad}^*_\square \Pi(\delta h/\delta\Pi) = \mathrm{ad}^*_{\delta h/\delta\Pi}\Pi\,,$$

so each matrix entry acts on its corresponding vector component.[3]

Higher dimensions Although it is beyond the scope of the present text, we shall make a few short comments about the meaning of the terms appearing in the Hamiltonian matrix (16.4). First, the notation indicates that the natural jump to higher dimensions has been made. This is done by using the spatial gradient to define the left-invariant auxiliary variable $\Xi \equiv g^{-1}\nabla g$ in higher dimensions. The lower left entry of the matrix (16.4) defines the ***covariant spatial gradient***, and its upper right entry defines the adjoint operator, the ***covariant spatial divergence***. More explicitly, in terms of indices and partial differential operators, this Hamiltonian matrix becomes,

$$\frac{\partial}{\partial t}\begin{bmatrix}\Pi_\alpha\\ \Xi^\alpha_i\end{bmatrix} = B_{\alpha\beta}\begin{bmatrix}\delta h/\delta\Pi_\beta\\ \delta h/\delta\Xi^\beta_j\end{bmatrix}, \tag{13.4}$$

where the Hamiltonian structure matrix $B_{\alpha\beta}$ is given explicitly as

$$B_{\alpha\beta} = \begin{bmatrix}-\Pi_\kappa\, t^{\,\kappa}_{\alpha\beta} & \delta^\beta_\alpha\partial_j + t^{\,\beta}_{\alpha\kappa}\Xi^\kappa_j\\ \delta^\alpha_\beta\partial_i - t^{\,\alpha}_{\beta\kappa}\Xi^\kappa_i & 0\end{bmatrix}. \tag{13.5}$$

Here, the summation convention is enforced on repeated indices. Superscript Greek indices refer to the Lie algebraic basis set, subscript Greek indices refer to the dual basis and Latin indices refer to the spatial reference frame. The partial derivative $\partial_j = \partial/\partial x_j$, say, acts to the right on all terms in a product by the chain rule.

Lie–Poisson bracket For the case that $t^{\,\alpha}_{\beta\kappa}$ are structure constants for the Lie algebra $so(3)$, then $t^{\,\alpha}_{\beta\kappa} = \epsilon_{\alpha\beta\kappa}$ with $\epsilon_{123} = +1$. By using the hat map (**??**), the Lie–Poisson Hamiltonian matrix in (13.5) may be rewritten for the $so(3)$ case in $\mathbb{R}^3$ vector form as

$$\frac{\partial}{\partial t}\begin{bmatrix}\boldsymbol{\Pi}\\ \boldsymbol{\Xi}_i\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Pi}\times & \partial_j + \boldsymbol{\Xi}_j\times\\ \partial_i + \boldsymbol{\Xi}_i\times & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta\boldsymbol{\Pi}\\ \delta h/\delta\boldsymbol{\Xi}_j\end{bmatrix}. \tag{13.6}$$

Returning to one dimension, stationary solutions $\partial_t \to 0$ and spatially independent solutions $\partial_x \to 0$ both satisfy equations of the same $se(3)$ form as the heavy top. For example, the time-independent solutions satisfy, with $\boldsymbol{\Omega} = \delta h/\delta\boldsymbol{\Pi}$ and $\boldsymbol{\Lambda} = \delta h/\delta\boldsymbol{\Xi}$,

$$\frac{d}{dx}\boldsymbol{\Lambda} = -\,\boldsymbol{\Xi}\times\boldsymbol{\Lambda} - \boldsymbol{\Pi}\times\boldsymbol{\Omega} \quad\text{and}\quad \frac{d}{dx}\boldsymbol{\Omega} = -\,\boldsymbol{\Xi}\times\boldsymbol{\Omega}\,.$$

[3]This is the lower right corner of the Hamiltonian matrix for a perfect complex fluid [Ho2002, GBRa2008]. It also appears in the Lie–Poisson brackets for Yang–Mills fluids [GiHoKu1982] and for spin glasses [HoKu1988].

That the equations have the same form is to be expected because of the exchange symmetry under $t \leftrightarrow x$ and $\boldsymbol{\Omega} \leftrightarrow \boldsymbol{\Xi}$. Perhaps less expected is that the heavy-top form reappears.

For $G = SO(3)$ and the Lagrangian $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ in one spatial dimension $\ell(\boldsymbol{\Omega}, \boldsymbol{\Xi})$ the Euler–Poincaré equation and its Hamiltonian form are given in terms of vector operations in $\mathbb{R}^3$, as follows. First, the Euler–Poincaré Equation (12.2) becomes

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \boldsymbol{\Omega}} = -\boldsymbol{\Omega} \times \frac{\delta \ell}{\delta \boldsymbol{\Omega}} - \frac{\partial}{\partial x}\frac{\delta \ell}{\delta \boldsymbol{\Xi}} - \boldsymbol{\Xi} \times \frac{\delta \ell}{\delta \boldsymbol{\Xi}} \,. \tag{13.7}$$

Choices for the Lagrangian

- Interesting choices for the Lagrangian include those symmetric under exchange of $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}$, such as

  $$\ell_\perp = |\boldsymbol{\Omega} \times \boldsymbol{\Xi}|^2/2 \quad \text{and} \quad \ell_\parallel = (\boldsymbol{\Omega} \cdot \boldsymbol{\Xi})^2/2 \,,$$

  for which the variational derivatives are, respectively,

  $$\frac{\delta \ell_\perp}{\delta \boldsymbol{\Omega}} = \boldsymbol{\Xi} \times (\boldsymbol{\Omega} \times \boldsymbol{\Xi}) =: |\boldsymbol{\Xi}|^2 \boldsymbol{\Omega}_\perp \,,$$
  $$\frac{\delta \ell_\perp}{\delta \boldsymbol{\Xi}} = \boldsymbol{\Omega} \times (\boldsymbol{\Xi} \times \boldsymbol{\Omega}) =: |\boldsymbol{\Omega}|^2 \boldsymbol{\Xi}_\perp \,,$$

  for $\ell_\perp$ and the complementary quantities,

  $$\frac{\delta \ell_\parallel}{\delta \boldsymbol{\Omega}} = (\boldsymbol{\Omega} \cdot \boldsymbol{\Xi})\boldsymbol{\Xi} =: |\boldsymbol{\Xi}|^2 \boldsymbol{\Omega}_\parallel \,,$$
  $$\frac{\delta \ell_\parallel}{\delta \boldsymbol{\Xi}} = (\boldsymbol{\Omega} \cdot \boldsymbol{\Xi})\boldsymbol{\Omega} =: |\boldsymbol{\Omega}|^2 \boldsymbol{\Xi}_\parallel \,,$$

  for $\ell_\parallel$. With either of these choices, $\ell_\perp$ or $\ell_\parallel$, Equation (13.7) becomes a local conservation law for spin angular momentum

  $$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \boldsymbol{\Omega}} = -\frac{\partial}{\partial x}\frac{\delta \ell}{\delta \boldsymbol{\Xi}} \,.$$

  The case $\ell_\perp$ is reminiscent of the ***Skyrme model***, a nonlinear topological model of pions in nuclear physics.

- Another interesting choice for $G = SO(3)$ and the Lagrangian $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ in one spatial dimension is

  $$\ell(\boldsymbol{\Omega}, \boldsymbol{\Xi}) = \frac{1}{2}\int_{-\infty}^{\infty} \boldsymbol{\Omega} \cdot \mathbb{A}\boldsymbol{\Omega} + \boldsymbol{\Xi} \cdot \mathbb{B}\boldsymbol{\Xi} \, dx \,,$$

for symmetric matrices $\mathbb{A}$ and $\mathbb{B}$, which may also be L^2-symmetric differential operators. In this case the variational derivatives are given by

$$\delta\ell(\boldsymbol{\Omega},\,\boldsymbol{\Xi}) = \int_{-\infty}^{\infty} \delta\boldsymbol{\Omega}\cdot\mathbb{A}\boldsymbol{\Omega} + \delta\boldsymbol{\Xi}\cdot\mathbb{B}\boldsymbol{\Xi}\,dx\,,$$

and the Euler–Poincaré Equation (12.2) becomes

$$\frac{\partial}{\partial t}\mathbb{A}\boldsymbol{\Omega} + \boldsymbol{\Omega}\times\mathbb{A}\boldsymbol{\Omega} + \frac{\partial}{\partial x}\mathbb{B}\boldsymbol{\Xi} + \boldsymbol{\Xi}\times\mathbb{B}\boldsymbol{\Xi} = 0\,. \qquad (13.8)$$

This is the sum of two coupled rotors, one in space and one in time, again suggesting the one-dimensional spin glass, or spin chain. When $\mathbb{A}$ and $\mathbb{B}$ are taken to be the identity, Equation (13.8) recovers the ***chiral model***, or ***sigma model***, which is completely integrable.

Hamiltonian structures The Hamiltonian structures of these equations on $so(3)^*$ are obtained from the Legendre-transform relations

$$\frac{\delta\ell}{\delta\boldsymbol{\Omega}} = \boldsymbol{\Pi}\,,\quad \frac{\delta h}{\delta\boldsymbol{\Pi}} = \boldsymbol{\Omega}\quad\text{and}\quad \frac{\delta h}{\delta\boldsymbol{\Xi}} = -\,\frac{\delta\ell}{\delta\boldsymbol{\Xi}}\,.$$

Hence, the Euler–Poincaré Equation (12.2) becomes

$$\frac{\partial}{\partial t}\boldsymbol{\Pi} = \boldsymbol{\Pi}\times\frac{\delta h}{\delta\boldsymbol{\Pi}} + \frac{\partial}{\partial x}\frac{\delta h}{\delta\boldsymbol{\Xi}} + \boldsymbol{\Xi}\times\frac{\delta h}{\delta\boldsymbol{\Xi}}\,, \qquad (13.9)$$

and the auxiliary Equation (13.10) becomes

$$\frac{\partial}{\partial t}\boldsymbol{\Xi} = \frac{\partial}{\partial x}\frac{\delta h}{\delta\boldsymbol{\Pi}} + \boldsymbol{\Xi}\times\frac{\delta h}{\delta\boldsymbol{\Pi}}\,, \qquad (13.10)$$

which recovers the Lie–Poisson structure in Equation (13.6).

Finally, the reconstruction equations may be expressed using the hat map as

$$\begin{aligned}\partial_t O(t,x) &= O(t,x)\widehat{\Omega}(t,x)\quad\text{and}\\ \partial_x O(t,x) &= O(t,x)\widehat{\Xi}(t,x)\,.\end{aligned} \qquad (13.11)$$

Remark

13.1. *The Euler–Poincaré equations for the continuum spin chain discussed here and their Lie–Poisson Hamiltonian formulation provide a framework for systematically investigating three-dimensional orientation dynamics along a one-dimensional strand. These partial differential equations are interesting in their own right and they have many possible applications. For an idea of where the applications of these equations could lead, consult [SiMaKr1988,EGHPR2010].*

Exercise. Write the Euler–Poincaré equations of the continuum spin chain for $SE(3)$, in which each point is both rotating and translating. Recall that

$$\left(\frac{d}{dt}\frac{\delta l}{\delta \xi}\,,\,\frac{d}{dt}\frac{\delta l}{\delta \alpha}\right) = \mathrm{ad}^*_{(\xi\,,\,\alpha)}\left(\frac{\delta l}{\delta \xi}\,,\,\frac{\delta l}{\delta \alpha}\right). \tag{13.12}$$

Apply formula (13.12) to express the space-time Euler–Poincaré Equation (12.2) for $SE(3)$ in vector form.

Complete the computation of the Lie–Poisson Hamiltonian form for the continuum spin chain on $SE(3)$. ★

Exercise. Let the set of 2×2 matrices M_i with $i=1,2,3$ satisfy the defining relation for the symplectic Lie group $Sp(2)$,

$$M_i J M_i^T = J \quad \text{with} \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{13.13}$$

The corresponding elements of its Lie algebra $m_i = \dot{M}_i M_i^{-1} \in sp(2)$ satisfy $(Jm_i)^T = Jm_i$ for each $i=1,2,3$. Thus, $\mathsf{X}_i = Jm_i$ satisfying $\mathsf{X}_i^T = \mathsf{X}_i$ is a set of three symmetric 2×2 matrices. Define $\mathsf{X} = J\dot{M}M^{-1}$ with time derivative $\dot{M} = \partial M(t,x)/\partial t$ and $\mathsf{Y} = JM'M^{-1}$ with space derivative $M' = \partial M(t,x)/\partial x$. Then show that

$$\mathsf{X}' = \dot{\mathsf{Y}} + [\mathsf{X},\mathsf{Y}]_J\,, \tag{13.14}$$

for the J-bracket defined by

$$[\mathsf{X},\mathsf{Y}]_J := \mathsf{X}J\mathsf{Y} - \mathsf{Y}J\mathsf{X} =: 2\mathrm{sym}(\mathsf{X}J\mathsf{Y}) =: \mathrm{ad}^J_{\mathsf{X}}\mathsf{Y}\,.$$

In terms of the J-bracket, compute the continuum Euler–Poincaré equations for a Lagrangian $\ell(\mathsf{X},\mathsf{Y})$ defined on the symplectic Lie algebra $\mathfrak{sp}(2)$.

Compute the Lie–Poisson Hamiltonian form of the system comprising the continuum Euler–Poincaré equations on $\mathfrak{sp}(2)^*$ and the compatibility equation (13.14) on $\mathfrak{sp}(2)$. ★

14 EPDiff and Shallow Water Waves

Figure 7: This section is about using EPDiff to model unidirectional shallow water wave trains and their interactions in one dimension.

14.1 The Euler-Poincaré equation for EPDiff

Exercise. (Worked exercise: Deriving the Euler-Poincaré equation for EPDiff in one dimension)

The EPDiff($\mathbb{R}$) equation for the H^1 norm of the velocity u is obtained from the Euler-Poincaré reduction theorem for a right-invariant Lagrangian, when one defines the Lagrangian to be half the square of the H^1 norm $\|u\|_{H^1}$ of the vector field of velocity $u = \dot{g}g^{-1} \in \mathfrak{X}(\mathbb{R})$ on the real line $\mathbb{R}$ with $g \in \mathrm{Diff}(\mathbb{R})$. Namely,

$$l(u) = \frac{1}{2}\|u\|^2_{H^1} = \frac{1}{2}\int_{-\infty}^{\infty} u^2 + u_x^2 \, dx \,.$$

(Assume $u(x)$ vanishes as $|x| \to \infty$.)

(A) Derive the EPDiff equation on the real line in terms of its velocity u and its momentum $m = \delta l/\delta u = u - u_{xx}$ in one spatial dimension for this Lagrangian.
Hint: Prove a Lemma first, that $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \mathrm{ad}_u \eta$ with $\eta = \delta g g^{-1}$.

(B) Use the Clebsch constrained Hamilton's principle

$$S(u,p,q) = \int l(u)\,dt + \sum_{a=1}^{N} \int p_a(t)\big(\dot{q}_a(t) - u(q(_a t),t)\big)dt$$

to derive the peakon singular solution $m(x,t)$ of EPDiff as a momentum map in terms of canonically conjugate variables $q_a(t)$ and $p_a(t)$, with $a = 1,2,\dots,N$.

★

Answer.

(A) ***Lemma***

The definition of velocity $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \mathrm{ad}_u\eta$ with $\eta = \delta g g^{-1}$.

Proof. Write $u = \dot{g}g^{-1}$ 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

$$\dot{g} = u \circ g \quad \text{and} \quad g' = \eta \circ g\,.$$

By the chain rule, these definitions have mixed partial derivatives

$$\dot{g}' = u' = \nabla u \cdot \eta \quad \text{and} \quad \dot{g}' = \dot{\eta} = \nabla \eta \cdot u\,.$$

The difference of the mixed partial derivatives implies the desired formula,

$$u' - \dot{\eta} = \nabla u \cdot \eta - \nabla \eta \cdot u = -\,[u,\eta] =: -\mathrm{ad}_u\eta\,,$$

so that

$$u' = \dot{\eta} - \mathrm{ad}_u \eta \,.$$

In 3D, this becomes

$$\delta \mathbf{u} = \dot{\mathbf{v}} - \mathrm{ad}_{\mathbf{u}} \mathbf{v} \,. \tag{14.1}$$

This formula may be rederived as follows. 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 (14.1), since

$$\mathbf{u}' - \dot{\mathbf{v}} = \nabla \mathbf{u} \cdot \mathbf{v} - \nabla \mathbf{v} \cdot \mathbf{u} = -\,[\mathbf{u}, \mathbf{v}] = -\,\mathrm{ad}_{\mathbf{u}} \mathbf{v} \,.$$

□

The EPDiff(H^1) equation on $\mathbb{R}$. The EPDiff(H^1) equation is written on the real line in terms of its velocity u and its momentum $m = \delta l/\delta u$ in one spatial dimension as

$$m_t + u m_x + 2 m u_x = 0 \,, \quad \text{where} \quad m = u - u_{xx} \tag{14.2}$$

where subscripts denote partial derivatives in x and t.

Proof. This equation is derived from the variational principle with $l(u) = \frac{1}{2}\|u\|^2_{H^1}$ as follows.

$$
\begin{aligned}
0 = \delta S &= \delta \int l(u)dt = \frac{1}{2}\,\delta \iint u^2 + u_x^2\,dx\,dt \\
&= \iint (u - u_{xx})\,\delta u\,dx\,dt =: \iint m\,\delta u\,dx\,dt \\
&= \iint m\,(\eta_t - \mathrm{ad}_u\eta)\,dx\,dt \\
&= \iint m\,(\eta_t + u\eta_x - \eta u_x)\,dx\,dt \\
&= -\iint (m_t + (um)_x + mu_x)\,\eta\,dx\,dt \\
&= -\iint (m_t + \mathrm{ad}^*_u m)\,\eta\,dx\,dt\,,
\end{aligned}
$$

where $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \mathrm{ad}_u\eta$ with $\eta = \delta g g^{-1}$. □

Exercise. Follow this proof in 3D and write out the resulting Euler-Poincaré equation. ★

(B) The constrained Clebsch action integral is given as

$$
S(u,p,q) = \int l(u)\,dt + \sum_{a=1}^N \int p_a(t)\big(\dot{q}_a(t) - u(q_a(t),t)\big)dt
$$

whose variation in u is gotten by inserting a delta function, so that

$$
\begin{aligned}
0 = \delta S &= \int \left(\frac{\delta l}{\delta u} - \sum_{a=1}^N p_a \delta(x - q_a(t)) \right) \delta u\,dx\,dt \\
&\quad - \int \left(\dot{p}_a(t) + \frac{\partial u}{\partial q_a}\,p_a(t) \right) \delta q_a - \delta p_a \Big(\dot{q}_a(t) - u(q_a(t),t) \Big) dt\,.
\end{aligned}
$$

The singular momentum solution $m(x,t)$ of EPDiff(H^1) is written as the cotangent-lift momentum map

$$
m(x,t) = \delta l/\delta u = \sum_{a=1}^N p_a(t)\delta(x - q_a(t)) \tag{14.3}
$$

Inserting this solution into the Legendre transform

$$h(m) = \langle\, m, u\,\rangle - l(u)$$

yields the conserved energy

$$e = \frac{1}{2}\int m(x,t)u(x,t)\,dx = \frac{1}{2}\sum_{a=1}^{N}\int p_a(t)\delta(x-q_a(t))u(x,t)\,dx = \frac{1}{2}\sum_{a=1}^{N} p_a(t)u(q_a(t),t) \tag{14.4}$$

Consequently, the variables (q_a, p_a) satisfy equations,

$$\dot q_a(t) = u(q_a(t),t)\,, \qquad \dot p_a(t) = -\frac{\partial u}{\partial q_a}\,p_a(t)\,, \tag{14.5}$$

with the pulse-train solution for velocity

$$u(q_a,t) = \sum_{b=1}^{N} p_b K(q_a,q_b) = \frac{1}{2}\sum_{b=1}^{N} p_b e^{-|q_a-q_b|} \tag{14.6}$$

where $K(x,y) = \frac{1}{2}e^{-|x-y|}$ is the Green's function kernel for the Helmholtz operator $1-\partial_x^2$. Each pulse in the pulse-train solution for velocity (14.6) has a sharp peak. For that reason, these pulses are called *peakons*. In fact, equations (14.5) are Hamilton's canonical equations with Hamiltonian obtained from equations (14.4) for energy and (14.6) for velocity, as given in [CaHo1993],

$$H_N = \frac{1}{2}\sum_{a,b=1}^{N} p_a\,p_b\,K(q_a,q_b) = \frac{1}{4}\sum_{a,b=1}^{N} p_a\,p_b\,\mathrm{e}^{-|q_a-q_b|}\,. \tag{14.7}$$

The first canonical equation in eqn (14.5) implies that the peaks at the positions $x = q^a(t)$ in the peakon-train solution (14.6) move with the flow of the fluid velocity u at those positions, since $u(q^a(t),t) = \dot q^a(t)$. This means the positions $q^a(t)$ are ***Lagrangian coordinates*** frozen into the flow of EPDiff. Thus, the singular solution obtained from the cotangent-lift momentum map (14.3) is the map from Lagrangian coordinates to Eulerian coordinates (that is, the ***Lagrange-to-Euler map***) for the momentum.

▲

Remark

14.1 (Solution behaviour of EPDiff(H^1)). *The peakon-train solutions of EPDiff are an* ***emergent phenomenon****. A wave train of peakons emerges in solving the initial-value problem for the EPDiff equation (14.2) for essentially any spatially confined initial condition. A numerical simulation of the solution behaviour for EPDiff(H^1) given in Figure 8 shows the emergence of a wave train of peakons from a Gaussian initial condition.*

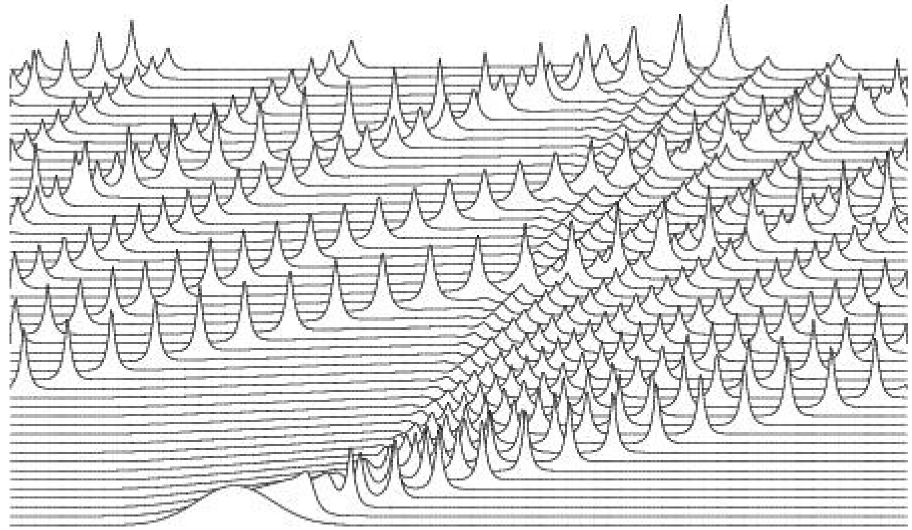

Figure 8: Under the evolution of the EPDiff(H^1) equation (14.2), an ordered ***wave train of peakons*** emerges from a smooth localized initial condition (a Gaussian). The spatial profiles at successive times are offset in the vertical to show the evolution. The peakon wave train eventually wraps around the periodic domain, thereby allowing the leading peakons to overtake the slower peakons from behind in collisions that conserve momentum and preserve the peakon shape but cause phase shifts in the positions of the peaks, as discussed in [CaHo1993].

Exercise. Compute the Lie–Poisson Hamiltonian form of the EPDiff equation (14.2). ★

Answer.

Lie-Poisson Hamiltonian form of EPDiff. In terms of m, the conserved energy Hamiltonian for the EPDiff equation (14.2) is obtained by Legendre transforming the kinetic-energy Lagrangian $l(u)$, as

$$h(m) = \Big\langle m\,,u\Big\rangle - l(u)\,.$$

Thus, the Hamiltonian depends on m, as

$$h(m) = \frac{1}{2}\int m(x)K(x-y)m(y)\,dxdy\,,$$

which also reveals the geodesic nature of the EPDiff equation (14.2) and the role of $K(x,y)$ in the kinetic energy metric on the Hamiltonian side.

The corresponding ***Lie-Poisson bracket*** for EPDiff as a Hamiltonian evolution equation is given by,

$$\partial_t m = \big\{m, h\big\} = -\mathrm{ad}^*_{\delta h/\delta m} m = -\left(\partial_x m + m\partial_x\right)\frac{\delta h}{\delta m} \quad \text{and} \quad \frac{\delta h}{\delta m} = u\,,$$

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,

$$\big\{f, h\big\} = -\int \frac{\delta f}{\delta m}(\partial_x m + m\partial_x)\frac{\delta h}{\delta m}dx = -\int m\left[\frac{\delta f}{\delta m}\,,\,\frac{\delta h}{\delta m}\right]dx$$

where $[\cdot\,,\,\cdot]$ denotes the Lie algebra bracket of vector fields. That is,

$$\left[\frac{\delta f}{\delta m}\,,\,\frac{\delta h}{\delta m}\right] = \frac{\delta f}{\delta m}\partial_x\frac{\delta h}{\delta m} - \frac{\delta h}{\delta m}\partial_x\frac{\delta f}{\delta m}\,.$$

▲

Exercise. What is the Casimir for this Lie Poisson bracket? What does it mean from the viewpoint of coadjoint orbits? What is the 3D version of this Lie Poisson bracket? Does it have a Casimir in 3D? ★

14.2 The CH equation is bi-Hamiltonian

The completely integrable CH equation for unidirectional shallow water waves first derived in [CaHo1993],

$$m_t + um_x + 2mu_x = \underbrace{-c_0u_x + \gamma u_{xxx}}_{\text{Linear Dispersion}}\,, \qquad m = u - \alpha^2 u_{xx}\,, \qquad u = K*m \quad \text{with} \quad K(x,y) = \frac{1}{2}e^{-|x-y|}\,. \tag{14.8}$$

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.

The term bi-Hamiltonian means the equation may be written in two compatible Hamiltonian forms, namely as

$$m_t = -B_2\,\frac{\delta H_1}{\delta m} = -B_1\,\frac{\delta H_2}{\delta m} \tag{14.9}$$

with

$$\begin{aligned} H_1 &= \frac{1}{2}\int (u^{\,2} + \alpha^{\,2} u_x^2)\;dx\,, \quad \text{and} \quad B_2 = \partial_x m + m\partial_x + c_0\partial_x + \gamma\,\partial_x^3 \\ H_2 &= \frac{1}{2}\int u^{\,3} + \alpha^{\,2} u\,u_x^2 + c_0 u^2 - \gamma\,u_x^2\;dx\,, \quad \text{and} \quad B_1 = \partial_x - \alpha^{\,2}\partial_x^3\,. \end{aligned} \tag{14.10}$$

These bi-Hamiltonian forms restrict properly to those for KdV when $\alpha^{\,2} \to 0$, and to those for EPDiff when $c_0, \gamma \to 0$. Compatibility of B_1 and B_2 is assured, because $(\partial_x m + m\partial_x)$, ∂_x and ∂_x^3 are all mutually compatible Hamiltonian operators. That is, any linear combination of these operators defines a Poisson bracket,

$$\{f, h\}(m) = -\int \frac{\delta f}{\delta m}(c_1 B_1 + c_2 B_2)\frac{\delta h}{\delta m}\;dx\,, \tag{14.11}$$

as a bilinear skew-symmetric operation which satisfies the Jacobi identity. Moreover, no further deformations of these Hamiltonian operators involving higher order partial derivatives would be compatible with B_2, as shown in [Ol2000]. This was already known in the literature for KdV, whose bi-Hamilton structure has $B_1 = \partial_x$ and B_2 the same as CH.

14.3 Magri's theorem

Let's start on the Hamiltonian side in 1D with the Lie–Poisson version of the Euler–Poincaré equation, namely

$$0 = m_t + (\partial m + m\partial)\frac{\delta H}{\delta m}$$

Write $m = u_x$ and let $H = \int |u_x|\,dx$ be bounded over the real line, so that $\lim_{|x|\to\infty}|u_x| = 0$. Then $\delta H/\delta u_x = \mathrm{sgn}(u_x)$ and the Lie–Poisson equation becomes

$$0 = \partial(u_t + |u_x|) + 2u_x\delta(u_x)$$

Consequently,

$$const = u_t + |u_x| = 0\,,$$

since $\partial^{-1}(2u_x\delta(u_x)) = 0$ and $\lim_{|x|\to\infty}|u_x| = 0$. This equation may have pretty explosive solutions, but at least it is evolutionary.

This approach also gives an evolutionary equation when using the vorticity bracket in $\mathbb{R}^3$ where

$$\partial_t\omega = -\text{curl}\left(\omega \times \text{curl}\frac{\delta H}{\delta\omega}\right)$$

and with $H(\omega) = \int |\omega|\, d^3x$ so that $\delta H/\delta\omega = \hat{\omega} = \omega/|\omega|$. With the vorticity bracket in 2D, this approach gives a trivial equation, because in that case, $H(\omega) = \int |\omega|\, d^2x$ is a Casimir.

As we shall see, because equation (14.8) is bi-Hamiltonian, it has an infinite number of conservation laws. These laws can be constructed by defining the transpose operator $\mathcal{R}^T = B_1^{-1}B_2$ that leads from the variational derivative of one conservation law to the next, according to

$$\frac{\delta H_n}{\delta m} = \mathcal{R}^T\frac{\delta H_{n-1}}{\delta m}\,, \quad n = -1, 0, 1, 2, \dots. \tag{14.12}$$

The operator $\mathcal{R}^T = B_1^{-1}B_2$ recursively takes the variational derivative of H_{-1} to that of H_0, to that of H_1, to then that of H_2. The next steps are not so easy for the integrable CH hierarchy, because each application of the recursion operator introduces an additional convolution integral into the sequence. Correspondingly, the recursion operator $\mathcal{R} = B_2B_1^{-1}$ leads to a hierarchy of commuting flows, defined by $K_{n+1} = \mathcal{R}K_n$, for $n = 0, 1, 2, \dots$,

$$m_t^{(n+1)} = K_{n+1}[m] = -B_1\,\frac{\delta H_n}{\delta m} = -\,B_2\frac{\delta H_{n-1}}{\delta m} = B_2B_1^{-1}K_n[m]\,. \tag{14.13}$$

The first three flows in the "positive hierarchy" when $c_0, \gamma \to 0$ are

$$m_t^{(1)} = 0\,, \quad m_t^{(2)} = -\,m_x\,, \quad m_t^{(3)} = -\,(m\partial + \partial m)u\,, \tag{14.14}$$

the third being EPDiff. The next flow is too complicated to be usefully written here. However, by construction, all of these flows commute with the other flows in the hierarchy, so they each conserve H_n for $n = 0, 1, 2, \dots$.

The recursion operator can also be continued for negative values of n. The conservation laws generated this way do not introduce convolutions, but care must be taken to ensure the conserved densities are integrable. All the Hamiltonian densities in the negative hierarchy are expressible in terms of m only and do not involve u. Thus, for instance, the first few Hamiltonians in the negative hierarchy of EPDiff are given by

$$H_0 = \int_{-\infty}^{\infty} m\,dx\,, \qquad H_{-1} = \int_{-\infty}^{\infty} \sqrt{m}\,dx\,, \tag{14.15}$$

and

$$H_{-2}=\frac{1}{2}\int_{-\infty}^{\infty}\left[\frac{\alpha^2}{4}\frac{m_x^2}{m^{5/2}}-\frac{2}{\sqrt{m}}\right]. \tag{14.16}$$

The flow defined by (14.13) for these is thus,

$$m_t^{(0)}=-B_1\,\frac{\delta H_{-1}}{\delta m}=-\,B_2\frac{\delta H_{-2}}{\delta m}=-(\partial-\alpha^2\partial^3)\left(\frac{1}{2\sqrt{m}}\right). \tag{14.17}$$

This flow is similar to the Dym equation,

$$u_{xxt}=\partial^3\left(\frac{1}{2\sqrt{u_{xx}}}\right). \tag{14.18}$$

14.4 Equation (14.8) is isospectral

The isospectral eigenvalue problem associated with equation (14.8) may be found by using the recursion relation of the bi-Hamiltonian structure, following the standard technique of [GeDo1979]. Let us introduce a spectral parameter λ and multiply by λ^n the $n-$th step of the recursion relation (14.13), then summing yields

$$B_1\sum_{n=0}^{\infty}\lambda^n\frac{\delta H_n}{\delta m}=\lambda B_2\sum_{n=0}^{\infty}\lambda^{(n-1)}\frac{\delta H_{n-1}}{\delta m}\,, \tag{14.19}$$

or, by introducing

$$\psi^2(x,t;\lambda):=\sum_{n=-1}^{\infty}\lambda^n\frac{\delta H_n}{\delta m}\,, \tag{14.20}$$

one finds, formally,

$$B_1\psi^2(x,t;\lambda)=\lambda B_2\psi^2(x,t;\lambda)\,. \tag{14.21}$$

This is a third-order eigenvalue problem for the squared-eigenfunction ψ^2, which turns out to be equivalent to a second order Sturm-Liouville problem for ψ. It is straightforward to show that if ψ satisfies

$$\lambda\left(\frac{1}{4}-\alpha^2\partial_x^2\right)\psi=\left(\frac{c_0}{4}+\frac{m(x,t)}{2}+\gamma\,\partial_x^2\right)\psi\,, \tag{14.22}$$

then ψ^2 is a solution of (14.21). Now, assuming λ will be independent of time, we seek, in analogy with the KdV equation, an evolution equation for ψ of the form,

$$\psi_t=a\psi_x+b\psi\,, \tag{14.23}$$

where a and b are functions of u and its derivatives to be determined by the requirement that the compatibility condition $\psi_{xxt} = \psi_{txx}$ between (14.22) and (14.23) implies (14.8). Cross differentiation shows

$$b = -\frac{1}{2}a_x\,, \quad \text{and} \quad a = -(\lambda + u)\,. \tag{14.24}$$

Consequently,

$$\psi_t = -(\lambda + u)\psi_x + \frac{1}{2}u_x\psi\,, \tag{14.25}$$

is the desired evolution equation for ψ.

Summary of the isospectral property of equation (14.8) Thus, according to the standard Gelfand-Dorfman theory of [GeDo1979] for obtaining the isospectral problem for equation via the squared-eigenfunction approach, its bi-Hamiltonian property implies that the nonlinear shallow water wave equation (14.8) arises as a compatibility condition for two linear equations. These are the *isospectral eigenvalue problem*,

$$\lambda\left(\frac{1}{4} - \alpha^2\partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi\,, \tag{14.26}$$

and the *evolution equation* for the eigenfunction ψ,

$$\psi_t = -(u + \lambda)\,\psi_x + \frac{1}{2}u_x\,\psi\,.$$

Compatibility of these linear equations ($\psi_{xxt} = \psi_{txx}$) together with isospectrality

$$d\lambda/dt = 0\,,$$

imply equation (14.8). Consequently, the nonlinear water wave equation (14.8) admits the IST method for the solution of its initial value problem, just as the KdV equation does. In fact, the isospectral problem for equation (14.8) restricts to the isospectral problem for KdV (i.e., the Schrödinger equation) when $\alpha^2 \to 0$.

Dispersionless case In the dispersionless case $c_0 = 0 = \gamma$, the shallow water equation (14.8) becomes the 1D geodesic equation EPDiff(H^1) in (14.2)

$$m_t + um_x + 2mu_x = 0\,, \qquad m = u - \alpha^2 u_{xx}\,, \tag{14.27}$$

and the spectrum of its eigenvalue problem (14.26) becomes *purely disctrete*. The traveling wave solutions of 1D EPDiff (14.27) in this dispersionless case are the "peakons," described by the reduced, or collective, solutions (14.5) for EPDiff equation (14.2) with traveling waves

$$u(x,t) = c\,K(x - ct) = c\,e^{-|x-ct|/\alpha}\,.$$

In this case, the EPDiff equation (14.2) may also be written as a conservation law for momentum,

$$\partial_t m = -\partial_x \Big(um + \frac{1}{2}u^2 - \frac{\alpha^2}{2}u_x^2 \Big) . \tag{14.28}$$

Its isospectral problem forms the basis for completely integrating the EPDiff equation as a Hamiltonian system and, thus, for finding its soliton solutions. Remarkably, the isospectral problem (14.26) in the dispersionless case $c_0 = 0 = \Gamma$ has purely discrete spectrum on the real line and the $N-$soliton solutions for this equation have the peakon form,

$$u(x,t) = \sum_{i=1}^{N} p_i(t) e^{-|x-q_i(t)|/\alpha} . \tag{14.29}$$

Here $p_i(t)$ and $q_i(t)$ satisfy the finite dimensional geodesic motion equations obtained as canonical Hamiltonian equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = - \frac{\partial H}{\partial q_i} , \tag{14.30}$$

when the Hamiltonian is given by,

$$H = \frac{1}{2} \sum_{i,j=1}^{N} p_i \, p_j \, e^{-|q_i - q_j|/\alpha} . \tag{14.31}$$

Thus, the CH peakons turn out to be an integrable subcase of the pulsons.

Integrability of the $N-$peakon dynamics One may verify integrability of the $N-$peakon dynamics by substituting the $N-$peakon solution (14.29) (which produces the sum of delta functions in (14.3) for the momentum map m) into the isospectral problem (14.26). This substitution reduces (14.26) to an $N \times N$ matrix eigenvalue problem.

In fact, the canonical equations (14.30) for the peakon Hamiltonian (14.31) may be written directly in Lax matrix form,

$$\frac{dL}{dt} = [L, A] \quad \Longleftrightarrow \quad L(t) = U(t) L(0) U^{\dagger}(t) , \tag{14.32}$$

with $A = \dot{U} U^{\dagger}(t)$ and $UU^{\dagger} = Id$. Explicitly, L and A are $N \times N$ matrices with entries

$$L_{jk} = \sqrt{p_j p_k} \, \phi(q_i - q_j) , \quad A_{jk} = -2\sqrt{p_j p_k} \, \phi'(q_i - q_j) . \tag{14.33}$$

Here $\phi'(x)$ denotes derivative with respect to the argument of the function ϕ, given by $\phi(x) = e^{-|x|/2\alpha}$. The Lax matrix L in (14.33) evolves by time-dependent unitary transformations, which leave its spectrum invariant. Isospectrality then implies that the traces $\operatorname{tr} L^n$, $n = 1, 2, \dots, N$ of the powers of the matrix L (or, equivalently, its N eigenvalues) yield N constants of the motion. These turn out to be independent, nontrivial and in involution. Hence, the canonically Hamiltonian $N-$peakon dynamics (14.30) is integrable.

Exercise. Show that the peakon Hamiltonian H_N in (14.31) is expressed as a function of the invariants of the matrix L, as

$$H_N = -\mathrm{tr}\, L^2 + 2(\mathrm{tr}\, L)^2 \,. \tag{14.34}$$

Show that evenness of H_N implies

1. The N coordinates q_i, $i = 1, 2, \ldots, N$ keep their initial ordering.
2. The N conjugate momenta p_i, $i = 1, 2, \ldots, N$ keep their initial signs.

This means no difficulties arise, either due to the nonanalyticity of $\phi(x)$, or the sign in the square-roots in the Lax matrices L and A. ★

14.5 Steepening Lemma and peakon formation

We now address the mechanism for the formation of the peakons, by showing that initial conditions exist for which the solution of the EPDiff(H^1) equation,

$$\partial_t m + u m_x + 2 u_x m = 0 \quad \text{with} \quad m = u - \alpha^2 u_{xx} \,, \tag{14.35}$$

can develop a vertical slope in its velocity $u(t, x)$, in finite time. The mechanism turns out to be associated with inflection points of negative slope, such as occur on the leading edge of a rightward propagating velocity profile. In particular, we have the following steepening lemma.

Lemma

14.2 (Steepening Lemma)**.**
Suppose the initial profile of velocity $u(0, x)$ has an inflection point at $x = \overline{x}$ to the right of its maximum, and otherwise it decays to zero in each direction sufficiently rapidly for the Hamiltonian H_1 in equation (14.10) to be finite. Then the negative slope at the inflection point will become vertical in finite time.

Proof. Consider the evolution of the slope at the inflection point. Define $s = u_x(\overline{x}(t), t)$. Then the EPDiff($H^1$) equation (14.35), rewritten as,

$$(1 - \alpha^2 \partial^2)(u_t + u u_x) = -\,\partial\Big(u^2 + \frac{\alpha^2}{2} u_x^2\Big) \,, \tag{14.36}$$

yields an equation for the evolution of s. Namely, using $u_{xx}(\overline{x}(t),t)=0$ leads to

$$\frac{ds}{dt}=-\frac{1}{2}s^2+\frac{1}{2}\int_{-\infty}^{\infty}\operatorname{sgn}(\overline{x}-y)e^{-|\overline{x}-y|}\partial_y\left(u^2+\frac{1}{2}u_y^2\right)dy\,. \tag{14.37}$$

Integrating by parts and using the inequality $a^2+b^2\ge 2ab$, for any two real numbers a and b, leads to

$$\begin{aligned}\frac{ds}{dt} &= -\frac{1}{2}s^2-\frac{1}{2}\int_{-\infty}^{\infty}e^{-|\overline{x}-y|}\left(u^2+\frac{1}{2}u_y^2\right)dy+u^2(\overline{x}(t),t)\\ &\le -\frac{1}{2}s^2+2u^2(\overline{x}(t),t)\,.\end{aligned} \tag{14.38}$$

Then, provided $u^2(\overline{x}(t),t)$ remains finite, say less than a number $M/4$, we have

$$\frac{ds}{dt} = -\frac{1}{2}s^2+\frac{M}{2}\,, \tag{14.39}$$

which implies, for negative slope initially $s\le-\sqrt{M}$, that

$$s \le \sqrt{M}\coth\left(\sigma+\frac{t}{2}\sqrt{M}\right)\,, \tag{14.40}$$

where σ is a negative constant that determines the initial slope, also negative. Hence, at time $t=-2\sigma/\sqrt{M}$ the slope becomes negative and vertical. The assumption that M in (14.39) exists is verified in general by a Sobolev inequality. In fact, $M=8H_1$, since

$$\max_{x\in\mathbb{R}}u^2(x,t)\le\int_{-\infty}^{\infty}\left(u^2+u_x^2\right)dx=2H_1=const\,. \tag{14.41}$$

□

Remark

14.3. If the initial condition is antisymmetric, then the inflection point at $u=0$ is fixed and $d\overline{x}/dt=0$, due to the symmetry $(u,x)\to(-u,-x)$ admitted by equation (14.8). In this case, $M=0$ and no matter how small $|s(0)|$ (with $s(0)<0$) verticality $s\to-\infty$ develops at $\overline{x}$ in finite time.

The steepening lemma indicates that traveling wave solutions of EPDiff(H^1) in (14.35) must not have the usual sech^2 shape, since inflection points with sufficiently negative slope can lead to unsteady changes in the shape of the profile if inflection points are present. In fact, numerical simulations show that the presence of an inflection point in any confined initial velocity distribution is the ***mechanism*** for the formation of the peakons. Namely. the initial (positive) velocity profile "leans" to the right and steepens, then produces a peakon which is taller than the initial profile, so it propagates away to the right. This leaves a profile behind with an inflection point of negative slope; so the process repeats, thereby producing a train of peakons with the tallest and fastest ones moving rightward in order of height. This discrete process of peakon creation corresponds to the discreteness of the isospectrum for the eigenvalue problem (14.26) in the dispersionless case, when $c_0 = 0 = \gamma$. These discrete eigenvalues correspond in turn to the asymptotic speeds of the peakons. The discreteness of the isospectrum means that only peakons will emerge in the initial value problem for EPDiff(H^1) in 1D.

15 The Euler-Poincaré framework: fluid dynamics à la [HoMaRa1998a]

15.1 Left and right momentum maps

The basic idea for the description of fluid dynamics by the action of diffeomorphisms is sketched in Fig 9.

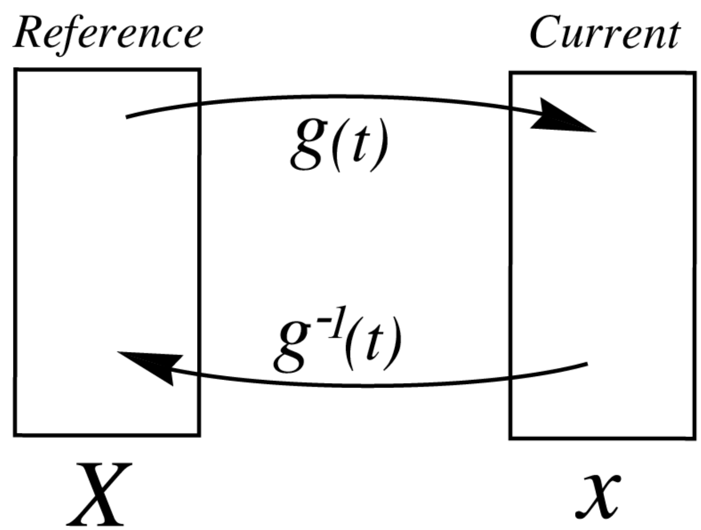

Figure 9: The forward and inverse group actions $g(t)$ and $g^{-1}(t)$ that represent ideal fluid flow are sketched here.

The forward and inverse maps sketched in Fig 9 represent ideal fluid flow by *left* group action of $g_t \in$ Diff on reference $(X \in M)$

and current ($x \in M$) coordinates. They are denoted as,

$$g_t : x(t,X) = g_t X \quad \text{and} \quad g_t^{-1} : X(t,x) = g_t^{-1} x\,, \tag{15.1}$$

so that taking time derivatives yields

$$\dot{x}(t,X) = \dot{g}_t X = (\dot{g}_t g_t^{-1}) x = \pounds_u x =: u(x,t) = u_t \circ g_t X\,, \tag{15.2}$$

and

$$\dot{X}(t,x) = (T_x g_t^{-1})(\dot{g}_t g_t^{-1} x) = T_x X \cdot u = \pounds_u X =: V(X,t) = V_t \circ g_t^{-1} x\,. \tag{15.3}$$

Here $u = \dot{g}_t g_t^{-1}$ is called the *Eulerian* velocity, and $V = \mathrm{Ad}_{g_t^{-1}} u$ is called the *convective* velocity. For $O_t \in SO(3)$, these correspond to the *spatial* angular velocity $\omega = \dot{O}_t O_t^{-1}$ and the *body* angular velocity $\Omega = \mathrm{Ad}_{O_t^{-1}} \omega = O_t^{-1} \dot{O}_t$. We shall mainly deal with the Eulerian fluid velocity in these notes.

Exercise. Use the Clebsch method to compute the momentum maps for the left group actions in (15.1). ★

15.2 The Euler-Poincaré framework for ideal fluids [HoMaRa1998a]

Almost all fluid models of interest admit the following general assumptions. These assumptions form the basis of the Euler-Poincaré theorem for ideal fluids 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 in [HoMaRa1998a], 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.[4]

[4] For fluid dynamics, right G–invariance of the Lagrangian function L is traditionally called "particle relabeling symmetry."

- 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, $\ell : \mathfrak{g} \times V^* \to \mathbb{R}$ 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 $\ell : \mathfrak{g} \times V^* \to \mathbb{R}$ a function $L : TG \times V^* \to \mathbb{R}$ that is right G–invariant, up to relabeling of a_0.

- 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 $\dot{a}(t) = -\, a(t)u(t) = \pounds_u a(t)$, where the right 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 by (summing on repeated indices)
$$[\mathbf{u}, \mathbf{v}]^i = u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j} =: -\,(\mathrm{ad}_{\mathbf{u}}\, \mathbf{v})^i\,. \tag{15.4}$$
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. This Lie algebra is given by the smooth right-invariant vector fields, $\mathfrak{g} = \mathfrak{X}$.

- Identify the Lie algebra of vector fields $\mathfrak{g}$ with its dual $\mathfrak{g}^*$ by using the L^2 pairing
$$\langle \mathbf{u}, \mathbf{v} \rangle = \int_{\mathcal{D}} \mathbf{u} \cdot \mathbf{v}\, dV\,. \tag{15.5}$$

- 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,
$$\langle \mathbf{m} \otimes dV, \mathbf{u} \rangle = \int_{\mathcal{D}} \mathbf{m} \cdot \mathbf{u}\, dV \tag{15.6}$$
where $\mathbf{m} \cdot \mathbf{u}$ is the standard contraction of a one–form $\boldsymbol{m}$ 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

$$\langle \mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV), \mathbf{v} \rangle = \int_{\mathcal{D}} \mathbf{m} \cdot \mathrm{ad}_{\mathbf{u}} \mathbf{v} \, dV = -\int_{\mathcal{D}} \mathbf{m} \cdot [\mathbf{u}, \mathbf{v}] \, dV \tag{15.7}$$

  and its expression is

$$\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV) = (\pounds_{\mathbf{u}} \mathbf{m} + (\mathrm{div}_{dV}\, \mathbf{u})\mathbf{m}) \otimes dV = \pounds_{\mathbf{u}}(\mathbf{m} \otimes dV)\,, \tag{15.8}$$

  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***

$$\Big(\mathrm{ad}^*_{\mathbf{u}}\, \mathbf{m}\Big)_i dx^i \;=\; \left(u^j \frac{\partial m_i}{\partial x^j} + m_j \frac{\partial u^j}{\partial x^i} + (\mathrm{div}_{dV}\, \mathbf{u}) m_i \right) dx^i \;=\; \left(\frac{\partial}{\partial x^j}(u^j m_i) + m_j \frac{\partial u^j}{\partial x^i} \right) dx^i\,. \tag{15.9}$$

  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.)

Definition

15.1. 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

15.2. The ***Lagrangian of a continuum mechanical system*** is a function

$$L : T\,\mathrm{Diff}(\mathcal{D}) \times V^* \to \mathbb{R}\,,$$

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 $\ell : \mathfrak{g}(\mathcal{D}) \times V^* \to \mathbb{R}$ 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,

$$\mathbf{U} = \mathbf{u} \circ \eta \quad \text{and} \quad a_0 = \eta^* a\,. \tag{15.10}$$

The evolution of a is by right action, given by the equation

$$\dot{a} = -\,\pounds_{\mathbf{u}}\, a = -\,a\mathbf{u}. \tag{15.11}$$

The solution of this equation, for the initial condition a_0, is

$$a(t) = \eta_{t*} a_0 = a_0 g^{-1}(t)\,, \tag{15.12}$$

where the lower star denotes the push forward operation and η_t is the flow of $\mathbf{u} = \dot{g} g^{-1}(t)$.

Definition

15.3. ***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 (15.11), or its solution (15.12) states that the tensor field density $a(t)$ (which may include mass density and other Eulerian quantities) is advected.

Remark

15.4 (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

15.5. 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

$$\langle b \diamond a, \mathbf{w} \rangle = -\int_{\mathcal{D}} b \cdot \pounds_{\mathbf{w}}\, a\,, \tag{15.13}$$

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

$$\dot{a} + \pounds_{\mathbf{u}} a = 0. \tag{15.14}$$

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 [HoMaRa1998a].

Theorem

15.6 (Euler–Poincaré Theorem for Continua.). *Consider 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*

$$\delta \int_{t_1}^{t_2} L\left(X, \mathbf{U}_t(X), a_0(X)\right) \, dt = 0 \tag{15.15}$$

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*

$$\delta \int_{t_1}^{t_2} \ell(\mathbf{u}, a) \, dt = 0 \tag{15.16}$$

holds on $\mathfrak{g}(\mathcal{D}) \times V^*$, using variations of the form

$$\delta \mathbf{u} = \frac{\partial \mathbf{w}}{\partial t} + [\mathbf{u}, \mathbf{w}] = \frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}} \mathbf{w} \,, \qquad \delta a = -\pounds_{\mathbf{w}} \, a, \tag{15.17}$$

where $\mathbf{w}_t = \delta\eta_t \circ \eta_t^{-1}$ vanishes at the endpoints.

iv *The Euler–Poincaré equations for continua*

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{u}} = -\, \mathrm{ad}^*_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a = -\pounds_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a \,, \tag{15.18}$$

hold, with auxiliary equations $(\partial_t + \pounds_{\mathbf{u}})a = 0$ for each advected quantity $a(t)$. The $\diamond$ operation defined in (15.13) 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 (15.8) in the last step of equation (15.18).

We refer to [HoMaRa1998a] 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**:

$$
\begin{aligned}
0 &= \delta \int_{t_1}^{t_2} l(\mathbf{u}, a) dt = \int_{t_1}^{t_2} \left(\frac{\delta l}{\delta \mathbf{u}} \cdot \delta \mathbf{u} + \frac{\delta l}{\delta a} \cdot \delta a \right) dt \\
&= \int_{t_1}^{t_2} \left[\frac{\delta l}{\delta \mathbf{u}} \cdot \left(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\, \mathbf{w} \right) - \frac{\delta l}{\delta a} \cdot \pounds_{\mathbf{w}}\, a \right] dt \\
&= \int_{t_1}^{t_2} \mathbf{w} \cdot \left[- \frac{\partial}{\partial t} \frac{\delta l}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}} \frac{\delta l}{\delta \mathbf{u}} + \frac{\delta l}{\delta a} \diamond a \right] dt \,. \qquad (15.19)
\end{aligned}
$$

The rest of the proof follows essentially the same track as the proof of the pure Euler-Poincaré theorem, modulo slight changes to accomodate the advected quantities.

In the absence of dissipation, most Eulerian fluid equations[5] can be written in the EP form in equation (15.18),

$$
\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} = \frac{\delta \ell}{\delta a} \diamond a \,, \quad \text{with} \quad \big(\partial_t + \pounds_{\mathbf{u}} \big) a = 0 \,. \qquad (15.20)
$$

Equation (15.20) 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 (15.13). Thus, Newton's Law is written in the Eulerian fluid representation as,[6]

$$
\frac{d}{dt}\Big|_{Lag} \mathbf{m} := \big(\partial_t + \pounds_{\mathbf{u}} \big) \mathbf{m} = \frac{\delta \ell}{\delta a} \diamond a \,, \quad \text{with} \quad \frac{d}{dt}\Big|_{Lag} a := \big(\partial_t + \pounds_{\mathbf{u}} \big) a = 0 \,. \qquad (15.21)
$$

- The left side of the EP equation in (15.21) 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

[5]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.

[6]In coordinates, a one-form density takes the form $\mathbf{m} \cdot d\mathbf{x} \otimes dV$ and the EP equation (15.18) is given neumonically by

$$
\frac{d}{dt}\Big|_{Lag} \big(\mathbf{m} \cdot d\mathbf{x} \otimes dV \big) = \underbrace{\frac{d\mathbf{m}}{dt}\Big|_{Lag} \cdot d\mathbf{x} \otimes dV}_{\text{Advection}} + \underbrace{\mathbf{m} \cdot d\mathbf{u} \otimes dV}_{\text{Stretching}} + \underbrace{\mathbf{m} \cdot d\mathbf{x} \otimes (\nabla \cdot \mathbf{u}) dV}_{\text{Expansion}} = \frac{\delta \ell}{\delta a} \diamond a
$$

with $\frac{d}{dt}\Big|_{Lag} d\mathbf{x} := \big(\partial_t + \pounds_{\mathbf{u}} \big) d\mathbf{x} = d\mathbf{u} = \mathbf{u}_{,j} dx^j$, 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 (15.9).

of its evolution, with respect to this norm. See [Ar1966, Ar1979, ArKh1998] 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 (15.21) 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 (15.13). For the L^2 pairing, this is integration by parts of (minus) the Lie derivative in (15.13).

- The quantity a is typically a tensor (e.g., 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 (15.21) 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 (15.21) arises in the dynamics of continuous media are given in [HoMaRa1998a]. The EP form of the Eulerian fluid description in (15.21) 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 [Po1901]. For modern discussions of the EP theory, see, e.g., [MaRa1994], or [HoMaRa1998a].

Exercise. For what types of tensors a_0 can one recast the EP equations for continua (15.18) as geodesic motion, perhaps by using a version of the Kaluza-Klein construction? ★

Exercise. State the EP theorem and write the EP equations for the *convective* velocity. ★

15.3 Corollary of the EP theorem: the Kelvin-Noether circulation theorem

Corollary

15.7 (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***

$$\frac{\partial a}{\partial t} + \pounds_{\mathbf{u}} a = 0. \tag{15.22}$$

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

$$I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}}\,. \tag{15.23}$$

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 (15.22) 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***,

$$\frac{d}{dt} I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta a}\diamond a\ . \tag{15.24}$$

Both an abstract proof of the Kelvin-Noether Circulation Theorem and a proof tailored for the case of continuum mechanical systems are given in [HoMaRa1998a]. 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 (15.18), 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

15.8. Since the last expression holds for every loop γ_t, we may write it as

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right)\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} = \frac{1}{D}\frac{\delta l}{\delta a}\diamond a\,. \tag{15.25}$$

Remark

15.9. 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 (15.24) 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.

16 The Hamiltonian formulation of ideal fluid dynamics

Legendre transform Taking the Legendre-transform of the Lagrangian $l(u,a)$: $\mathfrak{g}\times V \to \mathbb{R}$ yields the Hamiltonian $h(m,a) : \mathfrak{g}^*\times V \to \mathbb{R}$, given by

$$h(m,a) = \Big\langle m\,,\,u\Big\rangle - l(u,a)\,. \tag{16.1}$$

Differentiating the Hamiltonian determines its partial derivatives:

$$\begin{aligned}
\delta h &= \Big\langle \delta m\,,\,\frac{\delta h}{\delta m}\Big\rangle + \Big\langle \frac{\delta h}{\delta a}\,,\,\delta a\Big\rangle \\
&= \Big\langle \delta m\,,\,u\Big\rangle + \Big\langle m - \frac{\delta l}{\delta u}\,,\,\delta u\Big\rangle - \Big\langle \frac{\delta \ell}{\delta a}\,,\,\delta a\Big\rangle \\
&\Rightarrow\ \frac{\delta l}{\delta u} = m\,, \quad \frac{\delta h}{\delta m} = u \quad \text{and} \quad \frac{\delta h}{\delta a} = -\,\frac{\delta \ell}{\delta a}\,.
\end{aligned}$$

The middle term vanishes because $m - \delta l/\delta u = 0$ defines m. These derivatives allow one to rewrite the Euler–Poincaré equation for continua in (15.18) solely in terms of momentum m and advected quantities a as

$$\begin{aligned}
\partial_t m &= -\,\mathrm{ad}^*_{\delta h/\delta m}\, m - \frac{\delta h}{\delta a}\diamond a\,, \\
\partial_t a &= -\,\pounds_{\delta h/\delta m}\, a\,.
\end{aligned} \tag{16.2}$$

Hamiltonian equations The corresponding Hamiltonian equation for any functional of $f(m,a)$ is then

$$\begin{aligned}
\frac{d}{dt}f(m,a) &= \Big\langle \partial_t m\,,\,\frac{\delta f}{\delta m}\Big\rangle + \Big\langle \partial_t a\,,\,\frac{\delta f}{\delta a}\Big\rangle \\
&= -\Big\langle \mathrm{ad}^*_{\delta h/\delta m}\, m + \frac{\delta h}{\delta a}\diamond a\,,\,\frac{\delta f}{\delta m}\Big\rangle - \Big\langle \pounds_{\delta h/\delta m}\, a\,,\,\frac{\delta f}{\delta a}\Big\rangle \\
&= -\Big\langle m\,,\,\Big[\frac{\delta f}{\delta m},\,\frac{\delta h}{\delta m}\Big]\Big\rangle + \Big\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\Big\rangle \\
&=: \{f\,,\,h\}(m,a)\,,
\end{aligned} \tag{16.3}$$

which is plainly antisymmetric under the exchange $f \leftrightarrow h$. Assembling these equations into Hamiltonian form gives, symbolically,

$$\frac{\partial}{\partial t}\begin{bmatrix} m \\ a \end{bmatrix} = -\begin{bmatrix} \mathrm{ad}^*_\Box m & \Box\diamond a \\ \pounds_\Box a & 0 \end{bmatrix}\begin{bmatrix} \delta h/\delta m \\ \delta h/\delta a \end{bmatrix} \tag{16.4}$$

The boxes $\square$ in Equation (16.4) indicate how the various operations are applied in the matrix multiplication. For example,

$$\mathrm{ad}^*_{\square} m(\delta h/\delta m) = \mathrm{ad}^*_{\delta h/\delta m} m\,,$$

so each matrix entry acts on its corresponding vector component.

Remark

16.1. *The expression*

$$\{f\,,\,h\}(m,a) = -\left\langle m\,,\,\left[\frac{\delta f}{\delta m},\,\frac{\delta h}{\delta m}\right]\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\right\rangle$$

in (16.3) *defines the Lie-Poisson bracket on the dual to the semidirect-product Lie algebra* $\mathfrak{X}\circledS V^*$ *with Lie bracket*

$$\mathrm{ad}_{(u,\alpha)}\,(\overline{u},\,\overline{\alpha}) = \big(\mathrm{ad}_u\overline{u},\,\pounds^T_u\overline{\alpha} - \pounds^T_{\overline{u}}\alpha\big)$$

The coordinates are velocity vector field $u\in\mathfrak{X}$ *dual to momentum density* $m\in\mathfrak{X}^*$ *and* $\alpha\in V^*$ *dual to the vector space of advected quantities* $a\in V$*.*

Proof. We check that

$$\begin{aligned}\frac{df}{dt}(m,a) = \{f\,,\,h\}(m,a) &= \left\langle m\,,\,\mathrm{ad}_{\frac{\delta f}{\delta m}}\frac{\delta h}{\delta m}\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\right\rangle\\ &= -\left\langle \mathrm{ad}^*_{\frac{\delta h}{\delta m}} m\,,\,\frac{\delta f}{\delta m}\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a}\right\rangle + \left\langle -\pounds_{\delta h/\delta m}a\,,\,\frac{\delta f}{\delta a}\right\rangle\\ &= -\left\langle \mathrm{ad}^*_{\frac{\delta h}{\delta m}} m + \frac{\delta h}{\delta a}\diamond a\,,\,\frac{\delta f}{\delta m}\right\rangle - \left\langle \pounds_{\delta h/\delta m}a\,,\,\frac{\delta f}{\delta a}\right\rangle\end{aligned}$$

Note that the angle brackets refer to different types of pairings. This should cause no confusion. $\square$

17 Example: Euler–Poincaré theorem for GFD (geophysical fluid dynamics)

Figure 10 shows a screen shot of numerical simulations of damped and driven geophysical fluid dynamics (GFD) equations of the type studied in this section, taken from `http://www.youtube.com/watch?v=ujBi9Ba8hqs&feature=youtu.be`. The variations in space and time of the driving and damping by the Sun are responsible for the characteristic patterns of the flow. The nonlinear GFD equations in the absence of damping and driving are formulated in this section by using the Euler–Poincaré theorem.

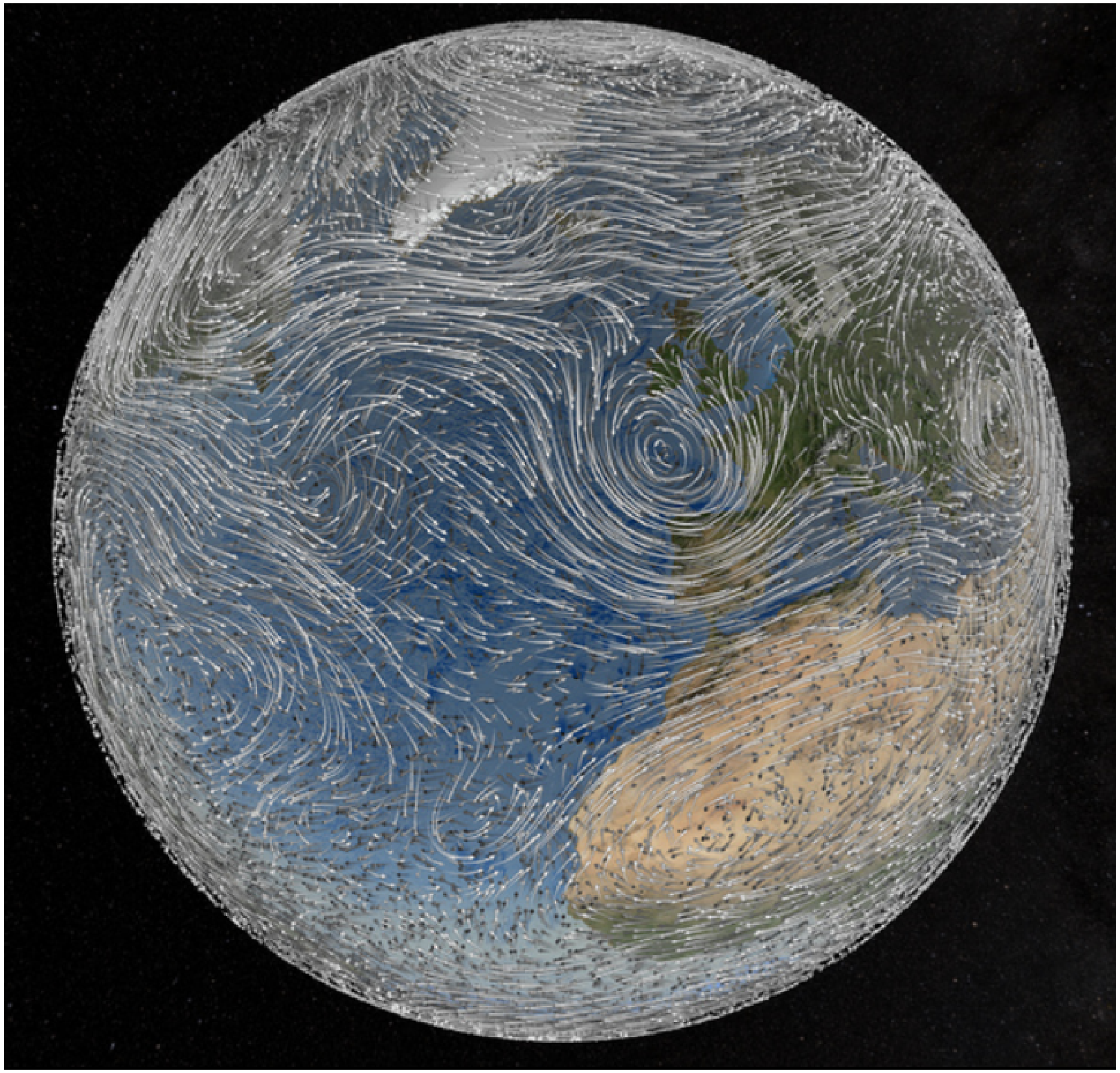

Figure 10: Atmospheric flows on Earth (wind currents) are driven by the Sun and its interaction with the surface and they are damped primarily by friction with the surface.

17.1 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

$$a \in \{b, D\, d^3x\}\,. \tag{17.1}$$

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,

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) b = 0, \quad \text{or} \quad \frac{\partial b}{\partial t} = -\ \mathbf{u}\cdot\nabla\, b\,, \tag{17.2}$$

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) D\, d^3x = 0, \quad \text{or} \quad \frac{\partial D}{\partial t} = -\ \nabla\cdot(D\mathbf{u})\,. \tag{17.3}$$

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 15.6, equation (15.16), 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

$$\begin{aligned} \delta b &= -\ \pounds_{\mathbf{w}}\, b = -\ \mathbf{w}\cdot\nabla\, b\,, \\ \delta D\ d^3x &= -\ \pounds_{\mathbf{w}}\,(D\,d^3x) = -\ \nabla\cdot(D\mathbf{w})\ d^3x\,. \end{aligned} \tag{17.4}$$

Hence, Hamilton's principle (15.16) with this dependence yields

$$\begin{aligned} 0 &= \delta\int dt\ \ell(\mathbf{u}; b, D) \\ &= \int dt\ \left[\frac{\delta\ell}{\delta\mathbf{u}}\cdot\delta\mathbf{u} + \frac{\delta\ell}{\delta b}\ \delta b + \frac{\delta\ell}{\delta D}\ \delta D\right] \\ &= \int dt\ \left[\frac{\delta\ell}{\delta\mathbf{u}}\cdot\left(\frac{\partial\mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\,\mathbf{w}\right) - \frac{\delta\ell}{\delta b}\ \mathbf{w}\cdot\nabla\, b - \frac{\delta\ell}{\delta D}\ \Big(\nabla\cdot(D\mathbf{w})\Big)\right] \\ &= \int dt\ \mathbf{w}\cdot\left[-\ \frac{\partial}{\partial t}\frac{\delta\ell}{\delta\mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}}\ \frac{\delta\ell}{\delta\mathbf{u}} - \frac{\delta\ell}{\delta b}\ \nabla\, b + D\ \nabla\frac{\delta\ell}{\delta D}\right] \\ &= -\int dt\ \mathbf{w}\cdot\left[\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right)\frac{\delta\ell}{\delta\mathbf{u}} + \frac{\delta\ell}{\delta b}\ \nabla\, b - D\ \nabla\frac{\delta\ell}{\delta D}\right], \end{aligned} \tag{17.5}$$

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.

17.2 Euler–Poincaré framework for GFD

The Euler–Poincaré equations for continua (15.18) may now be summarized in vector form for advected Eulerian variables a in the set (17.1). We adopt the notational convention of the circulation map I in equations (15.23) and (15.24) 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 $(\partial/\partial t+\pounds_{\mathbf{u}})Dd^3x=0$. Then, the Euclidean components of the Euler–Poincaré equations for continua in equation (17.5) are expressed in Kelvin theorem form (15.25) with a slight abuse of notation as

$$\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)\Big(\frac{1}{D}\frac{\delta\ell}{\delta\mathbf{u}}\cdot d\mathbf{x}\Big)\,+\,\frac{1}{D}\frac{\delta\ell}{\delta b}\nabla b\cdot d\mathbf{x}\,-\,\nabla\Big(\frac{\delta\ell}{\delta D}\Big)\cdot d\mathbf{x}=0\,, \tag{17.6}$$

in which the variational derivatives of the Lagrangian ℓ are to be computed according to the usual physical conventions, i.e., as Fréchet derivatives. Formula (17.6) is the Kelvin–Noether form of the equation of motion for ideal continua. Hence, we have the explicit Kelvin theorem expression, cf. equations (15.23) and (15.24),

$$\frac{d}{dt}\oint_{\gamma_t(\mathbf{u})}\frac{1}{D}\frac{\delta\ell}{\delta\mathbf{u}}\cdot d\mathbf{x}=-\oint_{\gamma_t(\mathbf{u})}\frac{1}{D}\frac{\delta\ell}{\delta b}\nabla b\cdot d\mathbf{x}\,, \tag{17.7}$$

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

$$\frac{\partial q}{\partial t}+\mathbf{u}\cdot\nabla q=0\,,\quad\text{where}\quad q=\frac{1}{D}\nabla b\cdot\operatorname{curl}\left(\frac{1}{D}\frac{\delta\ell}{\delta\mathbf{u}}\right). \tag{17.8}$$

This is also called ***PV convection***. Equations (17.6-17.8) embody most of the panoply of equations for GFD. The vector form of equation (17.6) is,

$$\underbrace{\Big(\frac{\partial}{\partial t}+\mathbf{u}\cdot\nabla\Big)\Big(\frac{1}{D}\frac{\delta l}{\delta\mathbf{u}}\Big)+\frac{1}{D}\frac{\delta l}{\delta u^j}\nabla u^j}_{\textbf{Geodesic Nonlinearity: Kinetic energy}}\quad=\quad\underbrace{\nabla\frac{\delta l}{\delta D}-\frac{1}{D}\frac{\delta l}{\delta b}\nabla b}_{\textbf{Potential energy}} \tag{17.9}$$

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

$$I = \oint_{\gamma(t)} \frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} \cdot d\mathbf{x}\,.$$

17.3 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

$$l(\mathbf{u}, b, D) = \int \rho_0 D(1+b)\left(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - gz\right) - p(D-1)\,d^3x\,, \tag{17.10}$$

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,

$$\begin{aligned} \frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} &= \rho_0(1+b)(\mathbf{u}+\mathbf{R})\,, \quad \frac{\delta l}{\delta b} = \rho_0 D\left(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\right),\\ \frac{\delta l}{\delta D} &= \rho_0(1+b)\left(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\right) - p\,, \quad \frac{\delta l}{\delta p} = -\,(D-1)\,. \end{aligned} \tag{17.11}$$

Hence, from the Euclidean component formula (17.9) for Hamilton principles of this type and the fundamental vector identity,

$$(\mathbf{b}\cdot\nabla)\mathbf{a} + a_j\nabla b^j = -\,\mathbf{b}\times(\nabla\times\mathbf{a}) + \nabla(\mathbf{b}\cdot\mathbf{a})\,, \tag{17.12}$$

we find the motion equation for an Euler fluid in three dimensions,

$$\frac{d\mathbf{u}}{dt} - \mathbf{u}\times\operatorname{curl}\mathbf{R} + g\hat{\mathbf{z}} + \frac{1}{\rho_0(1+b)}\nabla p = 0\,, \tag{17.13}$$

where $\operatorname{curl}\mathbf{R} = 2\boldsymbol{u}(\mathbf{x})$ is the Coriolis parameter (i.e., 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 (17.2). The pressure p is determined by requiring preservation of the constraint $D = 1$, for which the continuity equation (17.3) implies $\operatorname{div}\mathbf{u} = 0$. The Euler motion equation (17.13) 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

$$\frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q = 0\,, \quad \text{where, on using } D = 1\,, \quad q = \nabla b\cdot \operatorname{curl}\big(\mathbf{u}+\mathbf{R}\big). \tag{17.14}$$

Semidirect-product Lie-Poisson bracket for compressible ideal fluids.

1. Compute the Legendre transform for the Lagrangian,

$$l(\mathbf{u}, b, D) : \mathfrak{X}\times\Lambda^0\times\Lambda^3 \mapsto \mathbb{R}$$

   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

$$l(\mathbf{u}, b, D) = \int D\Big(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - e(D,b)\Big)\, d^3x\,, \tag{17.15}$$

   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?
5. Write the equations of motion and confirm their Kelvin-Noether circulation theorem.
6. Use the Kelvin-Noether circulation theorem for this theory to determine its potential vorticity and obtain the corresponding conservation laws. Write these conservation laws explicitly.

★

References

[AbMa1978] Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, 2nd ed. Reading, MA: Addison-Wesley.

[Ar1966] Arnold V.I. [1966], Sur la géometrie differentialle des groupes de Lie de dimiension infinie et ses applications à l'hydrodynamique des fluids parfaits. *Ann. Inst. Fourier, Grenoble* **16**, 319-361.

[Ar1979] Arnold V.I. [1979], *Mathematical Methods of Classical Mechanics*. Volume 60 of *Graduate Texts in Mathematics*. Springer.

[ArKh1998] Arnold, V. I and Khesin, B. A. [1998] *Topological Methods in Hydrodynamics.* Springer: New York.

[Bl1965] Bloembergen, N. [1965] *Nonlinear Optics*. New York: Benjamin.

[BoMa2009] Bou-Rabee, N. and Marsden, J. E. [2009]

[BoWo1965] Born, M. and Wolf, E. [1965] *Principles of Optics*. Oxford: Pergamon Press, p. 31.

[CaHo1993] Camassa, R. and Holm, D. D. [1993] An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.* **71**, 1661-64.

[FeVeHe1957] Feynman, R., Vernon, F. and Hellwarth, R. [1957] Geometrical representation of the Schrödinger equation for solving maser problems. *J. Appl. Phys.* **28**, 49–52.

[GBRa2008] Gay-Balmaz, F. and Ratiu, T.S. [2008] The geometric structure of complex fluids. *Adv. in Appl. Math.* **42** (2), 176–275.

[GeDo1979] Gelfand, I. M. and Dorfman, I. Ya. R. [1979] *Funct. Anal. Appl.* **13**, 248.

[GiHoKu1982] Gibbons, J., Holm, D. D. and Kupershmidt, B. A. [1982] Gauge-invariant Poisson brackets for chromohydrodynamics. *Phys. Lett. A* **90**, 281–283.

[Ho2002] Holm, D. D. [2002] Euler–Poincaré dynamics of perfect complex fluids. In *Geometry, Mechanics, and Dynamics: In Honor of the 60th Birthday of Jerrold E. Marsden*, edited by P. Newton, P. Holmes and A. Weinstein. New York: Springer, pp. 113–167.

[HoKu1988] Holm, D. D. and Kupershmidt, B. A. [1988] The analogy between spin glasses and Yang–Mills fluids. *J. Math Phys.* **29**, 21–30.

[Ho2005] Holm, D. D. [2005] The Euler–Poincaré variational framework for modeling fluid dynamics. In *Geometric Mechanics and Symmetry: The Peyresq Lectures*, edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306. Cambridge: Cambridge University Press.

[Ho2011GM1] Holm, D. D. [2011]
Geometric Mechanics I: Dynamics and Symmetry,
Second edition, World Scientific: Imperial College Press, Singapore, .

[Ho2011GM2] Holm, D. D. [2011]
Geometric Mechanics II: Rotating, Translating & Rolling,
Second edition, World Scientific: Imperial College Press, Singapore, .

[Ho2011] Holm, D. D. [2011]
Applications of Poisson geometry to physical problems,
Geometry & Topology Monographs **17**, 221–384.

[HoMaRa1998a] Holm, D. D., Marsden, J. E. and Ratiu, T. S. [1998a] The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. in Math.* **137**, 1–81.

[HoSmSt2009] Holm, D. D., Schmah, T. and Stoica, C. [2009]
Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions,
Oxford University Press.

[MaRa1994] Marsden, J. E. and Ratiu, T. S. [1994]
Introduction to Mechanics and Symmetry.
Texts in Applied Mathematics, Vol. 75. New York: Springer-Verlag.

[Ol2000] P. J. Olver [2000] *Applications of Lie Groups to Differential Equations*. Springer: New York.

[Po1901] H. Poincaré, Sur une forme nouvelle des équations de la méchanique,
C.R. Acad. Sci. **132** (1901) 369-371.
English translation in [Ho2011GM2], Appendix D.

[RaTuSbSoTe2005] Ratiu, T. S., Tudoran, R., Sbano, L., Sousa Dias, E. and Terra, G. [2005]
A crash course in geometric mechanics.
In *Geometric Mechanics and Symmetry: The Peyresq Lectures*,
edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306.
Cambridge: Cambridge University Press.

[Sh1984] Shen, Y. R. [1984] *The Principles of Nonlinear Optics*. New York: Wiley-lnterscience.

[St1852] Stokes, G. G. [1852] On the composition and resolution of streams of polarized light from different sources. *Trans. Camb. Phil. Soc.* **9**, 399–423.