GMC Network

GMC Notes, Number 2 (2012)

Darryl D Holm

Euler-Poincaré Theory from the Rigid Body to Solitons

6^{th} GMC Summer School Lectures, Miraflores de La Sierra, 22–26 June 2012

available at http://gmcnetwork.org

About GMC Notes:

GMC Notes are based in some material given in one of the courses at activities organized by the GMC network. These notes will be shared in a free only-digital edition. For further details about GMC Notes, please contact Ignazio Lacirasella (University of Bari, Italy) at lacirasella@dm.uniba.it.

GMC Network is an international thematic project about Geometry, Mechanics and Control. The organization of this network is based in on-going collaborations that have already been materialized in publications and joint activities between the different research groups.
GMC Network is supported by MEC (Spain) grant MTM2010-12116-E. For more informations, please visit the web page http://gmcnetwork.org/.

EULER-POINCARÉ THEORY FROM THE RIGID BODY TO SOLITONS
VI GMC SUMMER SCHOOL LECTURES, MIRAFLORES, 22–26 JUNE 2012

DARRYL D HOLM   IMPERIAL COLLEGE LONDON   D.HOLM@IC.AC.UK

http://www.ma.ic.ac.uk/~dholm/

The present notes are delivered for the attendants of the *VI Young Researchers Workshop on Geometry, Mechanics and Control*, held in Miraflores de la Sierra, Madrid (Spain), in June 2012.

Darryl D Holm is a Professor of Applied Mathematics at Imperial College London and a Laboratory Fellow in the Computation and Computer Science Division at Los Alamos National Laboratory.

Texts for the course include:
Geometric Mechanics I & II, by Darryl D Holm,
World Scientific: Imperial College Press, Singapore, Second edition (2011).
ISBN 978-1-84816-195-5

Main topic: The Euler-Poincaré theorem for geodesic motion on Lie groups

(1) **The Euler-Poincaré theorem is reviewed**
(2) **Rigid body dynamics is written as geodesic motion on $SO(3)$ and $SO(n)$**
   (a) Euler's RB equations are derived from the Euler-Poincaré theorem
   (b) The Hamiltonian formulation summons Lie-Poisson brackets and Nambu brackets on $\mathbb{R}^3$.
   (c) The Hamilton-Pontryagin variational principle imposes the reconstruction equation.
   (d) The Clebsch approach (constrained by infinitesimal transformation) leads to an interpretation of RB motion in terms of cotangent-lift momentum maps that are left-equivariant:

$$\begin{tikzcd} T^*G \arrow[r, "\Phi_{g(t)}"] \arrow[d, "J(0)"'] \arrow[dr, phantom, "\text{Left-equivariant Momentum Map}"] & T^*G \arrow[d, "J(t)"] \\ \mathfrak{g}^* \arrow[r, "\mathrm{Ad}^*_{g(t)^{-1}}"] & \mathfrak{g}^* \simeq T^*G/G \end{tikzcd}$$

   (e) One may interpret the EP equation as the infinitesimal equivariance relation

$$\xi = g^{-1}\dot{g}(t) \text{ or } \dot{g}(t)g^{-1} \qquad J := \frac{\delta \ell}{\delta \xi} \qquad \mathrm{Ad}^*_{g(t)^{-1}} J(0) = J(t) \qquad \frac{dJ}{dt} = \pm \mathrm{ad}^*_\xi J$$

   (f) RB dynamics is cast as an isospectral problem, written in Manakov's commutator form.
(3) **EPDiff as geodesic motion on** $\mathrm{Diff}(\mathbb{R}^n)$
   (a) The EPDiff equation is derived from the Euler-Poincaré theorem.
   (b) The Clebsch variational principle (constrained by the infinitesimal transformation) yields cotangent lift (CL) momentum maps.
   (c) CL momentum maps are infinitesimally equivariant, since $J := \frac{\delta \ell}{\delta \xi}$ obeys $\dot{J} = \pm \mathrm{ad}^*_\xi J$.
   (d) The bi-Hamiltonian formulation of EPDiff in 1D sets up and Magri's Lemmas.
   (e) The isospectral problem for EPDiff is found in 1D, and the N-soliton (peakon) dynamics is written in commutator (LAX) form.

Contents

Of course, not all of these letures could have been given at the five-day summer school at Miraflores, To help select topics, the students were asked at the end of each lecture to turn in a piece of paper including their **name, date and email address** on which they were supposed to answer the following two questions *in sentences*.

(1) What was this lecture about?
(2) Write a question that you would like to see addressed in a subsequent lecture.

Subsequent lectures at the summer school then selected material from these notes that emphasized the questions and interests that had been expressed by the students.

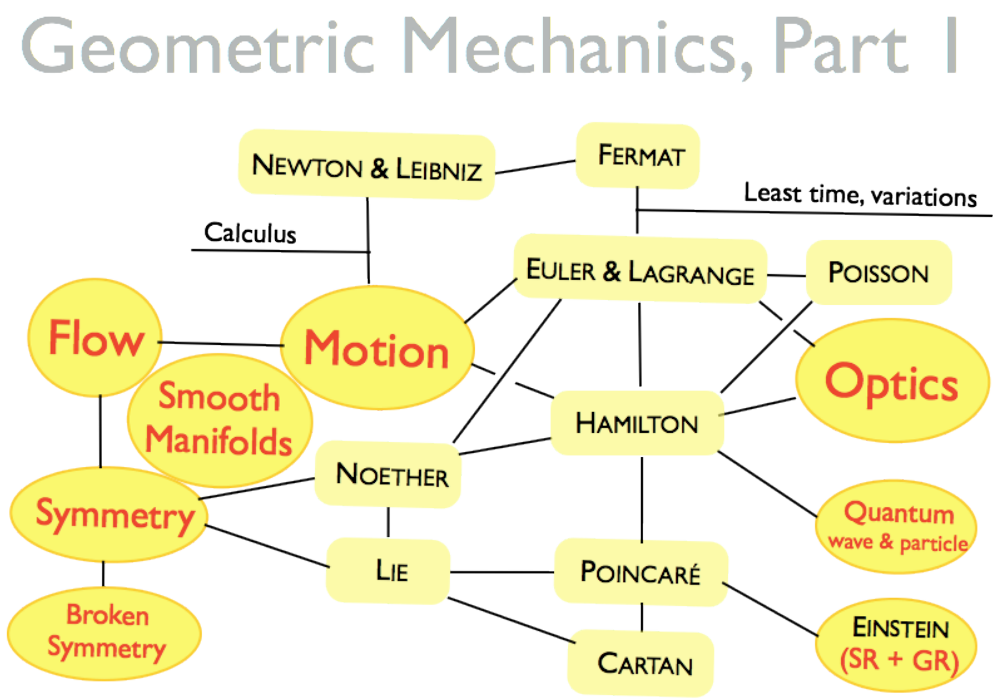


FIGURE 1. The fabric of geometric mechanics is woven by a network of fundamental contributions by at least a dozen people to the dual fields of optics and motion.

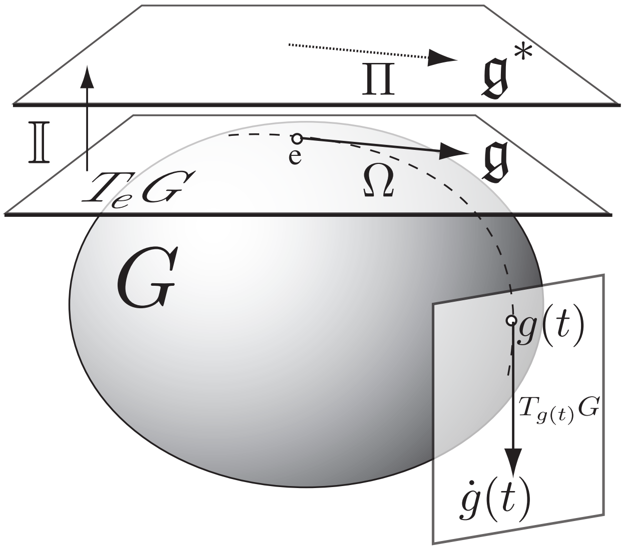
1. Mathematical setting

- The mathematical setting for geometric mechanics involves manifolds, (matrix) Lie groups and (later) diffeomorphisms
  - Manifold $M \simeq_{loc} \mathbb{R}^n$ e.g., $n = 1$ (scalars), $n = m$ (m-vectors), $n = m \times m$ (matrices),
  - Motion equation on TM: $\dot{q}(t) = f(q) \Longrightarrow$ transformation theory (pullbacks and all that)
  - Hamilton's principle for Lagrangian $L : TM \to \mathbb{R}$ vector fields
    * Euler–Lagrange equations on T^*M
    * Hamilton's canonical equations on T^*M
    * Euler–Poincaré eqns on $T_e^*G \simeq \mathfrak{g}^*$ for reduced Lagrangian $\ell : \mathfrak{g} \to \mathbb{R}$, e.g., *rigid body*.

A Lie group G is a manifold. Its tangent space at the identity T_eG is its Lie algebra $\mathfrak{g}$.

2. Euler–Poincaré Theorem

The definition of an invariant (or symmetric) function under a group action is as follows:



Definition

2.1. *Let G act on TG by left translation. A function $F : TG \to \mathbb{R}$ is called* ***left invariant*** *if and only if*

$$F(h(g,\dot g)) = F(g,\dot g) \quad \textit{for all } (g,\dot g) \in TG\,,$$

where

$$h(g,\dot g) := (hg, h\dot g)\,.$$

Example

2.2. *If the Lagrangian $L(g,\dot g) : TG \to \mathbb{R}$ in Hamilton's principle $\delta S = 0$ with $S = \int L dt$ is left invariant under the Lie group G, then:*

$$L(g,\dot g) = L(g^{-1}g, g^{-1}\dot g) = L(e, g^{-1}\dot g) = L(e,\xi) \quad \textit{for all} \quad (g,\dot g) \in TG,$$

where $\xi := g^{-1}\dot g$. Note that in this case the Lagrangian satisfies

$$L(g,\dot g) = L(e,\xi),$$

so it is left-invariant *under G.*

The Euler-Lagrange equation.

$$\frac{d}{dt}\frac{\delta L}{\delta \dot g} = \frac{\delta L}{\delta g}$$

can be re-expressed as

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

where l is defined to be the restriction of L to $\mathfrak{g}$:

$$l : \mathfrak{g} \to \mathbb{R}\,, \quad l(\xi) := L(e,\xi) \qquad \text{for all } g \in \xi\,.$$

The following theorem can be easily verified [HoMaRa1998]:

Theorem

2.3 (Euler–Poincaré reduction)**.** *Let G be a matrix Lie group and let $L : TG \to \mathbb{R}$ be a* ***left-invariant*** *Lagrangian. Define the* ***reduced Lagrangian****,*

$$l : \mathfrak{g} \to \mathbb{R}, \quad l(\xi) := L(e,\xi)\,,$$

as the restriction of L to $\mathfrak{g}$. For a curve $g(t) \in G$, let

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

Then, the following four statements are equivalent:

(i) *The variational principle*

$$\delta \int_a^b L(g(t), \dot{g}(t))\mathrm{d}t = 0$$

holds, for variations among paths with fixed endpoints.

(ii) *$g(t)$ satisfies the Euler–Lagrange equations for Lagrangian L defined on G.*

(iii) *The variational principle*

$$\delta \int_a^b l(\xi(t))\mathrm{d}t = 0$$

holds on $\mathfrak{g}$, using variations of the form

$$\delta\xi = \dot{\eta} + [\xi, \eta]\,,$$

where $\eta(t)$ is an arbitrary path in $\mathfrak{g}$ that vanishes at the endpoints in time, i.e. $\eta(a) = 0 = \eta(b)$.

(iv) *The (left invariant)* ***Euler–Poincaré equations*** *hold:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta\xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta\xi}\,, \tag{2.1}$$

where $\langle \mathrm{ad}^*_\xi \mu,\, \eta\rangle := \langle \mu,\, \mathrm{ad}_\xi \eta\rangle$, for $\mu \in \mathfrak{g}^*$ and $\xi,\, \eta \in \mathfrak{g}$.

Proof. A direct computation proves Theorem 2.3. Later, we will explain the source of the constraint $\delta\xi = \dot{\eta} + [\xi, \eta]$ 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, stationarity $\delta S = 0$ 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 (2.3) in the statement of the theorem. □

Exercise. Prove the Euler–Poincaré reduction Theorem 2.3 when G is a matrix Lie group. ★

Corollary

2.4 (Noether's theorem for Euler–Poincaré). *If η is an in finitesimal 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 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 are satisfied. □

Remark

2.5. *The form of the variation in* $\delta\xi = \dot\eta + [\xi, \eta]$ *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, cf. Proposition (3.3) for the rigid body,

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

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

$$\xi\,' - \dot\eta = [\,\xi\,,\,\eta\,] \quad \textit{with} \quad [\,\xi\,,\,\eta\,] := \xi\,\eta - \eta\,\xi = \mathrm{ad}_\xi\,\eta\,. \tag{2.2}$$

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

Remark

2.6. *A similar statement holds, with obvious changes for* ***right-invariant*** *Lagrangian systems on* TG*. In this case the Euler-Poincaré equations are given by:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta\xi} = -\,\mathrm{ad}^*_\xi\frac{\delta l}{\delta\xi}\,, \tag{2.3}$$

with the opposite sign.

Reconstruction.

Definition

2.7. *The reconstruction of the solution* $g(t)$ *of the Euler–Lagrange equations, with initial conditions* $g(0) = g_0$ *and* $\dot g(0) = v_0$*, is as follows:*

First, solve the initial value problem for the left-invariant Euler–Poincaré equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta\xi} = ad^*_\xi\frac{\delta l}{\delta\xi} \quad \textit{with} \quad \xi(0) = \xi_0 := g_0^{-1}v_0\,.$$

Second, using the solution for $\xi(t)$ *of the equation above, find the curve* $g(t) \in G$ *by solving the* ***reconstruction equation***

$$\dot g(t) = g(t)\xi(t) \quad \textit{with} \quad g(0) = g_0\,,$$

which is a differential equation with time-dependent coefficients.

Exercise. Write out the proof of the Euler–Poincaré reduction theorem for right-invariant Lagrangians and describe the corresponding reconstruction procedure. ★

Exercise. Consider the following action of a Lie group G on a product space $G \times Y$, where Y is some manifold:

$$(g\,,(h, y)) \to (gh, y).$$

Let $L : T(G \times Y) \to \mathbb{R}$ be invariant with respect to this action. Define $l : \mathfrak{g} \times TY \to \mathbb{R}$ as the restriction of L, i.e.

$$l(\xi, y, \dot y) := L(e, \xi, y, \dot y).$$

Deduce the reduced Hamilton's principle for l and show that the equations of motion are given by

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi}\,, \qquad \frac{d}{dt}\frac{\delta l}{\delta \dot y} = \frac{\delta l}{\delta y}\,.$$

★

WHAT WILL WE INVESTIGATE ABOUT THE RIGID BODY (RB)?

(1) Euler–Poincaré eqn on $\mathfrak{so}(3)^* \simeq \mathbb{R}^3$
(2) Hamilton-Pontryagin matrix form
(3) Noether theorem (coadjoint motion)
(4) Isospectral eigenvalue problem
(5) Manakov's matrix commutator form
(6) Hamiltonian forms (both Lie-Poisson and Nambu)
(7) Clebsch variational form (momentum map)
(8) RB Variants: $SO(4)$, $SP(2)$, $\mathbb{C}^2$, $\mathbb{C}^3$
(9) Bichrons (2 time variables)
(10) Coupled RBs
(11) Including potential energy
(12) Euler–Poincaré eqn on $\mathfrak{X}^*(\mathbb{R})$

3. Euler–Poincaré form of rigid-body motion

Euler's equations for rigid-body motion in principal axis coordinates, without external torques are

$$\begin{aligned} I_1\dot\Omega_1 &= (I_2 - I_3)\Omega_2\Omega_3,\\ I_2\dot\Omega_2 &= (I_3 - I_1)\Omega_3\Omega_1,\\ I_3\dot\Omega_3 &= (I_1 - I_2)\Omega_1\Omega_2, \end{aligned} \tag{3.1}$$

or, equivalently,

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega}\times\mathbf{\Omega}\,. \tag{3.2}$$

Here $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is the body angular velocity vector and $\mathbb{I} = \mathrm{diag}(I_1, I_2, I_3)$ is the moment of inertia tensor, which is diagonal in the principal axis frame of the rigid body.

Quadratic form. The moment of inertia $\mathbb{I}$ defines the following *quadratic form* $\mathbb{I}\mathbf{a}\cdot\mathbf{b}$ associated to the bilinear symmetric form for $\mathbb{R}^3$ vectors $\mathbf{a}$ and $\mathbf{b}$ in the body's principal axis frame,

$$(\mathbf{a},\,\mathbf{b}) := \int_{\mathcal{B}} \rho_0(\mathbf{X})(\mathbf{a}\times\mathbf{X})\cdot(\mathbf{b}\times\mathbf{X})d^3\mathbf{X} = \mathbb{I}\mathbf{a}\cdot\mathbf{b} = a^i\mathbb{I}_{ij}b^j\,. \tag{3.3}$$

Riemannian metric. Thus, the body's distribution of mass density $\rho_0(\mathbf{X})$ induces a Riemannian metric $\mathbb{I}$ for lowering indices of vectors in the body frame. That is, $\mathbb{I}:\mathbb{R}^3\to\mathbb{R}^{3^*}\simeq\mathbb{R}^3$. By the hat map then $\mathbb{I}:\mathfrak{so}(3)\to\mathfrak{so}(3)^*\simeq\mathbb{R}^3$.

Question? We ask whether Equations (3.1) may be expressed using Hamilton's principle on $\mathbb{R}^3$. For this, we will need to define the variational derivative of a functional $S[(\mathbf{\Omega}]$.

Definition

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

$$\delta S[\mathbf{\Omega}] := \lim_{s\to 0}\frac{S[\mathbf{\Omega}+s\delta\mathbf{\Omega}] - S[\mathbf{\Omega}]}{s} = \frac{d}{ds}\Big|_{s=0} S[\mathbf{\Omega}+s\delta\mathbf{\Omega}] =: \Big\langle \frac{\delta S}{\delta\mathbf{\Omega}}\,, \delta\mathbf{\Omega}\Big\rangle\,,$$

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

Theorem

3.2 (Euler's rigid-body equations)**.**
Euler's rigid-body equations are equivalent to Hamilton's principle

$$\delta S(\boldsymbol{\Omega}) = \delta \int_a^b l(\boldsymbol{\Omega})\, dt = 0, \tag{3.4}$$

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

$$l(\boldsymbol{\Omega}) = \frac{1}{2}(\boldsymbol{\Omega},\boldsymbol{\Omega}) := \frac{1}{2}\,\mathbb{I}\boldsymbol{\Omega}\cdot\boldsymbol{\Omega} = \frac{1}{2}\left(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2\right), \tag{3.5}$$

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

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega}\times\boldsymbol{\Xi}\,, \tag{3.6}$$

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

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

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

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

$$\boldsymbol{\Xi}(a) = 0 = \boldsymbol{\Xi}(b)\,.$$

Since $\boldsymbol{\Xi}$ is otherwise arbitrary, (3.4) is equivalent to

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

which recovers Euler's Equations (3.1) in vector form. □

3.1. **The hat map.** The Lie algebra $(\mathbb{R}^3, \times)$ of vectors in $\mathbb{R}^3$ with vector product $\times$ maps to the Lie algebra $(\mathfrak{so}(3), [\cdot,\cdot])$ of 3×3 skew-symmetric matrices with matrix commutator bracket $[\,\cdot\,,\,\cdot\,]$, by the following linear isomorphism, called the *hat map*,

$$\mathbf{u} := (u^1, u^2, u^3) \in \mathbb{R}^3 \quad\mapsto\quad \widehat{u} := \begin{bmatrix} 0 & -u^3 & u^2 \\ u^3 & 0 & -u^1 \\ -u^2 & u^1 & 0 \end{bmatrix} \in \mathfrak{so}(3)\,. \tag{3.7}$$

In matrix and vector components, the linear isomorphism is

$$\widehat{u}_{ij} := -\,\epsilon_{ijk}u^k\,,$$

where ϵ_{ijk} is the totally antisymmetric Levi-Civita symbol, with $\epsilon_{123} = 1$. Equivalently, this isomorphism is given by

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

The hat map $\widehat{\ }: (\mathbb{R}^3,\times) \to (\mathfrak{so}(3), [\cdot,\cdot])$ may also be defined using

$$\widehat{u} = \mathbf{u}\cdot\widehat{\boldsymbol{J}} = u^a\widehat{J}_a\,,$$

which holds for the $\mathfrak{so}(3)$ basis set of skew-symmetric 3×3 matrices $\widehat{J}_a \in \mathfrak{so}(3)$, with $a = 1,2,3$ defined in the relation (3.7).

Exercise. Verify the following formulas for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$:

$$
\begin{aligned}
(\mathbf{u}\times\mathbf{v})\widehat{\ } &= \widehat{u}\,\widehat{v} - \widehat{v}\,\widehat{u} =: [\widehat{u},\widehat{v}]\,,\\
[\widehat{u},\widehat{v}]\,\mathbf{w} &= (\mathbf{u}\times\mathbf{v})\times\mathbf{w}\,,\\
\big((\mathbf{u}\times\mathbf{v})\times\mathbf{w}\big)\widehat{\ } &= \big[[\widehat{u},\widehat{v}]\,,\,\widehat{w}\big]\,,\\
\mathbf{u}\cdot\mathbf{v} &= -\frac{1}{2}\,\mathrm{trace}(\widehat{u}\,\widehat{v})\\
&=: \big\langle\,\widehat{u}\,,\,\widehat{v}\,\big\rangle\,,
\end{aligned}
$$

in which the dot product of vectors is also the natural pairing of 3×3 skew-symmetric matrices. ★

Exercise. (Jacobi identity under the hat map) Verify that the Jacobi identity for the cross product of vectors in $\mathbb{R}^3$ is equivalent to the Jacobi identity for the commutator product of 3×3 skew matrices by proving the following identity satisfied by the hat map,

$$
\begin{aligned}
&\big((\mathbf{u}\times\mathbf{v})\times\mathbf{w} + (\mathbf{v}\times\mathbf{w})\times\mathbf{u} + (\mathbf{w}\times\mathbf{u})\times\mathbf{v}\big)\widehat{\ }\\
&= 0 = \big[[\widehat{u},\widehat{v}]\,,\,\widehat{w}\big] + \big[[\widehat{v},\widehat{w}]\,,\,\widehat{u}\big] + \big[[\widehat{w},\widehat{u}]\,,\,\widehat{v}\big]\,.
\end{aligned}
$$

★

3.2. **Restricted variations.**

Proposition

3.3 (Derivation of the restricted variation)**.**
The restricted variation in (3.6) arises via the following steps:
- *(i) Vary the definition of the body angular velocity, $\widehat{\Omega} = O^{-1}\dot{O}$.*
- *(ii) Take the time derivative of the variation, $\widehat{\Xi} = O^{-1}O'$.*
- *(iii) Use the equality of cross derivatives, $O^{\cdot\,\prime} = d^2O/dtds = O^{\prime\,\cdot}$.*
- *(iv) Apply the hat map.*

Proof. One computes directly that

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

On taking the difference, the cross derivatives cancel and one finds a variational formula equivalent to (3.6),

$$
\widehat{\Omega}\,' - \widehat{\Xi}^{\,\cdot} = \big[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,\big] \quad\text{with}\quad \big[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,\big] := \widehat{\Omega}\,\widehat{\Xi} - \widehat{\Xi}\,\widehat{\Omega}\,. \tag{3.8}
$$

Under the bracket relation

$$
\big[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,\big] = (\boldsymbol{\Omega}\times\boldsymbol{\Xi})\widehat{\ }
$$

for the hat map, this equation recovers the vector relation (3.6) in the form

$$
\boldsymbol{\Omega}\,' - \dot{\boldsymbol{\Xi}} = \boldsymbol{\Omega}\times\boldsymbol{\Xi}\,. \tag{3.9}
$$

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

$$
\mathbb{I}\dot{\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega}\,, \tag{3.10}
$$

do follow from the variational principle (3.4) with variations of the form (3.6) derived from the definition of body angular velocity $\widehat{\Omega}$. □

Remark

3.4. *The body angular velocity vector is expressed in terms of the spatial angular velocity vector by* $\boldsymbol{\Omega}(t) = O^{-1}(t)\boldsymbol{\omega}(t)$. *Consequently, the kinetic energy Lagrangian in (3.5) transforms as*

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

where

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

Exercise. Show that Hamilton's principle for the action

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

yields conservation of spatial angular momentum

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

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

Exercise. (Noether's theorem for the rigid body) What conservation law does Noether's theorem imply for the rigid-body Equations (3.2)?

Hint: Transform the endpoint terms arising on integrating the variation δS by parts in the proof of Theorem 3.2 into the spatial representation by setting $\boldsymbol{\Xi} = O^{-1}(t)\boldsymbol{\Gamma}$ and $\boldsymbol{\Omega} = O^{-1}(t)\boldsymbol{\omega}$. ★

Remark

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

$$\dot{O}(t) = O(t)\widehat{\Omega}(t)\,, \tag{3.11}$$

where the left-invariant skew-symmetric 3×3 *matrix* $\widehat{\Omega}$ *is defined by the hat map*

$$(O^{-1}\dot{O})_{jk} = \widehat{\Omega}_{jk} = -\,\Omega_i\epsilon_{ijk}\,. \tag{3.12}$$

Equation (3.11) is the ***reconstruction formula*** *for* $O(t) \in SO(3)$.

Once the time dependence of $\boldsymbol{\Omega}(t)$ *and hence* $\widehat{\Omega}(t)$ *is determined from the Euler equations, solving formula (3.11) as a linear differential equation with time-dependent coefficients yields the integral curve* $O(t) \in SO(3)$ *for the orientation of the rigid body.*

Exercise. [***Motion on*** $SO(4)$]
Write out the Euler–Poincaré equations in matrix form for a free rigid body fixed at its centre of mass in a 4-dimensional space. Use the analogue of the 'hat' map for $\mathfrak{so}(4)$ and write the $\mathbb{R}^6$ vector representation of the equations. ★

3.3. **Hamilton–Pontryagin constrained variations.** Formula (3.8) for the variation $\widehat{\Omega}$ of the skew-symmetric matrix

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

may be imposed as a constraint in Hamilton's principle and thereby provide a variational derivation of Euler's Equations (3.1) for rigid-body motion in principal axis coordinates. This constraint is incorporated into the matrix Euler equations, as follows.

Proposition

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

$$\frac{d\Pi}{dt} = -\left[\,\widehat{\Omega}\,,\,\Pi\,\right] \quad \textit{with} \quad \Pi = \mathbb{I}\widehat{\Omega} = \frac{\delta l}{\delta\widehat{\Omega}}\,, \tag{3.13}$$

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

$$l = \frac{1}{2}\Big\langle\,\mathbb{I}\widehat{\Omega}\,,\,\widehat{\Omega}\,\Big\rangle\,. \tag{3.14}$$

Here, the bracket

$$\big[\,\widehat{\Omega}\,,\,\Pi\big] := \widehat{\Omega}\Pi - \Pi\widehat{\Omega} \tag{3.15}$$

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

$$\Big\langle\,\Pi\,,\,\widehat{\Omega}\,\Big\rangle =: \frac{1}{2}\,\mathrm{trace}\,\Big(\Pi^T\,\widehat{\Omega}\Big)\,. \tag{3.16}$$

Remark

3.7. *Note that the symmetric part of* Π *does not contribute in the pairing and if set equal to zero initially, it will remain zero.*

Proposition

3.8 (Constrained variational principle)**.**
The matrix Euler Equations (3.13) are equivalent to stationarity $\delta S = 0$ *of the following* ***constrained action****:*

$$\begin{aligned} S(\widehat{\Omega}, O, \dot{O}, \Pi) &= \int_a^b l(\widehat{\Omega}, O, \dot{O}, \Pi)\,dt \\ &= \int_a^b \Big[l(\widehat{\Omega}) + \big\langle\,\Pi\,,\,(O^{-1}\dot{O} - \widehat{\Omega}\,)\,\big\rangle\Big]\,dt\,. \end{aligned} \tag{3.17}$$

Remark

3.9. *The integrand of the constrained action in (3.17) is similar to the formula for the Legendre transform, but its functional dependence is different. This variational approach is related to the classic* ***Hamilton–Pontryagin principle*** *for control theory. It has also be used to develop algorithms for geometric numerical integrations of rotating motion.*

Proof. The variations of S in formula (3.17) are given by

$$\begin{aligned} \delta S \;=\; & \int_a^b \bigg\{\Big\langle\,\frac{\partial l}{\partial\widehat{\Omega}} - \Pi\,,\,\delta\widehat{\Omega}\,\Big\rangle \\ & + \Big\langle\,\delta\Pi\,,\,(O^{-1}\dot{O} - \Omega)\,\Big\rangle + \Big\langle\,\Pi\,,\,\delta(O^{-1}\dot{O})\,\Big\rangle\bigg\}\,dt\,, \end{aligned}$$

where

$$\delta(O^{-1}\dot{O}) = \widehat{\Xi}^{\,\cdot} + \big[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,\big]\,, \tag{3.18}$$

and $\widehat{\Xi} = (O^{-1}\delta O)$ from Equation (3.8).

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

$$
\begin{aligned}
\int_a^b \Big\langle \Pi\,,\, \delta(O^{-1}\dot{O}) \Big\rangle\, dt
&= \int_a^b \Big\langle \Pi\,,\, \widehat{\Xi}^{\,\cdot} + [\,\widehat{\Omega}\,,\,\widehat{\Xi}\,] \Big\rangle\, dt \\
&= \int_a^b \Big\langle -\,\Pi^{\,\cdot} - [\,\widehat{\Omega}\,,\,\Pi\,]\,,\,\widehat{\Xi} \Big\rangle\, dt \\
&\quad + \Big\langle \Pi\,,\,\widehat{\Xi} \Big\rangle\Big|_a^b\,,
\end{aligned}
\tag{3.19}
$$

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

$$
\operatorname{trace}\left(\Pi^T\,\widehat{\Xi}\,\widehat{\Omega}\right) = \operatorname{trace}\left(\widehat{\Omega}\,\Pi^T\,\widehat{\Xi}\right)\,. \tag{3.20}
$$

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

$$
\frac{\partial l}{\partial \widehat{\Omega}} = \Pi\,, \quad O^{-1}\frac{dO}{dt} = \widehat{\Omega}\,, \quad \frac{d\Pi}{dt} = -[\,\widehat{\Omega}\,,\,\Pi\,]\,. \tag{3.21}
$$

These are the Euler rigid body equations in matrix form on $SO(n)$. □

Remark

3.10 (Interpreting the formulas in (3.21)).
*The first formula in (3.21) defines the angular momentum matrix Π as the **fibre derivative** of the Lagrangian with respect to the angular velocity matrix $\widehat{\Omega}$. The second formula is the reconstruction formula (3.11) for the solution curve $O(t) \in SO(3)$, given the solution $\widehat{\Omega}(t) = O^{-1}\dot{O}$. And the third formula is Euler's equation for rigid-body motion in matrix form.*

Exercise. Use the fibre derivative relation to compute the Hamiltonian $h(\Pi)$ via the Legendre transform,

$$
h(\Pi) = \langle \Pi,\, \widehat{\Omega} \rangle - l(\widehat{\Omega}) \tag{3.22}
$$

then express the matrix Euler rigid body equations in Hamiltonian form as a Poisson bracket relation. Notice that equation (3.17) for the Hamilton–Pontryagin principle also contains this Legendre transform. ★

Answer. The Hamiltonian $h(\Pi)$ satisfies

$$
dh(\Pi) = \left\langle d\Pi,\, \frac{\partial h}{\partial \Pi} \right\rangle = \left\langle d\Pi,\, \widehat{\Omega} \right\rangle - \left\langle \Pi - \frac{\partial l}{\partial \widehat{\Omega}},\, d\widehat{\Omega} \right\rangle
$$

so that

$$
\Pi = \frac{\partial l}{\partial \widehat{\Omega}}\,, \quad \frac{\partial h}{\partial \Pi} = \widehat{\Omega}
$$

The matrix Euler rigid body equations (3.21) are then expressed as

$$
\frac{d\Pi}{dt} = -\left[\frac{\partial h}{\partial \Pi}\,,\,\Pi\right] \tag{3.23}
$$

and a function $f(\Pi)$ has time derivative

$$
\begin{aligned}
\frac{d}{dt} f(\Pi) &= -\left\langle \frac{\partial f}{\partial \Pi},\, \left[\frac{\partial h}{\partial \Pi}\,,\,\Pi\right] \right\rangle \\
&= -\left\langle \Pi,\, \left[\frac{\partial f}{\partial \Pi}\,,\,\frac{\partial h}{\partial \Pi}\right] \right\rangle =: \Big\{f,\, h\Big\}(\Pi)\,.
\end{aligned}
\tag{3.24}
$$

The last expression defines the ***Lie-Poisson bracket***, which inherits the Jacobi property from the matrix commutator. ▲

Exercise. Use equation (3.22) to rewrite the Hamilton–Pontryagin variational principle (3.17) as $\delta S = 0$ for the action

$$S(O^{-1}\dot{O}, \Pi) = \int_a^b \Big(\big\langle\, \Pi \,,\, O^{-1}\dot{O} \,\big\rangle - h(\Pi) \Big)\, dt\,. \tag{3.25}$$

Take the variations using (3.18) and recover the Hamiltonian form of the matrix Euler rigid body equations (3.21). How does this compare with the results for $\delta S = 0$ with $S = \int_a^b \langle p, \dot{q}\rangle - H(q,p)\, dt$? ★

Exercise. Write the Lie-Poisson bracket in (3.24) in three dimensions for $\mathfrak{so}(3)^*$ in $\mathbb{R}^3$ vector form by using the hat map. Thereby, discover the Nambu bracket form of the rigid body equations. ★

Answer. In $\mathbb{R}^3$ vector form the Lie-Poisson bracket in (3.24) becomes

$$\frac{d}{dt} f(\mathbf{\Pi}) = -\,\frac{\partial f}{\partial \mathbf{\Pi}} \cdot \frac{\partial h}{\partial \mathbf{\Pi}} \times \mathbf{\Pi} = -\,\mathbf{\Pi} \cdot \frac{\partial f}{\partial \mathbf{\Pi}} \times \frac{\partial h}{\partial \mathbf{\Pi}} =: \Big\{ f,\, h \Big\}(\mathbf{\Pi})\,. \tag{3.26}$$

Euler's equations are recovered by setting $f(\mathbf{\Pi}) = \mathbf{\Pi}$

$$\frac{d\mathbf{\Pi}}{dt} = -\,\frac{\partial h}{\partial \mathbf{\Pi}} \times \mathbf{\Pi} =: \Big\{ \mathbf{\Pi},\, h \Big\}\,. \tag{3.27}$$

If we write $c(\mathbf{\Pi}) = \frac{1}{2}\|\mathbf{\Pi}\|^2$, then the Lie-Poisson bracket in (3.26) may be expressed in ***Nambu bracket*** form,

$$\frac{d}{dt} f(\mathbf{\Pi}) = -\,\frac{\partial c}{\partial \mathbf{\Pi}} \cdot \frac{\partial f}{\partial \mathbf{\Pi}} \times \frac{\partial h}{\partial \mathbf{\Pi}} =: \Big\{ c,\, f,\, h \Big\}(\mathbf{\Pi})\,, \tag{3.28}$$

which is the triple scalar product of gradients in $\mathbf{\Pi}$.[1] ▲

Remark

3.11 (Interpreting the endpoint terms in (3.19)).
We transform the endpoint terms in (3.19), arising on integrating the variation δS by parts in the proof of Theorem 3.2 into the spatial representation by setting $\widehat{\Xi}(t) =: O(t)\widehat{\xi}O^{-1}(t)$ and $\widehat{\Pi}(t) =: O(t)\widehat{\pi}(t)O^{-1}(t)$, as follows:

$$\Big\langle\, \Pi \,,\, \widehat{\Xi} \,\Big\rangle = \mathrm{trace}\Big(\Pi^T\, \widehat{\Xi} \Big) = \mathrm{trace}\Big(\pi^T\, \widehat{\xi} \Big) = \Big\langle\, \pi \,,\, \widehat{\xi} \,\Big\rangle\,. \tag{3.29}$$

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

$$\pi(a) = \pi(b)\,. \tag{3.30}$$

*This is **Noether's theorem** for the rigid body.*

Theorem

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

$$\pi = O^{-1}(t)\Pi(t)O(t) =: \mathrm{Ad}^*_{O^{-1}(t)}\Pi(t). \tag{3.31}$$

[1] The Lie-Poisson and Nambu brackets introduced by discovery in these two exercises will be discussed further below.

Proof.

$$
\begin{aligned}
\frac{d}{dt}\Big\langle \pi\,,\,\widehat{\xi}\Big\rangle &= \frac{d}{dt}\Big\langle O^{-1}\Pi O\,,\,\widehat{\xi}\Big\rangle = \frac{d}{dt}\,\mathrm{trace}\Big(\Pi^T O^{-1}\widehat{\xi}O\Big)\\
&= \Big\langle \frac{d}{dt}\Pi + \big[\,\widehat{\Omega}\,,\,\Pi\,\big]\,,\,O^{-1}\widehat{\xi}O\Big\rangle = 0\\
&=: \Big\langle \frac{d}{dt}\Pi - \mathrm{ad}^*_{\widehat{\Omega}}\Pi\,,\,\mathrm{Ad}_{O^{-1}}\widehat{\xi}\Big\rangle,\\
\frac{d}{dt}\Big\langle \mathrm{Ad}^*_{O^{-1}}\Pi\,,\,\widehat{\xi}\Big\rangle &= \Big\langle \mathrm{Ad}^*_{O^{-1}}\Big(\frac{d}{dt}\Pi - \mathrm{ad}^*_{\widehat{\Omega}}\Pi\Big)\,,\,\widehat{\xi}\Big\rangle. \qquad (3.32)
\end{aligned}
$$

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

Remark

3.13. *This proof of Noether's theorem for the rigid body when the constrained Hamilton–Pontryagin action is invariant under spatial rotations also proves a general result in Equation (3.32), with $\widehat{\Omega} = O^{-1}\dot{O}$ for a Lie group O, that*

$$
\boxed{\frac{d}{dt}\Big(\mathrm{Ad}^*_{O^{-1}}\Pi\Big) = \mathrm{Ad}^*_{O^{-1}}\Big(\tfrac{d}{dt}\Pi - \mathrm{ad}^*_{\widehat{\Omega}}\Pi\Big)} \qquad (3.33)
$$

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

$$
\boxed{\tfrac{d}{dt}\Pi - \mathrm{ad}^*_{\widehat{\Omega}}\Pi = 0\,, \quad \textit{has solution} \quad \Pi(t) = \mathrm{Ad}^*_{O(t)}\pi \quad \textit{with} \quad \pi = \Pi(0) \in \mathfrak{o}^*} \qquad (3.34)
$$

Exercise. Retrace the proof of the variational principle for the Euler–Poincaré equation, replacing the left-invariant quantity $\widehat{\Omega} = O^{-1}\dot{O}$ with the right-invariant quantity $\xi := \dot{g}g^{-1} \in \mathfrak{g}$ for $g \in G$, a Lie group.

Write Noether's theorem for a right-invariant Lagrangian $\ell(\xi)$.

★

Answer.
For a right-invariant system $\xi := \dot{g}g^{-1} \in \mathfrak{g}$ and $J := \delta\ell/\delta\xi \in \mathfrak{g}^*$ for a reduced Lagrangian $\ell(\xi)$, the Euler-Poincaré theorem delivers the following equation, which has one sign difference from before,

$$
\frac{d}{dt}J(t) + \mathrm{ad}^*_{\xi}J(t) = 0\,, \quad \text{with solution} \quad J(t) = \mathrm{Ad}^*_{g(t)}J(0). \qquad (3.35)
$$

▲

3.4. Manakov's formulation of the $SO(n)$ rigid body.

Proposition

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

$$
\frac{dM}{dt} = \big[\,M\,,\,\Omega\,\big] \quad \textit{with} \quad M = \mathbb{A}\Omega + \Omega\mathbb{A}\,, \qquad (3.36)
$$

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

Proof. Manakov's commutator form of the $SO(n)$ rigid-body Equations (3.36) follows as the Euler–Lagrange equations for Hamilton's principle $\delta S = 0$ with $S = \int l\, dt$ for the Lagrangian

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

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

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

after cyclically permuting the order of matrix multiplication under the trace and substituting $M := \mathbb{A}\Omega + \Omega\mathbb{A}$. Using the variational formula (3.18) for $\delta\Omega$ now leads to

$$\delta S \;=\; -\frac{1}{2}\int_a^b \mathrm{tr}\big((\Xi^{\cdot} + \Omega\Xi - \Xi\Omega)M\big)\,dt\,.$$

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

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

Finally, invoking stationarity for arbitrary Ξ implies the commutator form (3.36). □

3.5. **Matrix Euler–Poincaré equations.** Manakov's commutator form of the rigid-body equations recalls much earlier work by Poincaré [Po1901], who also noticed that the matrix commutator form of Euler's rigid-body equations suggests an additional mathematical structure going back to Lie's theory of groups of transformations depending continuously on parameters. In particular, Poincaré [Po1901] remarked that the commutator form of Euler's rigid-body equations would make sense for any Lie algebra, not just for $\mathfrak{so}(3)$. The proof of Manakov's commutator form (3.36) by Hamilton's principle is essentially the same as Poincaré's proof in [Po1901], which is translated into English and discussed thoroughly in [JKLOR2011].

Theorem

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

$$\frac{dM}{dt} = \big[\,M\,,\,\Omega\,\big] \quad \textit{with} \quad M = \frac{\delta l}{\delta\Omega}\,, \tag{3.37}$$

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

Proof. The proof here is the same as the proof of Manakov's commutator formula (3.36) via Hamilton's principle, modulo replacing $O^{-1}\dot{O} \in so(n)$ with $g^{-1}\dot{g} \in \mathfrak{g}$. □

Remark

3.16. *Poincaré's observation leading to the matrix Euler–Poincaré Equation (3.37) was reported in two pages with no references* [Po1901]. *The proof above shows that the matrix Euler–Poincaré equations possess a natural variational principle. Note that if $\Omega = g^{-1}\dot{g} \in \mathfrak{g}$, then $M = \delta l/\delta\Omega \in \mathfrak{g}^*$, where the dual is defined in terms of the matrix trace pairing.*

3.6. **An isospectral eigenvalue problem for the $SO(n)$ rigid body.** The solution of the $SO(n)$ rigid-body dynamics

$$\frac{dM}{dt} = [\,M\,,\,\Omega\,] \quad \text{with} \quad M = \mathbb{A}\Omega + \Omega\mathbb{A}\,, \tag{3.38}$$

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

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

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

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

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

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

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

  The proof of this statement follows from the corresponding property of similarity transformations.
- Its matrix invariants are preserved:

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

  for every non-negative integer power K.

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

Theorem

3.17. *Isospectrality allows the quadratic rigid-body dynamics (3.38) on $SO(n)$ to be rephrased as a system of two* coupled linear equations*: the eigenvalue problem for M and an evolution equation for its eigenvectors ψ, as follows:*

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

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

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

Now substitute the second equation to recover the $SO(n)$ rigid-body dynamics (3.38). □

3.7. **Manakov's integration of the $SO(n)$ rigid body.** Manakov [Man1976] observed that Equations (3.36) may be "deformed" into

$$\frac{d}{dt}(M + \lambda A) = [(M + \lambda A), (\Omega + \lambda B)]\,, \tag{3.39}$$

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

- The coefficent of λ^2 is

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

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

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

- The coefficent of λ is

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

  Therefore, by antisymmetry of M and Ω,

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

  which implies that

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

  Hence, angular velocity Ω is a linear function of angular momentum, M.

- Finally, the coefficent of λ^0 recovers the Euler equation
$$\frac{dM}{dt} = [M, \Omega]\,,$$
but now with the restriction that the moments of inertia are of the form
$$\Omega_{ij} = \frac{b_i - b_j}{a_i - a_j} M_{ij} \quad \text{(no sum)}.$$
This relation turns out to possess only five free parameters for $n = 4$.

Under these conditions, Manakov's deformation of the $SO(n)$ rigid-body equation into the commutator form (3.39) implies for every non-negative integer power K that

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

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

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

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

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

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

4. Hamiltonian form of rigid-body motion

The ***Legendre transform*** of the Lagrangian (3.5) in the variational principle (3.4) for Euler's rigid-body dynamics (3.10) on $\mathbb{R}^3$ will reveal its well-known Hamiltonian formulation.

Definition

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

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

The Legendre transformation defines the ***body angular momentum*** by the variations of the rigid body's reduced Lagrangian with respect to the body angular velocity. For the Lagrangian in (3.4), the $\mathbb{R}^3$ components of the body angular momentum are

$$\Pi_i = I_i \Omega_i = \frac{\partial l}{\partial \Omega_i}\,, \quad i = 1, 2, 3. \tag{4.1}$$

4.1. Hamiltonian form and Poisson bracket.

Definition

4.2 (Dynamical systems in Hamiltonian form)**.**
A dynamical system on a manifold M

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

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

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

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

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

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

Definition

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

- *It is* ***bilinear****.*
- *It is* ***skew-symmetric****,* $\{F\,,\,H\} = -\,\{H\,,\,F\}$*.*
- *It satisfies the* ***Leibniz rule*** *(product rule),*
$$\{FG\,,\,H\} = \{F\,,\,H\}G + F\{G\,,\,H\}\,,$$
for the product of any two functions F *and* G *on* M*.*
- *It satisfies the* ***Jacobi identity****,*
$$\{F\,,\,\{G\,,\,H\}\} + \{G\,,\,\{H\,,\,F\}\} + \{H\,,\,\{F\,,\,G\}\} = 0\,, \tag{4.2}$$
for any three functions F*,* G *and* H *on* M*.*

Remark

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

4.2. **Lie–Poisson Hamiltonian rigid-body dynamics.** The Legendre transform for this system is
$$h(\boldsymbol{\Pi}) := \boldsymbol{\Pi}\cdot\boldsymbol{\Omega} - l(\boldsymbol{\Omega})\,, \tag{4.3}$$
in terms of the vector dot product on $\mathbb{R}^3$. Hence, one finds the expected expression for the rigid-body Hamiltonian
$$h = \frac{1}{2}\boldsymbol{\Pi}\cdot\mathbb{I}^{-1}\boldsymbol{\Pi} := \frac{\Pi_1^2}{2I_1} + \frac{\Pi_2^2}{2I_2} + \frac{\Pi_3^2}{2I_3}\,. \tag{4.4}$$
The Legendre transform $\mathbb{F}l = \partial l/\partial\boldsymbol{\Omega} = \boldsymbol{\Pi}$ for this case is a diffeomorphism, so one may solve for the body angular velocity $\boldsymbol{\Omega}$ as the derivative of the reduced Hamiltonian with respect to the body angular momentum $\boldsymbol{\Pi}$ namely,
$$\frac{\partial h}{\partial\boldsymbol{\Pi}} = \mathbb{I}^{-1}\boldsymbol{\Pi} = \boldsymbol{\Omega}\,. \tag{4.5}$$
Hence, the reduced Euler–Lagrange equations for l may be expressed equivalently in angular momentum vector components in $\mathbb{R}^3$ and Hamiltonian h as
$$\frac{d}{dt}(\mathbb{I}\boldsymbol{\Omega}) = \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega} \Longleftrightarrow \dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi}\times\frac{\partial h}{\partial\boldsymbol{\Pi}} := \{\boldsymbol{\Pi}, h\}\,.$$
This expression suggests we introduce the following ***rigid-body Poisson bracket*** on functions of the $\boldsymbol{\Pi}$'s:
$$\{f, h\}(\boldsymbol{\Pi}) := -\,\boldsymbol{\Pi}\cdot\left(\frac{\partial f}{\partial\boldsymbol{\Pi}}\times\frac{\partial h}{\partial\boldsymbol{\Pi}}\right). \tag{4.6}$$
For the Hamiltonian (4.4), one checks that the Euler equations in terms of the rigid-body angular momenta,
$$\begin{aligned}
\dot{\Pi}_1 &= \left(\frac{1}{I_3} - \frac{1}{I_2}\right)\Pi_2\Pi_3\,,\\
\dot{\Pi}_2 &= \left(\frac{1}{I_1} - \frac{1}{I_3}\right)\Pi_3\Pi_1\,,\\
\dot{\Pi}_3 &= \left(\frac{1}{I_2} - \frac{1}{I_1}\right)\Pi_1\Pi_2\,,
\end{aligned} \tag{4.7}$$
that is, the equations
$$\dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi}\times\mathbb{I}^{-1}\boldsymbol{\Pi}\,, \tag{4.8}$$
are equivalent to
$$\dot{f} = \{f, h\}\,, \quad\text{with}\quad f = \boldsymbol{\Pi}\,.$$

4.3. **Lie–Poisson bracket.** The Poisson bracket proposed in (4.6) is an example of a ***Lie–Poisson bracket***.

It satisfies the defining relations of a Poisson bracket for a number of reasons, not least because it is the hat map to $\mathbb{R}^3$ of the following bracket defined by the general form in Equation (3.32) in terms of the $\mathfrak{so}(3)^* \times \mathfrak{so}(3)$ pairing $\langle\,\cdot\,,\,\cdot\,\rangle$ in Equation (3.19). Namely,

$$
\begin{aligned}
\frac{dF}{dt} &= \left\langle \frac{d}{dt}\Pi,\, \frac{\partial F}{\partial \Pi} \right\rangle = \left\langle \mathrm{ad}^*_{\widehat{\Omega}}\Pi,\, \frac{\partial F}{\partial \Pi} \right\rangle \\
&= \left\langle \Pi,\, \mathrm{ad}_{\widehat{\Omega}}\frac{\partial F}{\partial \Pi} \right\rangle = \left\langle \Pi,\, \left[\widehat{\Omega},\, \frac{\partial F}{\partial \Pi}\right] \right\rangle \\
&= -\left\langle \Pi,\, \left[\frac{\partial F}{\partial \Pi},\, \frac{\partial H}{\partial \Pi}\right] \right\rangle,
\end{aligned}
\tag{4.9}
$$

where we have used the equation corresponding to (4.5) under the inverse of the hat map

$$
\widehat{\Omega} = \frac{\partial H}{\partial \Pi}
$$

and applied antisymmetry of the matrix commutator. Writing Equation (4.9) as

$$
\frac{dF}{dt} = -\left\langle \Pi,\, \left[\frac{\partial F}{\partial \Pi},\, \frac{\partial H}{\partial \Pi}\right] \right\rangle =: \Big\{ F,\, H \Big\}
\tag{4.10}
$$

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

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

(4.11) $\{\Pi_i,\, \Pi_j\} =$

$\{\,\cdot\,,\,\cdot\,\}$	Π_1	Π_2	Π_3
Π_1	0	$-\Pi_3$	Π_2
Π_2	Π_3	0	$-\Pi_1$
Π_3	$-\Pi_2$	Π_1	0

or, in index notation, as

$$
\{\Pi_i\,,\, \Pi_j\} = -\epsilon_{ijk}\Pi_k = \widehat{\Pi}_{ij}\,.
\tag{4.12}
$$

Remark

4.5. *The Lie–Poisson bracket in the form (4.10) would apply to any Lie algebra. This Lie–Poisson Hamiltonian form of the rigid-body dynamics substantiates Poincaré's observation in* [Po1901] *that the corresponding equations could have been written on the dual of any Lie algebra by using the* ad^* *operation for that Lie algebra. See* [JKLOR2011] *for more discussion.*

The corresponding Poisson bracket in (4.6) in $\mathbb{R}^3$-vector form also satisfies the defining relations of a Poisson bracket because it is an example of a ***Nambu bracket***, to be discussed next.

4.4. **Nambu's $\mathbb{R}^3$ Poisson bracket.** The rigid-body Poisson bracket (4.6) is a special case of the Poisson bracket for functions of $\mathbf{x} \in \mathbb{R}^3$,

$$
\{f, h\} = -\,\nabla c \cdot \nabla f \times \nabla h\,.
\tag{4.13}
$$

This bracket generates the motion

$$
\dot{\mathbf{x}} = \{\mathbf{x}, h\} = \nabla c \times \nabla h\,.
\tag{4.14}
$$

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

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

$$\{f, h\} = -\nabla c \cdot \nabla f \times \nabla h \,. \tag{4.15}$$

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

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

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

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

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

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

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

Theorem

4.6 (Jacobi identity)**.**
The Nambu $\mathbb{R}^3$ bracket (4.15) satisfies the Jacobi identity.

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

$$\pounds_{X_G} X_H = [X_G, X_H] = -X_{\{G,H\}} \,.$$

Repeating the Lie derivative produces

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

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

Exercise. How is the $\mathbb{R}^3$ bracket related to the canonical Poisson bracket?
Hint: Restrict to level surfaces of the function $c(\mathbf{x})$. ★

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

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

Suppose the function $c(\mathbf{x})$ is chosen so that the $\mathbb{R}^3$ bracket (4.13) defines a proper Poisson bracket. What are the Casimirs for the $\mathbb{R}^3$ bracket (4.13)? Why? ★

Exercise. (Geometric interpretation of Nambu motion)

- Show that the Nambu motion equation (4.14)
$$\dot{\mathbf{x}} = \{\mathbf{x}, h\} = \nabla c \times \nabla h$$
for the $\mathbb{R}^3$ bracket (4.13) is invariant under a certain linear combination of the functions c and h. Interpret this invariance geometrically.
- Show that the rigid-body equations (4.7) for
$$\mathbb{I} = \mathrm{diag}(1,\,1/2,\,1/3)$$
may be interpreted as intersections in $\mathbb{R}^3$ of the spheres $x_1^2+x_2^2+x_3^2 = \text{constant}$ and the hyperbolic cylinders $x_1^2 - x_3^2 = \text{constant}$, as in Fig. 4.4.
- Show that the rigid-body equations (4.7) may be written as
$$\dot{x}_1 = -\,a_1a_3x_2x_3\,,\ \ \dot{x}_2 = -\,a_2a_3x_3x_1\,,\ \ \dot{x}_3 = a_1a_2x_1x_2\,, \tag{4.16}$$
with nonzero constants a_1, a_2 and a_3 that satisfy $1/a_1 + 1/a_2 = 1/a_3$. Write these equations as a Nambu motion equation on $\mathbb{R}^3$ of the form (4.14). Interpret the solutions of Equations (4.16) geometrically as intersections of orthogonal cylinders (elliptic or hyperbolic) for various values and signs of a_1, a_2 and a_3, as in Fig. 4.4.

★

Answer. $\dot{\mathbf{x}} := (\dot{x}_1, \dot{x}_2, \dot{x}_3)^T = \frac{1}{4}\nabla(a_1x_1^2 + a_3x_3^2) \times \nabla(a_2x_2^2 + a_3x_3^2)$, where (a_1, a_2, a_3) may be written in terms of (I_1, I_2, I_3), when they satisfy $1/a_1 + 1/a_2 = 1/a_3$. ▲

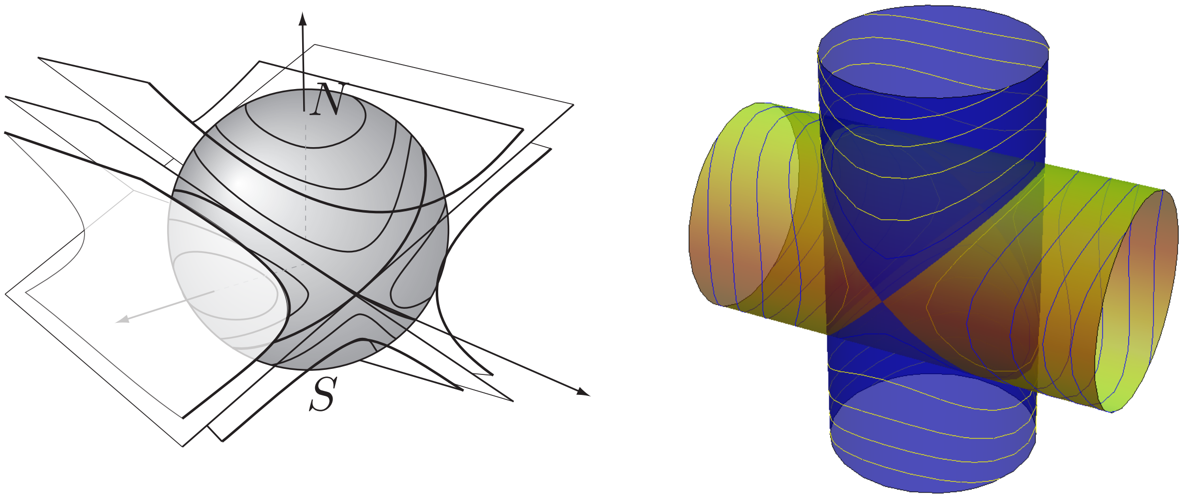


FIGURE 2. Left: Rigid body motions, seen as intersections in $\mathbb{R}^3$ of the sphere $x_1^2 + x_2^2 + x_3^2 = \text{constant}$ and the hyperbolic cylinders $x_1^2 - x_3^2 = \text{constant}$. Right: The same rigid body motions, seen as intersections in $\mathbb{R}^3$ of orthogonal elliptic cylinders.

4.5. Clebsch variational principle for the rigid body.

Proposition

4.7 (Clebsch variational principle).
The Euler rigid-body Equations (3.2) on $T\mathbb{R}^3$ are equivalent to the ***constrained variational principle****,*
$$\delta S(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P}) = \delta \int_a^b l(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P})\, dt = 0, \tag{4.17}$$

for a ***constrained action integral***

$$S(\boldsymbol{\Omega},\mathbf{Q},\dot{\mathbf{Q}}) = \int_a^b l(\boldsymbol{\Omega},\mathbf{Q},\dot{\mathbf{Q}})\,dt = \int_a^b \frac{1}{2}\boldsymbol{\Omega}\cdot\mathbb{I}\boldsymbol{\Omega} + \mathbf{P}\cdot\Big(\dot{\mathbf{Q}} + \boldsymbol{\Omega}\times\mathbf{Q}\Big)\,dt\,. \tag{4.18}$$

Remark

4.8 (Reconstruction as constraint)**.**

- *The first term in the Lagrangian (4.18),*

$$l(\boldsymbol{\Omega}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2) = \frac{1}{2}\boldsymbol{\Omega}^T\mathbb{I}\boldsymbol{\Omega}\,, \tag{4.19}$$

  is again the (rotational) kinetic energy of the rigid body.
- *The second term in the Lagrangian (4.18) introduces the Lagrange multiplier* $\mathbf{P}$ *which imposes the constraint*

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

  This ***reconstruction formula*** *has the solution*

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

  which satisfies

$$\begin{aligned}\dot{\mathbf{Q}}(t) &= -(O^{-1}\dot{O})O^{-1}(t)\mathbf{Q}(0)\\ &= -\hat{\Omega}(t)\mathbf{Q}(t) = -\boldsymbol{\Omega}(t)\times\mathbf{Q}(t)\,.\end{aligned} \tag{4.20}$$

Proof. The variations of S are given by

$$\begin{aligned}\delta S &= \int_a^b \Big(\frac{\delta l}{\delta\boldsymbol{\Omega}}\cdot\delta\boldsymbol{\Omega} + \frac{\delta l}{\delta\mathbf{P}}\cdot\delta\mathbf{P} + \frac{\delta l}{\delta\mathbf{Q}}\cdot\delta\mathbf{Q}\Big)\,dt\\ &= \int_a^b \Big[\Big(\frac{\delta l}{\delta\boldsymbol{\Omega}} - \mathbf{P}\times\mathbf{Q}\Big)\cdot\delta\boldsymbol{\Omega}\\ &\qquad + \delta\mathbf{P}\cdot\Big(\dot{\mathbf{Q}} + \boldsymbol{\Omega}\times\mathbf{Q}\Big) - \delta\mathbf{Q}\cdot\Big(\dot{\mathbf{P}} + \boldsymbol{\Omega}\times\mathbf{P}\Big)\Big]\,dt\,.\end{aligned}$$

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

$$\boldsymbol{\Pi} := \frac{\delta l}{\delta\boldsymbol{\Omega}} = \mathbf{P}\times\mathbf{Q}\,,\quad \dot{\mathbf{Q}} = -\,\boldsymbol{\Omega}\times\mathbf{Q}\,,\quad \dot{\mathbf{P}} = -\,\boldsymbol{\Omega}\times\mathbf{P}\,. \tag{4.21}$$

Euler's form of the rigid-body equations emerges from these ***symmetric equations***, upon elimination of $\mathbf{Q}$ and $\mathbf{P}$, as

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

which are Euler's equations for the rigid body in $T\mathbb{R}^3$ when $\boldsymbol{\Pi} = \mathbb{I}\boldsymbol{\Omega}$. □

Remark

4.9. *The Clebsch variational principle for the rigid body is a natural approach in developing geometric algorithms for numerical integrations of rotating motion.*

Remark

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

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

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

what are the Poisson brackets for $\mathbf{\Pi} = \mathbf{P} \times \mathbf{Q} \in \mathbb{R}^3$ in (4.21)? Show these Poisson brackets recover the rigid-body Poisson bracket (4.6). ★

Answer. The components of the angular momentum $\mathbf{\Pi} = \mathbb{I}\mathbf{\Omega}$ in (4.21) are

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

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

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

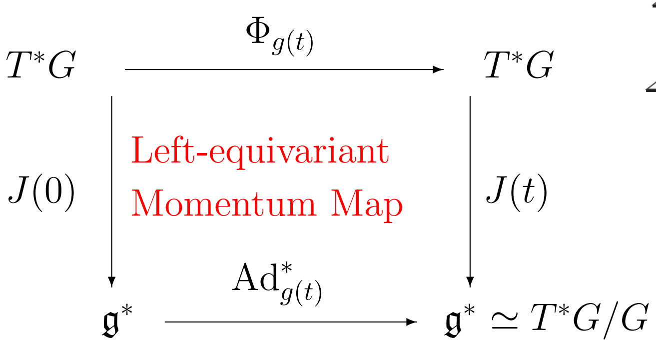
Consequently, the derivative property of the canonical Poisson bracket yields

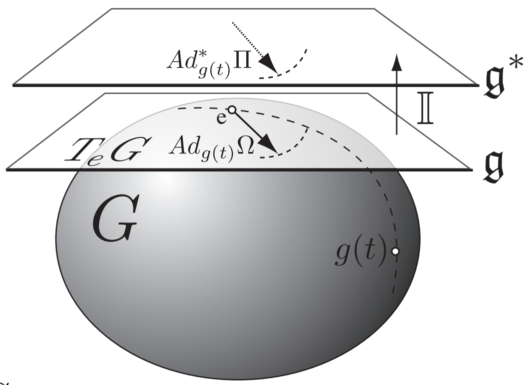
$$\{f, h\}(\mathbf{\Pi}) = \frac{\partial f}{\partial \Pi_a}\{\Pi_a, \Pi_i\}\frac{\partial h}{\partial \Pi_b} = -\,\epsilon_{abc}\Pi_c \frac{\partial f}{\partial \Pi_a}\frac{\partial h}{\partial \Pi_b}\,, \tag{4.22}$$

which is indeed the Lie–Poisson bracket in (4.6) on functions of the $\mathbf{\Pi}$'s. The correspondence with the hat map noted above shows that this Poisson bracket satisfies the Jacobi identity as a result of the Jacobi identity for the vector cross product on $\mathbb{R}^3$. ▲

Remark

4.11. *This exercise proves that the map $T^*\mathbb{R}^3 \to \mathbb{R}^3$ given by $\mathbf{\Pi} = \mathbf{P} \times \mathbf{Q} \in \mathbb{R}^3$ in (4.21) is Poisson. That is, the map takes Poisson brackets on one manifold into Poisson brackets on another manifold. This is one of the properties of a* ***momentum map****.*





Recall the set-up for equivariant momentum maps.

Definition

4.12 (Cotangent lift (CL) momentum map)**.** *The CL momentum map*

$$J : T^*M \mapsto \mathfrak{g}^*$$

is defined for the Lie algebra action $\xi_M(q)$ of $\xi \in \mathfrak{g}$ on q in manifold M by the pairings

$$\begin{aligned} J^\xi(p,q) \;&:=\; \Big\langle J(p,q),\, \xi \Big\rangle_{\mathfrak{g}^*\times\mathfrak{g}} \\ &=\; \Big\langle\!\Big\langle p_q,\, \xi_M(q) \Big\rangle\!\Big\rangle_{T^*M\times TM} \end{aligned}$$

where $p_q \in T^*_qM$ is the momentum at position $q \in M$ and $\xi_M(q)$ is the vector field tangent to the flow of $g(t) \in G$ at q.

Proposition

4.13. $J^\xi(p,q)$ *is the* Hamiltonian *for infinitesimal action* $\xi_M(q)$ *and its cotangent lift.*

Proof.

$$\dot q = \{q, J^\xi\} = \xi_M(q) \quad \text{and} \quad \dot p = \{p, J^\xi\} = -\,\frac{d\xi_M}{dq}^T \cdot p$$

□

Example

4.14 (Body angular momentum, $G = SO(3)$ and $M = \mathbb{R}^3$).

$$J(q,p) = p \times q, \quad \big(\textit{Body angular momentum } J \in \mathfrak{so}(3)^* \simeq \mathbb{R}^3\big)\,.$$

The Hamiltonian $J^\xi(q,p) = p\times q\cdot\xi$ *generates the infinitesimal* $SO(3)$ *rotations,*

$$q'(t) = \{q, J^\xi(q,p)\} = -\,\xi\times q(t), \quad p'(t) = \{p, J^\xi(q,p)\} = -\,\xi\times p(t),$$

for the canonical Poisson bracket $\{\,\cdot\,,\,\cdot\,\}$*. These imply the Euler-Poincaré (EP) equation for* $J(q,p) = p\times q \in \mathfrak{so}(3)^* \simeq \mathbb{R}^3$

$$J'(t) = -\,\xi\times J(t) = \mathrm{ad}^*_\xi J \quad \textit{for } \xi\in\mathfrak{so}(3) \textit{ and } J\in\mathfrak{so}(3)^*\,.$$

Proof.

$$\begin{aligned} J'(t) &= p'(t)\times q + p\times q'(t)\\ &= -(\xi\times p)\times q - p\times(\xi\times q)\\ &= -q\times(p\times\xi) - p\times(\xi\times q)\\ \text{(By Jacobi identity)} &= \xi\times(q\times p)\\ &= -\,\xi\times J\\ &= \mathrm{ad}^*_\xi J \end{aligned}$$

□

This calculation also illustrates the following theorem.

Theorem

4.15. *The CL momentum map* $J(p,q)$ *is infinitesimally equivariant.*

That is, the CL momentum map $J(p,q)$ satisfies the EP equation, when (p,q) satisfy the canonical equations for the Hamiltonian $J^\xi(p,q) = \langle p_q, \Phi_M(q)\rangle$. Consequently, (p,q) satisfy the equations of motion for the canonical transformation $\Phi_{g(t)}$ of T^*M and the momentum map satisfies $J'(t) = \mathrm{ad}^*_\xi J$, which is the infinitesimal (linearised) version of $J(t) = \mathrm{Ad}^*_{g(t)}J(0)$. To remind ourselves of the latter fact, we recall equation (3.33) in the present notation, as

$$\frac{d}{dt}\Big(J(0)\Big) = \frac{d}{dt}\Big(\mathrm{Ad}^*_{g^{-1}(t)}J\Big) = \mathrm{Ad}^*_{g^{-1}(t)}\Big(\frac{d}{dt}J - \mathrm{ad}^*_\xi J\Big) = 0.$$

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

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

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

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

for which M is the cotangent lift momentum map

$$M = \frac{\partial l}{\partial \Omega} = \frac{1}{2}(P^T Q - Q^T P)$$

and $Q, P \in SO(n)$ satisfy the following symmetric equations reminiscent of those in (4.21),

$$\dot{Q} = Q\Omega \quad \text{and} \quad \dot{P} = P\Omega\,, \tag{4.23}$$

as a result of the constraints.

Show that M satisfies the Euler-Poincaré equation

$$\frac{dM}{dt} = \operatorname{ad}^*_\Omega M = -\left[\,\Omega,\, M\,\right],$$

as it should, since it is a cotangent lift momentum map and those are equivariant.

★

4.6. **Rotating motion with an added quadratic potential energy.** Manakov's method for showing the integrability of the n-dimensional rigid body illustrates the conditions necessary to prove isospectral integrability for any Lie–Poisson system. For example, consider the problem of a rigid body in a quadratic potential, first studied in [Bo1985].

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

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

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

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

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

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

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

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

Integrating by parts, invoking homogeneous endpoint conditions, then rearranging as in the proof of Proposition 3.4 and using the variational relation (3.18), rewritten here as

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

finally yields the following formula for the variation,

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

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

$$\frac{dM}{dt} = [\,M\,,\,\Omega\,] + [\,\mathbb{S}(t)\,,\,\mathbb{A}\,]\,. \tag{4.24}$$

A differential equation for $\mathbb{S}(t)$ follows from the time derivative of its definition $\mathbb{S}(t) := O^{-1}(t)\mathbb{S}_0 O(t)$, as

$$\frac{d\mathbb{S}}{dt} = \big[\,\mathbb{S}\,,\,\Omega\,\big]\,. \tag{4.25}$$

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

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

$$\frac{d}{dt}\big(\mathbb{S} + \lambda M + \lambda^2 \mathbb{A}^2\big) = \big[\,\mathbb{S} + \lambda M + \lambda^2\mathbb{A}^2\,,\,\Omega + \lambda\mathbb{A}\,\big]\,. \tag{4.26}$$

The commutator form (4.26) implies for every non-negative integer power K that

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

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

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

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

$$\mathrm{tr}(\mathbb{S} + \lambda M + \lambda^2\mathbb{A}^2)^K = \text{constant}\,, \tag{4.27}$$

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

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

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

Determine the Lie algebra involved. ★

Exercise. Explicitly compute the conservation laws in (4.27) for $n = 4$. ★

Exercise. What is the dimension of the generic solution of the system of equations (4.24) and (4.25)? That is, what is the sum of the dimensions of $so(n)$ and the symmetric $n \times n$ matrices, minus the number of conservation laws? ★

Exercise. Write the equations of motion and their Lie–Poisson Hamiltonian formulation in $\mathbb{R}^3$-vector form for the case when

$$\Omega(t) = O^{-1}(t)\dot{O}(t) \in \mathfrak{so}(3)$$

by using the hat map. List the conservation laws in this case. ★

Exercise. How would the variational calculation of the system (4.24) and (4.25) have changed if the Lie group had been unitary instead of orthogonal and the matrices $\mathbb{S}_0$, $\mathbb{A}$ and $\mathbb{S}(t)$ were Hermitian, rather than symmetric? ★

5. Variations on rigid-body dynamics

5.1. Rotations in the language of quaternions.

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

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

Hamilton was correct: quaternions are special. For example, they form the only associative division ring containing both real and complex numbers. For us, they also form a natural introduction to geometric mechanics. In particular, quaternions will introduce us to ***mechanics on Lie groups***; namely, mechanics on the Lie group $SU(2)$ of 2×2 special unitary matrices.

5.1.1. *Multiplying quaternions using Pauli matrices.*

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

$$\mathbb{J}_1\mathbb{J}_2\mathbb{J}_3 = -\mathbb{J}_0\,, \tag{5.1}$$

and the squares

$$\mathbb{J}_1^2 = \mathbb{J}_2^2 = \mathbb{J}_3^2 = -\mathbb{J}_0\,, \tag{5.2}$$

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

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

$$\mathfrak{q} = [q_0,\ \boldsymbol{q}] = q_0\mathbb{J}_0 + q_1\mathbb{J}_1 + q_2\mathbb{J}_2 + q_3\mathbb{J}_3 \in \mathbb{H}\,. \tag{5.3}$$

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

(5.4)

	$\mathbb{J}_0$	$\mathbb{J}_1$	$\mathbb{J}_2$	$\mathbb{J}_3$
$\mathbb{J}_0$	$\mathbb{J}_0$	$\mathbb{J}_1$	$\mathbb{J}_2$	$\mathbb{J}_3$
$\mathbb{J}_1$	$\mathbb{J}_1$	$-\mathbb{J}_0$	$\mathbb{J}_3$	$-\mathbb{J}_2$
$\mathbb{J}_2$	$\mathbb{J}_2$	$-\mathbb{J}_3$	$-\mathbb{J}_0$	$\mathbb{J}_1$
$\mathbb{J}_3$	$\mathbb{J}_3$	$\mathbb{J}_2$	$-\mathbb{J}_1$	$-\mathbb{J}_0$

.

Definition

5.1 (Multiplication of quaternions)**.** *The* ***multiplication rule*** *for two quaternions,*

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

may be defined in vector notation as

$$\mathfrak{q}\mathfrak{r} = [q_0,\ \mathbf{q}][r_0,\ \mathbf{r}] = [\,q_0r_0 - \mathbf{q}\cdot\mathbf{r}\,,\ q_0\mathbf{r} + r_0\mathbf{q} + \mathbf{q}\times\mathbf{r}\,]\ . \tag{5.5}$$

Remark

5.2. *The antisymmetric and symmetric parts of the quaternionic product correspond to* ***vector operations***[2]*:*

$$\frac{1}{2}\Big(\mathfrak{q}\mathfrak{r} - \mathfrak{r}\mathfrak{q}\Big) = [0\,,\,\mathbf{q}\times\mathbf{r}\,]\,, \tag{5.6}$$

$$\frac{1}{2}\Big(\mathfrak{q}\mathfrak{r} + \mathfrak{r}\mathfrak{q}\Big) = [q_0r_0 - \mathbf{q}\cdot\mathbf{r}\,,\,q_0\mathbf{r} + r_0\mathbf{q}\,]\,. \tag{5.7}$$

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

Theorem

5.3 (Isomorphism with Pauli matrix product)**.**
The multiplication rule (5.5) may be represented in a 2×2 matrix basis as

$$\mathfrak{q} = [q_0,\,\mathbf{q}] = q_0\sigma_0 - i\mathbf{q}\cdot\boldsymbol{\sigma}\,, \quad \textit{with}\quad \mathbf{q}\cdot\boldsymbol{\sigma} := \sum_{a=1}^{3} q_a\sigma_a\,, \tag{5.8}$$

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

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5.9}$$

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

$$\sigma_a\sigma_b = \delta_{ab}\,\sigma_0 + i\epsilon_{abc}\sigma_c \quad \text{for} \quad a,b,c = 1,2,3, \tag{5.10}$$

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 the basic quaternionic properties. $\square$

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

$$[\,\sigma_a\,,\,\sigma_b\,] := \sigma_a\sigma_b - \sigma_b\sigma_a = 2i\epsilon_{abc}\sigma_c \quad \text{for} \quad a,b,c = 1,2,3, \tag{5.11}$$

and their ***anticommutator relation***

$$\{\,\sigma_a\,,\,\sigma_b\,\}_+ := \sigma_a\sigma_b + \sigma_b\sigma_a = 2\delta_{ab}\sigma_0 \quad \text{for} \quad a,b = 1,2,3. \tag{5.12}$$

The corresponding relations among quaternions are given in (5.6) and (5.7), respectively. ★

Exercise. Verify the quaternionic multiplication rule expressed in the tetrad-bracket notation in (5.5) by using the isomorphism (5.8) and the product relation for the Pauli matrices in Equation (5.10). ★

Answer.

$$\begin{aligned} \mathfrak{q}\mathfrak{r} &= (q_0\sigma_0 - iq_a\sigma_a)(r_0\sigma_0 - ir_b\sigma_b) \\ &= (q_0r_0 - \boldsymbol{q}\cdot\boldsymbol{r})\,\sigma_0 - i(q_0\boldsymbol{r} + r_0\boldsymbol{q} + \boldsymbol{q}\times\boldsymbol{r})\cdot\boldsymbol{\sigma}\,. \end{aligned}$$

▲

[2] Hamilton introduced the word ***vector*** in 1846 as a synonym for a ***pure quaternion***, whose scalar part vanishes.

Exercise. Use Equations (5.11), (5.12) and isomorphism (5.8) to verify relations (5.6) and (5.7). ★

Exercise. Use formula (5.10) to verify the decomposition of a vector in Pauli matrices

$$\mathbf{q}\,\sigma_0 = (\mathbf{q}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma} - i\mathbf{q}\times\boldsymbol{\sigma}\,, \tag{5.13}$$

which is valid for three-vectors $\mathbf{q}\in\mathbb{R}^3$. Verify also that

$$-\,|\mathbf{q}\times\boldsymbol{\sigma}|^2 = 2|\mathbf{q}|^2\sigma_0 = 2(\mathbf{q}\cdot\boldsymbol{\sigma})^2.$$

★

Exercise. Use Equations (5.11) to verify the commutation relation

$$[\,\mathbf{p}\cdot\boldsymbol{\sigma},\,\mathbf{q}\cdot\boldsymbol{\sigma}\,] = 2i\mathbf{p}\times\mathbf{q}\cdot\boldsymbol{\sigma}$$

for three-vectors $\mathbf{p},\mathbf{q}\in\mathbb{R}^3$. ★

5.1.2. *Quaternionic conjugate.*

Remark

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

$$\mathfrak{p}(\mathfrak{q}\mathfrak{r}) = (\mathfrak{p}\mathfrak{q})\mathfrak{r}\,. \tag{5.14}$$

However, the quaternionic product is not commutative,

$$[\mathfrak{p},\,\mathfrak{q}] := \mathfrak{p}\mathfrak{q} - \mathfrak{q}\mathfrak{p} = [0,\,2\mathbf{p}\times\mathbf{q}\,]\,, \tag{5.15}$$

as we saw earlier in (5.6).

Definition

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

$$\mathfrak{q}^* = [q_0\,,\,-\mathbf{q}]\ . \tag{5.16}$$

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

$$\mathrm{Re}\,\mathfrak{q} \;:=\; \frac{1}{2}(\mathfrak{q}+\mathfrak{q}^*) = [\,q_0,\,0\,]\,, \tag{5.17}$$

$$\mathrm{Im}\,\mathfrak{q} \;:=\; \frac{1}{2}(\mathfrak{q}-\mathfrak{q}^*) = [\,0\,,\,\mathbf{q}\,]\,. \tag{5.18}$$

Lemma

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

$$(\mathfrak{p}\mathfrak{q})^* \;=\; \mathfrak{q}^*\mathfrak{p}^* \quad \textit{(note reversed order)}, \tag{5.19}$$

$$\begin{aligned}\mathrm{Re}(\mathfrak{p}\mathfrak{q}^*) \;&:=\; \frac{1}{2}(\mathfrak{p}\mathfrak{q}^*+\mathfrak{q}\mathfrak{p}^*)\\ &=\; [p_0q_0+\mathbf{p}\cdot\mathbf{q}\,,\,0\,] \quad \textit{(yields real part)}.\end{aligned} \tag{5.20}$$

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

Definition

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

$$\begin{aligned} \mathfrak{p}\cdot\mathfrak{q} &= [p_0\,,\,\mathbf{p}\,]\cdot[q_0\,,\,\mathbf{q}\,] \\ &:= [p_0q_0+\mathbf{p}\cdot\mathbf{q}\,,\,0\,] = \mathrm{Re}(\mathfrak{p}\mathfrak{q}^*)\,. \end{aligned} \tag{5.21}$$

Definition

5.8 (Pairing of quaternions)**.** *The quaternionic dot product (5.21) defines a real symmetric pairing* $\langle\,\cdot\,,\,\cdot\,\rangle : \mathbb{H}\times\mathbb{H}\mapsto\mathbb{R}$*, denoted as*

$$\langle\,\mathfrak{p}\,,\,\mathfrak{q}\,\rangle = \mathrm{Re}(\mathfrak{p}\mathfrak{q}^*) := \mathrm{Re}(\mathfrak{q}\mathfrak{p}^*) = \langle\,\mathfrak{q}\,,\,\mathfrak{p}\,\rangle\,. \tag{5.22}$$

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

Definition

5.9 (Magnitude of a quaternion)**.** *The* ***magnitude*** *of a quaternion* $\mathfrak{q}$ *may be defined by*

$$|\mathfrak{q}| := (\mathfrak{q}\cdot\mathfrak{q})^{1/2} = ({q_0}^2+\mathbf{q}\cdot\mathbf{q})^{1/2}\,. \tag{5.23}$$

Remark

5.10. *A level set of* $|\mathfrak{q}|$ *defines a three-sphere* S^3*.*

Definition

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

$$|\mathfrak{q}|^2 := \mathfrak{q}\mathfrak{q}^* = (\mathfrak{q}\cdot\mathfrak{q})\mathfrak{e}\,, \tag{5.24}$$

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

$$\mathfrak{q}^{-1} := \mathfrak{q}^*/|\mathfrak{q}|^2 \tag{5.25}$$

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

Exercise. Does a quaternion $\mathfrak{q}$ have a square root? Prove it. ★

Exercise. Show that the magnitude of the product of two quaternions is the product of their magnitudes. ★

Answer. From the definitions of the quaternionic multiplication rule (5.5), inner product (5.21) and magnitude (5.23), one verifies that

$$\begin{aligned} |\mathfrak{p}\mathfrak{q}|^2 &= (p_0q_0-\mathbf{p}\cdot\mathbf{q})^2+|p_0\mathbf{q}+\mathbf{p}q_0+\mathbf{p}\times\mathbf{q}|^2 \\ &= ({p_0}^2+|\mathbf{p}|^2)({q_0}^2+|\mathbf{q}|^2) = |\mathfrak{p}|^2|\mathfrak{q}|^2\,. \end{aligned}$$

▲

Definition

5.12. *A quaternion* $\mathfrak{q}$ *with magnitude* $|\mathfrak{q}|=1$ *is called a* ***unit quaternion****.*

Definition

5.13. *A quaternion with no scalar (or real) part* $\mathfrak{q} = [0, \mathbf{q}]$ *is called a* ***pure quaternion****, or equivalently a* ***vector*** *(a term introduced by Hamilton in 1846 [Ne1997]).*

Exercise. Show that the antisymmetric and symmetric parts of the product of two pure quaternions $\mathfrak{v} = [0\,,\,\boldsymbol{v}]$ and $\mathfrak{w} = [0\,,\,\boldsymbol{w}]$ yield, respectively, the cross product and (minus) the scalar product of the two corresponding vectors $\boldsymbol{v}$, $\boldsymbol{w}$. ★

Answer. The quaternionic product of pure quaternions $\mathfrak{v} = [0\,,\,\boldsymbol{v}]$ and $\mathfrak{w} = [0\,,\,\boldsymbol{w}]$ is defined as

$$\mathfrak{v}\mathfrak{w} = \big[\, -\boldsymbol{v}\cdot\boldsymbol{w},\ \boldsymbol{v}\times\boldsymbol{w} \,\big]\,.$$

Its antisymmetric (vector) part yields the cross product of the corresponding vectors:

$$\mathrm{Im}(\mathfrak{v}\mathfrak{w}) = \frac{1}{2}\Big(\mathfrak{v}\mathfrak{w} - \mathfrak{w}\mathfrak{v}\Big) = [0\,,\ \boldsymbol{v}\times\boldsymbol{w}] \qquad (\text{vanishes for } \boldsymbol{v}\|\boldsymbol{w})\,.$$

Its symmetric (or real) part yields minus the scalar product of the vectors:

$$\mathrm{Re}(\mathfrak{v}\mathfrak{w}) = \frac{1}{2}\Big(\mathfrak{v}\mathfrak{w} + \mathfrak{w}\mathfrak{v}\Big) = [-\,\boldsymbol{v}\cdot\boldsymbol{w}\,,\ 0] \qquad (\text{vanishes for } \boldsymbol{v}\perp\boldsymbol{w})\,.$$

▲

Remark

5.14 ($\mathbb{H}_0 \simeq \mathbb{R}^3$)**.** *Being three-dimensional linear spaces possessing the same vector and scalar products, pure quaternions in* $\mathbb{H}_0$ *(with no real part) are equivalent to vectors in* $\mathbb{R}^3$*.*

5.1.3. *Decomposition of three-vectors.* Pure quaternions have been identified with vectors in $\mathbb{R}^3$. Under this identification, the two types of products of pure quaternions $[0,\,\mathbf{v}]$ and $[0,\,\mathbf{w}]$ are given by

$$[0,\,\mathbf{v}]\cdot[0,\,\mathbf{w}] = [\mathbf{v}\cdot\mathbf{w},\,0] \quad \text{and} \quad [0,\,\mathbf{v}][0,\,\mathbf{w}] = [-\mathbf{v}\cdot\mathbf{w},\,\mathbf{v}\times\mathbf{w}]\,.$$

Thus, the dot $(\,\cdot\,)$ and cross $(\,\times\,)$ products of three-vectors may be identified with these two products of pure quaternions. The product of an arbitrary quaternion $[\alpha,\,\boldsymbol{\chi}]$ with a pure unit quaternion $[0,\,\hat{\boldsymbol{\omega}}]$ produces another pure quaternion, provided $\boldsymbol{\chi}\cdot\hat{\boldsymbol{\omega}} = 0$. In this case, one computes

$$[\alpha,\,\boldsymbol{\chi}][0,\,\hat{\boldsymbol{\omega}}] = [-\boldsymbol{\chi}\cdot\hat{\boldsymbol{\omega}},\,\alpha\,\hat{\boldsymbol{\omega}} + \boldsymbol{\chi}\times\hat{\boldsymbol{\omega}}] =: [0,\,\mathbf{v}],\ \text{for } \boldsymbol{\chi}\cdot\hat{\boldsymbol{\omega}} = 0\,. \tag{5.26}$$

Remark

5.15. *Quaternions are summoned whenever a three-vector* $\mathbf{v}$ *is decomposed into its components parallel* $(\|)$ *and perpendicular* $(\perp)$ *to a unit three-vector direction* $\hat{\boldsymbol{\omega}}$*, according to*

$$\mathbf{v} = \alpha\,\hat{\boldsymbol{\omega}} + \boldsymbol{\chi}\times\hat{\boldsymbol{\omega}} = [\alpha,\,\boldsymbol{\chi}][0,\,\hat{\boldsymbol{\omega}}] = \mathbf{v}_\| + \mathbf{v}_\perp\,. \tag{5.27}$$

Here $\alpha = \hat{\boldsymbol{\omega}}\cdot\mathbf{v}$ *and* $\boldsymbol{\chi} = \hat{\boldsymbol{\omega}}\times\mathbf{v}$ *so that* $\boldsymbol{\chi}\cdot\hat{\boldsymbol{\omega}} = 0$ *and one uses* $\hat{\boldsymbol{\omega}}\cdot\hat{\boldsymbol{\omega}} = 1$ *to find* $\mathbf{v}\cdot\mathbf{v} = \alpha^2 + \chi^2$ *with* $\chi := |\boldsymbol{\chi}|$*. The vector decomposition (5.27) is precisely the quaternionic product (5.26), in which the vectors* $\mathbf{v}$ *and* $\hat{\boldsymbol{\omega}}$ *are treated as pure quaternions.*

This remark may be summarised by the following.

Proposition

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

5.1.4. *Quaternionic conjugation: Cayley–Klein parameters.*

Definition

5.17 (Quaternionic conjugation).
Quaternionic conjugation *is defined as the map under the quaternionic product (recalling that this product is associative),*

$$\mathfrak{r} \to \mathfrak{r}' = \hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^*\,, \tag{5.28}$$

where $\hat{\mathfrak{q}} = [q_0\,,\,\mathbf{q}]$ *is a unit quaternion,* $\hat{\mathfrak{q}}\cdot\hat{\mathfrak{q}} = 1$*, so* $\hat{\mathfrak{q}}\hat{\mathfrak{q}}^* = \mathfrak{e} = [1,0]$*. The inverse map is*

$$\mathfrak{r} = \hat{\mathfrak{q}}^*\mathfrak{r}'\hat{\mathfrak{q}}\,.$$

Exercise. Show that the product of a quaternion $\mathfrak{r} = [r_0, \boldsymbol{r}]$ with a unit quaternion $\hat{\mathfrak{q}} = [q_0, \boldsymbol{q}]$, whose inverse is $\hat{\mathfrak{q}}^* = [q_0, -\,\boldsymbol{q}]$, satisfies

$$\begin{aligned}
\mathfrak{r}\hat{\mathfrak{q}}^* &= \big[\mathfrak{r}\cdot\hat{\mathfrak{q}}, -\,r_0\boldsymbol{q} + q_0\boldsymbol{r} + \boldsymbol{q}\times\boldsymbol{r}\big]\,,\\
\hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^* &= \big[r_0|\hat{\mathfrak{q}}|^2, \boldsymbol{r} + 2q_0\boldsymbol{q}\times\boldsymbol{r} + 2\boldsymbol{q}\times(\boldsymbol{q}\times\boldsymbol{r})\big]\,,
\end{aligned}$$

where $\mathfrak{r}\cdot\hat{\mathfrak{q}} = r_0q_0 + \boldsymbol{r}\cdot\boldsymbol{q}$ and $|\hat{\mathfrak{q}}|^2 = \hat{\mathfrak{q}}\cdot\hat{\mathfrak{q}} = {q_0}^2 + \boldsymbol{q}\cdot\boldsymbol{q} = 1$ according to the definitions of the dot product in (5.21) and magnitude in (5.23). ★

Remark

5.18. *The same products using the pure unit quaternion* $\hat{\mathfrak{z}} = [0, \hat{\mathbf{z}}]$ *with* $\hat{\mathbf{z}} = (0,0,1)^T$ *and the unit quaternion* $\hat{\mathfrak{q}} = [q_0, \mathbf{q}]$ *satisfy*

$$\begin{aligned}
\hat{\mathfrak{z}}\hat{\mathfrak{q}}^* &= \big[q_3, q_0\hat{\mathbf{z}} + \mathbf{q}\times\hat{\mathbf{z}}\big]\,,\\
\hat{\mathfrak{q}}\,\hat{\mathfrak{z}}\,\hat{\mathfrak{q}}^* &= \big[0, \hat{\mathbf{z}} + 2q_0\mathbf{q}\times\hat{\mathbf{z}} + 2\mathbf{q}\times(\mathbf{q}\times\hat{\mathbf{z}})\big]\,,
\end{aligned}$$

which produces a complete set of unit vectors.

Remark

5.19. *Conjugation* $\hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^*$ *is a wise choice, as opposed to, say, choosing the apparently less meaningful triple product*

$$\hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}} \;=\; [0,\,\mathbf{r}] + (r_0q_0 - \mathbf{r}\cdot\mathbf{q})[q_0,\,\mathbf{q}]$$

for quaternions $\mathfrak{r} = [r_0, \mathbf{r}]$ *and* $\hat{\mathfrak{q}} = [q_0, \mathbf{q}]$ *with* $|\mathfrak{q}|^2 = {q_0}^2 + \mathbf{q}\cdot\mathbf{q} = 1$*.*

Exercise. For $\mathfrak{q}^* = [q_0, -\,\boldsymbol{q}]$, such that $\mathfrak{q}^*\mathfrak{q} = \mathbb{J}_0|\mathfrak{q}|^2$, verify that

$$2\mathfrak{q}^* = -\mathbb{J}_0\mathfrak{q}\mathbb{J}_0^* + \mathbb{J}_1\mathfrak{q}\mathbb{J}_1^* + \mathbb{J}_2\mathfrak{q}\mathbb{J}_2^* + \mathbb{J}_3\mathfrak{q}\mathbb{J}_3^*\,.$$

What does this identity mean geometrically? Does the complex conjugate z^* for $z\in\mathbb{C}$ satisfy such an identity? Prove it. ★

Lemma

5.20. *As a consequence of Remark 5.18 and the Exercise just before it, one finds that conjugation* $\hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^*$ *of a quaternion* $\mathfrak{r}$ *by a unit quaternion* $\hat{\mathfrak{q}}$ *preserves the sphere* $S^3_{|\mathfrak{r}|}$ *given by any level set of* $|\mathfrak{r}|$*. That is, the value of* $|\mathfrak{r}|^2$ *is invariant under conjugation by a unit quaternion:*

$$|\hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^*|^2 = |\mathfrak{r}|^2 = {r_0}^2 + \mathbf{r}\cdot\mathbf{r}\,. \tag{5.29}$$

Definition

5.21 (Conjugacy classes). *The set*

$$C(\mathfrak{r}) := \left\{ \mathfrak{r}' \in \mathbb{H} \,\middle|\, \mathfrak{r}' = \hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^* \right\} \tag{5.30}$$

is called the ***conjugacy class*** *of the quaternion* $\mathfrak{r}$*.*

Corollary

5.22. *The conjugacy classes of the three-sphere* $S^3_{|\mathfrak{r}|}$ *under conjugation by a unit quaternion* $\hat{\mathfrak{q}}$ *are the two-spheres given by*

$$\left\{ \mathbf{r} \in \mathbb{R}^3 \,\middle|\, \mathbf{r}\cdot\mathbf{r} = |\mathfrak{r}|^2 - {r_0}^2 \right\}. \tag{5.31}$$

Proof. The proof is a straightforward exercise. □

Remark

5.23. *The expressions in Remark 5.18 correspond to* ***spatial rotations*** *when* $r_0 = 0$ *so that* $\mathfrak{r} = [0,\, \mathbf{r}\,]$.

Lemma

5.24 (Euler–Rodrigues formula). *If* $\mathfrak{r} = [0, \mathbf{r}]$ *is a pure quaternion and* $\hat{\mathfrak{q}} = [q_0\,,\, \mathbf{q}]$ *is a unit quaternion, then under quaternionic conjugation,*

$$\begin{aligned} \mathfrak{r}' &= \hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^* = \big[0, \mathbf{r}'\big] \\ &= \big[0, \mathbf{r} + 2q_0(\mathbf{q}\times\mathbf{r}) + 2\mathbf{q}\times(\mathbf{q}\times\mathbf{r})\big]\,. \end{aligned} \tag{5.32}$$

For $\hat{\mathfrak{q}} := \pm[\cos\frac{\theta}{2}\,,\, \sin\frac{\theta}{2}\,\hat{\mathbf{n}}]$*, we have*

$$\big[0, \mathbf{r}'\big] = \left[\cos\frac{\theta}{2}\,,\, \sin\frac{\theta}{2}\,\hat{\mathbf{n}}\right]\big[0, \mathbf{r}\big]\left[\cos\frac{\theta}{2}\,, -\sin\frac{\theta}{2}\,\hat{\mathbf{n}}\right],$$

so that

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} + 2\cos\frac{\theta}{2}\sin\frac{\theta}{2}\,(\hat{\mathbf{n}}\times\mathbf{r}) + 2\sin^2\frac{\theta}{2}\big(\hat{\mathbf{n}}\times(\hat{\mathbf{n}}\times\mathbf{r})\big) \\ &= \mathbf{r} + \sin\theta\,(\hat{\mathbf{n}}\times\mathbf{r}) + (1-\cos\theta)\big(\hat{\mathbf{n}}\times(\hat{\mathbf{n}}\times\mathbf{r})\big) \\ &=: O^\theta_{\hat{\mathbf{n}}}\,\mathbf{r}\,. \end{aligned} \tag{5.33}$$

This is the famous ***Euler–Rodrigues formula*** *for the rotation* $O^\theta_{\hat{\mathbf{n}}}\,\mathbf{r}$ *of a vector* $\mathbf{r}$ *by an angle* θ *about the unit vector* $\hat{\mathbf{n}}$*.*

Exercise. Verify the Euler–Rodrigues formula (5.33) by a direct computation using quaternionic multiplication. ★

Exercise. Write formula (5.32) for conjugation of a pure quaternion by a unit quaternion ${q_0}^2 + \mathbf{q}\cdot\mathbf{q} = 1$ as a 3×3 matrix operation acting on a vector. ★

Answer. As a 3×3 matrix operation acting on a vector, $\boldsymbol{r}' = O_{3\times3}\boldsymbol{r}$, formula (5.32) becomes

$$\begin{aligned} \boldsymbol{r}' &= \boldsymbol{r} + 2q_0(\boldsymbol{q}\times\boldsymbol{r}) + 2\boldsymbol{q}\times(\boldsymbol{q}\times\boldsymbol{r}) \\ &= \Big[(2{q_0}^2 - 1)Id + 2q_0\widehat{q} + 2\mathbf{q}\mathbf{q}^T\Big]\boldsymbol{r} =: O_{3\times3}\boldsymbol{r}\,, \end{aligned}$$

where $\widehat{q} = \boldsymbol{q}\times$, or in components $\widehat{q}_{lm} = -q^k\epsilon_{klm}$ by the hat map in (3.7) and (3.12). When $\mathfrak{q} = [q_0, \boldsymbol{q}]$ is a unit quaternion, the Euler–Rodrigues formula implies $O_{3\times3} \in SO(3)$. ▲

Definition

5.25 (Euler parameters). *In the Euler–Rodrigues formula (5.33) for the rotation of vector* $\mathbf{r}$ *by angle* θ *about* $\hat{\mathbf{n}}$*, the quantities* θ, $\hat{\mathbf{n}}$ *are called the* ***Euler parameters****.*

Definition

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

$$\hat{\mathfrak{q}} := \pm\Big[\cos\frac{\theta}{2}\,,\, \sin\frac{\theta}{2}\,\hat{\mathbf{n}}\Big]\,. \tag{5.34}$$

The quantities $q_0 = \pm\cos\frac{\theta}{2}$ *and* $\mathbf{q} = \pm\sin\frac{\theta}{2}\,\hat{\mathbf{n}}$ *in (5.34) are called the* ***Cayley–Klein parameters****.*

Remark

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

$$\mathfrak{q} = |\mathfrak{q}|\hat{\mathfrak{q}} = |\mathfrak{q}|\Big[\cos\frac{\theta}{2},\, \sin\frac{\theta}{2}\,\hat{\mathbf{n}}\Big]\,. \tag{5.35}$$

The calculation of the Euler–Rodrigues formula (5.33) shows the equivalence of quaternionic conjugation and rotations of vectors. Moreover, compositions of quaternionic products imply the following.

Corollary

5.28. *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. Compute $O^{\pi}_{\hat{\boldsymbol{y}}}O^{\pi}_{\hat{\boldsymbol{x}}} - O^{\pi}_{\hat{\boldsymbol{x}}}O^{\pi}_{\hat{\boldsymbol{y}}}$ by quaternionic multiplication. Does it vanish? Prove it. ★

Remark

5.29 (Cayley–Klein parameters for three-vectors). *Consider the unit Cayley–Klein quaternion,* $\hat{\mathfrak{p}} := \pm[\cos\frac{\theta}{2}\,,\, \sin\frac{\theta}{2}\,\hat{\boldsymbol{\chi}}]$*. Then the decompositions for quaternions (5.26) and for vectors (5.27) may be set equal to find*

$$\begin{aligned}
[0,\ \hat{\mathbf{v}}] \;&:=\; |\mathbf{v}|^{-1}[0,\mathbf{v}] = \hat{\mathfrak{p}}[0,\hat{\boldsymbol{\omega}}]\hat{\mathfrak{p}}^*\\
&=\; [0, \cos\theta\,\hat{\boldsymbol{\omega}} + \sin\theta\,\hat{\boldsymbol{\chi}}\times\hat{\boldsymbol{\omega}}]\\
&=\; |\mathbf{v}|^{-1}[\alpha,\ \boldsymbol{\chi}][0,\hat{\boldsymbol{\omega}}]\\
&=\; (\alpha^2+\chi^2)^{-1/2}\,[0, \alpha\,\hat{\boldsymbol{\omega}} + \boldsymbol{\chi}\times\hat{\boldsymbol{\omega}}]\,.
\end{aligned}$$

Thus, the unit vector $\hat{\mathbf{v}} = |\mathbf{v}|^{-1}\mathbf{v}$ *is a rotation of* $\hat{\boldsymbol{\omega}}$ *by angle* θ *around* $\hat{\boldsymbol{\chi}}$ *with*

$$\cos\theta = \frac{\alpha}{(\alpha^2+\chi^2)^{1/2}} \quad \textit{and} \quad \sin\theta = \frac{\chi}{(\alpha^2+\chi^2)^{1/2}}\,.$$

Hence, the alignment parameters α and χ in (5.26) and (5.27) define the three-vector $\mathbf{v}$ in $[0,\mathbf{v}] = [\alpha,\boldsymbol{\chi}][0,\hat{\boldsymbol{\omega}}]$ as a stretching of $\hat{\boldsymbol{\omega}}$ by $(\alpha^2+\chi^2)^{1/2}$ and a rotation of $\hat{\boldsymbol{\omega}}$ by $\theta = \tan^{-1}\chi/\alpha$ about $\hat{\boldsymbol{\chi}}$. The Cayley–Klein angle θ is the ***relative angle*** *between the directions $\hat{\mathbf{v}}$ and $\hat{\boldsymbol{\omega}}$.*

5.1.5. *Pure quaternions, Pauli matrices and $SU(2)$.*

Exercise. Write the product of two pure unit quaternions as a multiplication of Pauli matrices. ★

Answer. By the quaternionic multiplication rule (5.5), one finds

$$[0\,,\,\hat{\mathbf{v}}][0\,,\,\hat{\mathbf{w}}] = [-\,\hat{\mathbf{v}}\cdot\hat{\mathbf{w}}\,,\,\hat{\mathbf{v}}\times\hat{\mathbf{w}}] =: [\cos\theta\,,\,\hat{\boldsymbol{n}}\sin\theta]\,. \tag{5.36}$$

Here $\hat{\mathbf{v}}\cdot\hat{\mathbf{w}} = -\cos\theta$, so that θ is the relative angle between the unit three-vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$, and $\hat{\mathbf{v}}\times\hat{\mathbf{w}} = \hat{\boldsymbol{n}}\sin\theta$ is their cross product, satisfying

$$|\hat{\mathbf{v}}\times\hat{\mathbf{w}}|^2 = |\hat{\mathbf{v}}|^2|\hat{\mathbf{w}}|^2 - (\hat{\mathbf{v}}\cdot\hat{\mathbf{w}})^2 = 1-\cos^2\theta = \sin^2\theta\,.$$

This is equivalent to following the multiplication of Pauli matrices,

$$\begin{aligned}(-i\hat{\mathbf{v}}\cdot\boldsymbol{\sigma})(-i\hat{\mathbf{w}}\cdot\boldsymbol{\sigma}) &= -\,\hat{\mathbf{v}}\cdot\hat{\mathbf{w}}\,\sigma_0 - i\,\hat{\mathbf{v}}\times\hat{\mathbf{w}}\cdot\boldsymbol{\sigma}\\ &= -\,(\cos\theta\,\sigma_0 + i\,\sin\theta\,\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma})\,,\end{aligned} \tag{5.37}$$

with, e.g., $\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma} = \sum_{a=1}^3 \hat{n}_a\sigma_a$. ▲

Proposition

5.30 (De Moivre's theorem for quaternions)**.** *De Moivre's theorem for complex numbers of unit modulus is*

$$(\cos\theta + i\sin\theta)^m = (\cos m\theta + i\sin m\theta)\,.$$

The analogue of De Moivre's theorem for unit quaternions is

$$[\cos\theta,\,\sin\theta\hat{\mathbf{n}}]^m = [\cos m\theta,\,\sin m\theta\hat{\mathbf{n}}]\,.$$

Proof. The proof follows immediately from the Cayley–Klein representation of a unit quaternion. □

Theorem

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

Proof. The matrix representation of a unit quaternion is given in (5.8). 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

$$\begin{aligned} Q &= q_0\sigma_0 - i\boldsymbol{q}\cdot\boldsymbol{\sigma}\\ &= \begin{bmatrix} q_0 - iq_3 & -iq_1 - q_2\\ -iq_1 + q_2 & q_0 + iq_3\end{bmatrix}.\end{aligned} \tag{5.38}$$

The matrix Q is a unitary 2×2 matrix ($QQ^\dagger = Id$) with unit determinant ($\det Q = 1$). That is, $Q\in SU(2)$. In fact, we may rewrite the map (5.28) for quaternionic conjugation of a vector $\mathfrak{r} = [0,\,\boldsymbol{r}]$ by a unit quaternion equivalently in terms of unitary conjugation of the Hermitian Pauli spin matrices as

$$\mathfrak{r}' = \hat{\mathfrak{q}}\,\mathfrak{r}\,\hat{\mathfrak{q}}^* \iff \mathbf{r}'\cdot\boldsymbol{\sigma} = Q\,\mathbf{r}\cdot\boldsymbol{\sigma}Q^\dagger\,, \tag{5.39}$$

with

$$\mathbf{r}\cdot\boldsymbol{\sigma} = \begin{bmatrix} r_3 & r_1 - ir_2\\ r_1 + ir_2 & -\,r_3\end{bmatrix}. \tag{5.40}$$

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

5.32. *A variant of the map (5.38), known as the Kustaanheimo–Stiefel map, establishes a relation between the solutions of a constrained isotropic harmonic oscillator in four dimensions and those of the Kepler problem in three dimensions. However, the KS map is beyond our present scope.*

Remark

5.33. *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 (5.38) 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 (5.39) arising from the isomorphism between unit quaternions and $SU(2)$. ★

Remark

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

5.1.6. *Tilde map:* $\mathbb{R}^3 \simeq su(2) \simeq so(3)$. The following ***tilde map*** may be defined by considering the isomorphism (5.8) for a pure quaternion $[0, \boldsymbol{q}]$. Namely,

$$
\begin{aligned}
\boldsymbol{q} \in \mathbb{R}^3 \mapsto -i\, \boldsymbol{q} \cdot \boldsymbol{\sigma} \;&=\; -\,i \sum_{j=1}^{3} q_j \sigma_j \\
&=\; \begin{bmatrix} -iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & iq_3 \end{bmatrix} =: \tilde{\boldsymbol{q}} \in su(2)\,.
\end{aligned}
\tag{5.41}
$$

The tilde map (5.41) is a Lie algebra isomorphism between $\mathbb{R}^3$ with the cross product of vectors 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,

$$
\boxed{\mathrm{Im}(\mathfrak{p}\mathfrak{q}) = \frac{1}{2}\Big(\mathfrak{p}\mathfrak{q} - \mathfrak{q}\mathfrak{p}\Big) = [\tilde{\boldsymbol{p}},\, \tilde{\boldsymbol{q}}] = (\boldsymbol{p} \times \boldsymbol{q})\tilde{}\,.} \tag{5.42}
$$

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

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

5.1.7. *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

$$
\breve{\boldsymbol{\mu}} \cdot \mathbf{q} := \big\langle \mu, \tilde{\mathbf{q}} \big\rangle_{su(2)^* \times su(2)}
$$

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

Then, for example,

$$
\breve{\boldsymbol{\mu}} \cdot (\boldsymbol{p} \times \boldsymbol{q}) := \big\langle \mu, [\tilde{\boldsymbol{p}},\, \tilde{\boldsymbol{q}}] \big\rangle_{su(2)^* \times su(2)},
$$

which foreshadows the adjoint and coadjoint actions of $SU(2)$ in rigid-body dynamics.

5.1.8. *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 (5.8) and the tilde map (5.41) 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

$$n_\alpha = \sum_{k,l=1}^{2} a_k^* \{\sigma_\alpha\}_{kl}\, a_l \quad \text{with} \quad \alpha = 0, 1, 2, 3\,. \tag{5.43}$$

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

$$\begin{aligned} n_0 &= |a_1|^2 + |a_2|^2\,, \\ n_3 &= |a_1|^2 - |a_2|^2\,, \\ n_1 + i\,n_2 &= 2a_1^* a_2\,. \end{aligned} \tag{5.44}$$

Remark

5.35. *One may motivate the definition of $n_\alpha \in \mathbb{R}^4$ in (5.43) by introducing the following Hermitian matrix,*

$$\rho = \mathbf{a} \otimes \mathbf{a}^* = \frac{1}{2}\Big(n_0\sigma_0 + \mathbf{n}\cdot\boldsymbol{\sigma}\Big)\,, \tag{5.45}$$

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

$$\mathbf{n} = \operatorname{tr}\rho\,\boldsymbol{\sigma} = a_l a_k^* \boldsymbol{\sigma}_{kl}\,. \tag{5.46}$$

The last equation recovers (5.43). We will return to the interpretation of this map when we discuss momentum maps in Chapter **??**. *For now, we simply observe that the components of the singular Hermitian matrix* (det $\rho = 0$)

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

are all invariant under the diagonal action

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

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

$$n_0^2 = n_1^2 + n_2^2 + n_3^2 = |\mathbf{n}|^2\,. \tag{5.47}$$

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.

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

5.1.9. *Poincaré's sphere and Hopf's fibration.* The same map $S^3 \mapsto S^2$ given by (5.43) from the $n_0 = const$ S^3 to the Poincaré sphere S^2 was later studied by Hopf, who found it to be a ***fibration*** of S^3 over S^2. That is, $S^3 \simeq S^2 \times S^1$ locally, where S^1 is the fibre. A fibre bundle structure is defined descriptively, as follows.

Definition

5.36 (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. However, we shall confine our attention here to the one particular case leading to the Poincaré sphere.*

Remark

5.37. *The **Hopf fibration**, or fibre bundle, $S^3 \simeq S^2 \times S^1$ has spheres as its total space, base space and fibre, respectively. In terms of the Poincaré sphere one may think of the Hopf fibration locally as a sphere S^2 which has a great circle S^1 attached at every point. The phase on the great circles at opposite points are orthogonal (rotated by $\pi/2$, not 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.*

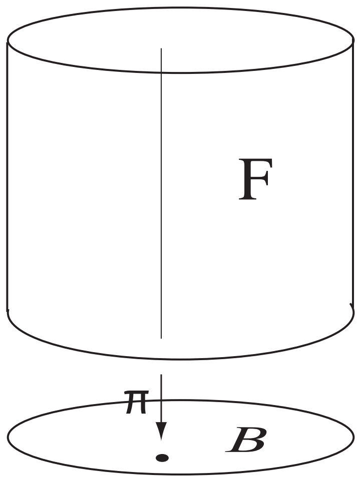


FIGURE 3. A fibre bundle E looks locally like the product space $B \times F$, of the base space B with the fibre space F. The map $\pi : E \approx B \times F \mapsto B$ projects E onto the base space B.

Remark

5.38. *The conjugacy classes of S^3 by unit quaternions yield the family of two-spheres S^2 in formula (5.31) of Corollary 5.22. These also produce a version of the Hopf fibration $S^3 \simeq S^2 \times S^1$, obtained by identifying the Poincaré sphere (5.47) from the definitions (5.44).*

Remark

5.39 (Hopf fibration/quaternionic conjugation)**.** *Conjugating the pure unit quaternion along the z-axis $[0, \hat{\mathbf{z}}]$ by the other unit quaternions yields the entire unit two-sphere S^2. This is to be expected from the complete set of unit vectors found by quaternionic conjugation in (5.29). However, it may be shown explicitly by computing the $SU(2)$ multiplication for $|a_1|^2 + |a_2|^2 = 1$,*

$$\begin{bmatrix} a_1 & -a_2^* \\ a_2 & a_1^* \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} a_1^* & a_2^* \\ -a_2 & a_1 \end{bmatrix} = \begin{bmatrix} -in_3 & -in_1 + n_2 \\ -in_1 - n_2 & in_3 \end{bmatrix}. \tag{5.48}$$

This is the tilde map (5.41) once again and (n_1, n_2, n_3) are the components of the Hopf fibration [MaRa1994].

In other words, cf. Equation (5.38),

$$-ig\sigma_3 g^\dagger = -i\mathbf{n}\cdot\boldsymbol{\sigma}.$$

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

Remark

5.40. *The isomorphism given in (5.8), (5.41) and (5.48) between the unit quaternions and $SU(2)$ expressed in terms of the Pauli spin matrices connects the quaternions to the mathematics of Poincaré's sphere $\mathbb{C}^2 \mapsto S^2$, Hopf's fibration $S^3 \simeq S^2 \times S^1$ and the geometry of fibre bundles. This deep network of connections would amply reward the efforts of further study.*

Exercise. Show that the Hopf fibration is a decomposition law for the group $SU(2)$.
Hint: Write the Hopf fibration in quaternionic form. ★

Exercise. Write the quaternionic version of unitary transformations of Hermitian matrices.
Hint: The Pauli spin matrices defined in (5.9) are Hermitian. To get started, you may want to take a look at Equation (5.39). ★

5.2. **Rotations in four dimensions: $SO(4)$.**

Scenario 5.41. *The Tets are yet another alien life form who also use one-dimensional time $t \in \mathbb{R}$ (we sigh with relief), but their spatial coordinates are $\mathbf{X} \in \mathbb{R}^4$, while ours are $\mathbf{x} \in \mathbb{R}^3$. They test us to determine whether we are an intelligent life form by requiring us to write the equations for rigid-body motion for four-dimensional rotations.*

Hint: The angular velocity of rotation $\widehat{\Psi} = O^{-1}\dot{O}(t)$ for rotations $O(t) \in SO(4)$ in four dimensions will be represented by a 4×4 skew-symmetric matrix. Write a basis for the 4×4 skew-symmetric matrices by adding a row and column to the 3×3 basis.

Answer. Any 4×4 skew-symmetric matrix may be represented as a linear combination of 4×4 basis matrices with three-dimensional vector coefficients Ω, $\Lambda \in \mathbb{R}^3$ in the form

$$
\begin{aligned}
\widehat{\Psi} &= \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 & -\Lambda_1 \\ \Omega_3 & 0 & -\Omega_1 & -\Lambda_2 \\ -\Omega_2 & \Omega_1 & 0 & -\Lambda_3 \\ \Lambda_1 & \Lambda_2 & \Lambda_3 & 0 \end{pmatrix} \\
&= \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K} \\
&= \Omega_a \widehat{J}_a + \Lambda_b \widehat{K}_b \,.
\end{aligned}
$$

This is the formula for the angular velocity of rotation in four dimensions.

The 4×4 basis set $\widehat{J} = (J_1,\, J_2,\, J_3)^T$ and $\widehat{K} = (K_1,\, K_2,\, K_3)^T$ consists of the following six linearly independent 4×4 skew-symmetric matrices, $\widehat{J}_a$, $\widehat{K}_b$ with $a, b = 1, 2, 3$:

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

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

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

The matrices $\widehat{J}_a$ with $a = 1, 2, 3$ embed the basis for 3×3 skew-symmetric matrices into the 4×4 matrices by adding a row and column of zeros. The skew matrices $\widehat{K}_a$ with $a = 1, 2, 3$ then extend the 3×3 basis to 4×4.

5.2.1. *Commutation relations.* The skew matrix basis $\widehat{J}_a$, $\widehat{K}_b$ with $a, b = 1, 2, 3$ satisfies the commutation relations,

$$
\begin{aligned}
\big[\,\widehat{J}_a,\, \widehat{J}_b\,\big] &= \widehat{J}_a\widehat{J}_b - \widehat{J}_b\widehat{J}_a = \epsilon_{abc}\widehat{J}_c\,, \\
\big[\,\widehat{J}_a,\, \widehat{K}_b\,\big] &= \widehat{J}_a\widehat{K}_b - \widehat{K}_b\widehat{J}_a = \epsilon_{abc}\widehat{K}_c\,, \\
\big[\,\widehat{K}_a,\, \widehat{K}_b\,\big] &= \widehat{K}_a\widehat{K}_b - \widehat{K}_b\widehat{K}_a = \epsilon_{abc}\widehat{J}_c\,.
\end{aligned}
$$

These commutation relations may be verified by a series of direct calculations, as $[\,\widehat{J}_1,\, \widehat{J}_2\,] = \widehat{J}_3$, etc.

5.2.2. *Hat map for* 4×4 *skew matrices.* The map above for the 4×4 skew matrix $\widehat{\Psi}$ may be written as

$$
\widehat{\Psi} = \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K} = \Omega_a\widehat{J}_a + \Lambda_b\widehat{K}_b\,, \ \text{ sum on } a, b = 1, 2, 3\,.
$$

This map provides the 4×4 version of the hat map, written now as $(\,\cdot\,)\widehat{\ }: \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathfrak{so}(4)$. Here $\mathfrak{so}(4)$ is the Lie algebra of the 4×4 special orthogonal matrices, which consists of the 4×4 skew matrices represented in the six-dimensional basis of $\widehat{J}$'s and $\widehat{K}$'s.

5.2.3. *Commutator as intertwined vector product.* The commutator of 4×4 skew matrices corresponds to an intertwined vector product, as follows. For any vectors Ω, Λ, ω, $\lambda \in \mathbb{R}^3$, one has

$$
\begin{aligned}
&\Big[\,\Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K}\,,\, \omega \cdot \widehat{J} + \lambda \cdot \widehat{K}\,\Big] \\
&\qquad = \Big(\Omega \times \omega + \Lambda \times \lambda\Big) \cdot \widehat{J} + \Big(\Omega \times \lambda - \Lambda \times \omega\Big) \cdot \widehat{K}\,.
\end{aligned}
$$

Likewise, the matrix pairing $\langle\, A, B\,\rangle = \operatorname{tr}(A^TB)$ is related to the vector dot-product pairing in $\mathbb{R}^3$ by

$$
\Big\langle\, \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K}\,,\, \omega \cdot \widehat{J} + \lambda \cdot \widehat{K}\,\Big\rangle = \Omega \cdot \omega + \Lambda \cdot \lambda\,.
$$

That is,

$$
\Big\langle\, \widehat{J}_a\,,\, \widehat{J}_b\,\Big\rangle = \delta_{ab} = \Big\langle\, \widehat{K}_a\,,\, \widehat{K}_b\,\Big\rangle \quad \text{and} \quad \Big\langle\, \widehat{J}_a\,,\, \widehat{K}_b\,\Big\rangle = 0\,.
$$

5.2.4. *Euler–Poincaré equation on* $\mathfrak{so}(4)^*$. For

$$
\Phi = O^{-1}\delta O(t) = \xi \cdot \widehat{J} + \eta \cdot \widehat{K} \in \mathfrak{so}(4)\,,
$$

Hamilton's principle $\delta S = 0$ for $S = \int_a^b \ell(\Psi)\,dt$ with

$$
\Psi = O^{-1}\dot{O}(t) = \Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K} \in \mathfrak{so}(4)
$$

leads to

$$
\delta S = \int_a^b \Big\langle \frac{\delta \ell}{\delta \Psi}\,,\, \delta\Psi \Big\rangle\, dt = \int_a^b \Big\langle \frac{\delta \ell}{\delta \Psi}\,,\, \dot{\Phi} + \operatorname{ad}_\Psi \Phi \Big\rangle\, dt\,,
$$

where

$$
\begin{aligned}
\operatorname{ad}_\Psi \Phi &= [\Psi,\, \Phi] = \big[\,\Omega \cdot \widehat{J} + \Lambda \cdot \widehat{K},\, \xi \cdot \widehat{J} + \eta \cdot \widehat{K}\,\big] \\
&= \Big(\Omega \times \xi + \Lambda \times \eta\Big) \cdot \widehat{J} + \Big(\Omega \times \eta - \Lambda \times \xi\Big) \cdot \widehat{K}\,.
\end{aligned}
$$

Thus,

$$\begin{aligned}
\delta S &= \int_a^b \Big\langle -\frac{d}{dt}\frac{\delta \ell}{\delta \Psi}\,,\, \Phi \Big\rangle + \Big\langle \frac{\delta \ell}{\delta \Psi}\,,\, \mathrm{ad}_\Psi \Phi \Big\rangle\, dt \\
&= \int_a^b \Big\langle -\frac{d}{dt}\frac{\delta \ell}{\delta \Omega}\cdot \widehat{J} - \frac{d}{dt}\frac{\delta \ell}{\delta \Lambda}\cdot \widehat{K}\,,\, \xi\cdot \widehat{J} + \eta\cdot \widehat{K} \Big\rangle\, dt \\
&\quad + \int_a^b \Big\langle \frac{\delta \ell}{\delta \Omega}\cdot \widehat{J} + \frac{\delta \ell}{\delta \Lambda}\cdot \widehat{K}\,, \\
&\qquad \Big(\Omega\times\xi + \Lambda\times\eta\Big)\cdot \widehat{J} + \Big(\Omega\times\eta - \Lambda\times\xi\Big)\cdot \widehat{K} \Big\rangle\, dt \\
&= \int_a^b \Big(-\frac{d}{dt}\frac{\delta \ell}{\delta \Omega} + \frac{\delta \ell}{\delta \Omega}\times\Omega - \frac{\delta \ell}{\delta \Lambda}\times\Lambda \Big)\cdot \xi \\
&\quad + \Big(-\frac{d}{dt}\frac{\delta \ell}{\delta \Lambda} + \frac{\delta \ell}{\delta \Lambda}\times\Omega + \frac{\delta \ell}{\delta \Omega}\times\Lambda \Big)\cdot \eta\, dt\,.
\end{aligned}$$

Hence, $\delta S = 0$ yields

$$\begin{aligned}
\frac{d}{dt}\frac{\delta \ell}{\delta \Omega} &= \frac{\delta \ell}{\delta \Omega}\times\Omega - \frac{\delta \ell}{\delta \Lambda}\times\Lambda \\
\text{and}\quad \frac{d}{dt}\frac{\delta \ell}{\delta \Lambda} &= \frac{\delta \ell}{\delta \Lambda}\times\Omega + \frac{\delta \ell}{\delta \Omega}\times\Lambda\,.
\end{aligned} \tag{5.49}$$

These are the $\widehat{J}$, $\widehat{K}$ basis components of the Euler–Poincaré equation on $\mathfrak{so}(4)^*$,

$$\frac{d}{dt}\frac{\delta \ell}{\delta \Psi} = \mathrm{ad}^*_\Psi \frac{\delta \ell}{\delta \Psi}\,,$$

written with $\Psi = \Omega\cdot\widehat{J} + \Lambda\cdot\widehat{K}$ in this basis.

5.2.5. *Hamiltonian form on* $\mathfrak{so}(4)^*$. Legendre-transforming yields the pairs

$$\Pi = \frac{\delta \ell}{\delta \Omega}\,, \qquad \Omega = \frac{\delta h}{\delta \Pi}\,, \quad \text{and} \quad \Xi = \frac{\delta \ell}{\delta \Lambda}\,, \qquad \Lambda = \frac{\delta h}{\delta \Xi}\,.$$

Hence, these equations may be expressed in Hamiltonian form as

$$\frac{d}{dt}\begin{bmatrix}\Pi\\ \Xi\end{bmatrix} = \begin{bmatrix}\Pi\times & \Xi\times\\ \Xi\times & \Pi\times\end{bmatrix}\begin{bmatrix}\delta h/\delta\Pi\\ \delta h/\delta\Xi\end{bmatrix}. \tag{5.50}$$

The corresponding Lie–Poisson bracket is given by

$$\begin{aligned}
\{f,\, h\} &= -\,\Pi\cdot\left(\frac{\delta f}{\delta \Pi}\times\frac{\delta h}{\delta \Pi} + \frac{\delta f}{\delta \Xi}\times\frac{\delta h}{\delta \Xi}\right) \\
&\quad -\,\Xi\cdot\left(\frac{\delta f}{\delta \Pi}\times\frac{\delta h}{\delta \Xi} - \frac{\delta h}{\delta \Pi}\times\frac{\delta f}{\delta \Xi}\right).
\end{aligned}$$

This Lie–Poisson bracket has an extra term proportional to Π, relative to the $se(3)^*$ bracket for the heavy top. Its Hamiltonian matrix has two null eigenvectors for the variational derivatives of $C_1 = |\Pi|^2 + |\Xi|^2$ and $C_2 = \Pi\cdot\Xi$. The functions C_1, C_2 are the Casimirs of the $\mathfrak{so}(4)$ Lie–Poisson bracket. That is, $\{C_1,\, H\} = 0 = \{C_2,\, H\}$ for every Hamiltonian $H(\Pi,\, \Xi)$.

The Hamiltonian matrix in Equation (5.50) is similar to that for the Lie–Poisson formulation of heavy-top dynamics, except for the one extra term $\{\Xi,\, \Xi\} \neq 0$. ▲

5.3. **Rotations in complex space.**

Scenario 5.42. *The Bers are another alien life form who use one-dimensional time $t \in \mathbb{R}$ (thankfully), but their spatial coordinates are complex $\mathbf{z} \in \mathbb{C}^3$, while ours are real $\mathbf{x} \in \mathbb{R}^3$. They test us to determine whether we are an intelligent life form by requiring us to write the equations for rigid-body motion for body angular momentum coordinates $\mathbf{L} \in \mathbb{C}^3$.*

Their definition of a rigid body requires its moment of inertia $\mathbb{I}$, rotational kinetic energy $\frac{1}{2}\mathbf{L}\cdot\mathbb{I}^{-1}\mathbf{L}$ and magnitude of body angular momentum $\sqrt{\mathbf{L}\cdot\mathbf{L}}$ all to be real. They also tell us these rigid-body equations must be invariant under the operations of parity $\mathcal{P}\,\mathbf{z} \to -\mathbf{z}^*$ and time reversal $\mathcal{T} : t \to -t$.

What equations should we give them? Are these equations the same as ours in real body angular momentum coordinates? Keep your approach general for as long as you like, but if you wish to simplify, work out your results with the simple example in which $\mathbb{I} = \mathrm{diag}(1, 2, 3)$.

Answer. Euler's equations for free rotational motion of a rigid body about its centre of mass may be expressed in real vector coordinates $\mathbf{L} \in \mathbb{R}^3$ ($\mathbf{L}$ is the body angular momentum vector) as

$$\dot{\mathbf{L}} = \frac{\partial C}{\partial \mathbf{L}} \times \frac{\partial E}{\partial \mathbf{L}}\,, \tag{5.51}$$

where C and E are conserved quadratic functions defined by

$$C(\mathbf{L}) = \frac{1}{2}\mathbf{L}\cdot\mathbf{L}\,, \quad E(\mathbf{L}) = \frac{1}{2}\mathbf{L}\cdot\mathbb{I}^{-1}\mathbf{L}\,. \tag{5.52}$$

Here, $\mathbb{I}^{-1} = \mathrm{diag}\,(I_1^{-1}, I_2^{-1}, I_3^{-1})$ is the inverse of the (real) moment of inertia tensor in principal axis coordinates. These equations are $\mathcal{PT}$-symmetric; they are invariant under spatial reflections of the angular momentum components in the body $P : \mathbf{L} \to \mathbf{L}$ composed with time reversal $T : \mathbf{L} \to -\mathbf{L}$. The simplifying choice $\mathbb{I}^{-1} = \mathrm{diag}(1, 2, 3)$ reduces the dynamics (5.51) to

$$\dot{L}_1 = L_2L_3\,, \quad \dot{L}_2 = -\,2L_1L_3\,, \quad \dot{L}_3 = L_1L_2\,, \tag{5.53}$$

which may also be written equivalently as

$$\dot{\mathbf{L}} = \mathbf{L} \times \mathsf{K}\mathbf{L}\,, \tag{5.54}$$

with $\mathsf{K} = \mathrm{diag}(-1, 0, 1)$.

Since $\mathbf{L}$ is complex, we set $\mathbf{L} = \mathbf{x} + i\mathbf{y}$ and obtain *four* conservation laws, namely the real and imaginary parts of $C(\mathbf{L}) = \frac{1}{2}\mathbf{L}\cdot\mathbf{L}$ and $H(\mathbf{L}) = \frac{1}{2}\mathbf{L}\cdot\mathsf{K}\mathbf{L}$, expressed as

$$C(\mathbf{L}) \;=\; \frac{1}{2}\mathbf{x}\cdot\mathbf{x} - \frac{1}{2}\mathbf{y}\cdot\mathbf{y} + i\mathbf{x}\cdot\mathbf{y}\,, \tag{5.55}$$

$$H(\mathbf{L}) \;=\; \frac{1}{2}\mathbf{x}\cdot\mathsf{K}\mathbf{x} - \frac{1}{2}\mathbf{y}\cdot\mathsf{K}\mathbf{y} + i\mathbf{x}\cdot\mathsf{K}\mathbf{y}\,. \tag{5.56}$$

The solutions to Euler's equations that have been studied in the past are the *real* solutions to (5.53), that is, the solutions for which $\mathbf{y} = 0$. For this case the phase space is three-dimensional and the two conserved quantities are

$$C = \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right), \quad H = -\frac{1}{2}x_1^2 + \frac{1}{2}x_3^2\,. \tag{5.57}$$

If we take $C = \frac{1}{2}$, then the phase-space trajectories are constrained to a sphere of radius 1. There are six critical points located at $(\pm1, 0, 0)$, $(0, \pm1, 0)$ and $(0, 0, \pm1)$. These are the conventional trajectories that are discussed in standard textbooks on dynamical systems [MaRa1994].

Exercise. When $H = 0$, show that the resulting equation is a first integral of the simple pendulum problem. ★

Let us now examine the complex $\mathcal{PT}$-symmetric solutions to Euler's equations. The equation set (5.52) is six-dimensional. However, a reduction in dimension occurs because the requirement of $\mathcal{PT}$ symmetry requires the constants of motion C and H in (5.56) to be real. The vanishing of the imaginary parts of C and H gives the two equations

$$\mathbf{x}\cdot\mathbf{y} = 0, \quad \mathbf{x}\cdot\mathsf{K}\mathbf{y} = 0\,. \tag{5.58}$$

These two bilinear constraints may be used to eliminate the $\mathbf{y}$ terms in the complex Equations (5.53). When this elimination is performed using the definition $\mathsf{K} = \mathrm{diag}(-1, 0, 1)$, one obtains the following real equations for $\mathbf{x}$ on the $\mathcal{PT}$ constraint manifolds (5.58):

$$\dot{\mathbf{x}} = \mathbf{x} \times \mathsf{K}\mathbf{x} + M(\mathbf{x})\,\mathbf{x}\,. \tag{5.59}$$

Here, the scalar function $M = PN/D$, where the functions P, N and D are given by

$$\begin{aligned} P(\mathbf{x}) &= 2x_1x_2x_3\,, \quad N(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - 1\,, \\ D(\mathbf{x}) &= \left|\mathrm{Re}\left(\frac{\partial C}{\partial \mathbf{L}} \times \frac{\partial H}{\partial \mathbf{L}}\right)\right|^2 = x_1^2x_2^2 + x_2^2x_3^2 + 4x_1^2x_3^2\,. \end{aligned} \tag{5.60}$$

The system (5.59) has nonzero divergence, so it cannot be Hamiltonian even though it arises from constraining a Hamiltonian system. Nonetheless, the system has two *additional* real conservation laws, and it reduces to the integrable form

$$
\begin{aligned}
\dot{x}_1 &= x_2x_3\left(1+2x_1^2N/D\right) , && (5.61)\\
\dot{x}_2 &= -2x_1x_3\left(1-x_2^2N/D\right) , &&\\
\dot{x}_3 &= x_1x_2\left(1+2x_3^2N/D\right) , && (5.62)
\end{aligned}
$$

on level sets of two conserved quantities:

$$
\begin{aligned}
A &= \frac{(N+1)^2N}{D} , && (5.63)\\
B &= \frac{x_1^2-x_3^2}{D}\left(2x_2^2x_3^2+4x_1^2x_3^2+x_2^4+2x_1^2x_2^2-x_2^2\right) .
\end{aligned}
$$

Hence, the motion takes place in $\mathbb{R}^3$ on the intersection of the level sets of these two conserved quantities. These quantities vanish when either $N=0$ (the unit sphere) or $x_3^2-x_1^2=0$ (the degenerate hyperbolic cylinder). On these level sets of the conserved quantities the motion Equations (5.59) restrict to Equations (5.53) for the original real rigid body. ▲

Remark

5.43. *We are dealing with rotations of the group of complex* 3×3 *orthogonal matrices with unit determinant acting on complex three-vectors. These are the linear maps,* $SO(3,\mathbb{C})\times\mathbb{C}^3\mapsto\mathbb{C}^3$.

Euler's Equations (5.51) for complex body angular momentum describe geodesic motion on $SO(3,\mathbb{C})$ *with respect to the metric given by the trace norm* $g(\Omega,\Omega)=\frac{1}{2}\mathrm{trace}\,(\Omega^T\mathbb{I}\Omega)$ *for the real symmetric moment of inertia tensor* $\mathbb{I}$ *and left-invariant Lie algebra element* $\Omega(t)=g^{-1}(t)\dot{g}(t)\in so(3,\mathbb{C})$. *Because* $SO(3,\mathbb{C})$ *is orthogonal,* $\Omega\in so(3,\mathbb{C})$ *is a* 3×3 *complex skew-symmetric matrix, which may be identified with complex vectors* $\widehat{\boldsymbol{\Omega}}\in\mathbb{C}^3$ *by* $(\Omega)_{jk}=-\widehat{\Omega}^i\epsilon_{ijk}$. *Euler's Equations (5.51) follow from Hamilton's principle in Euler–Poincaré or Lie–Poisson form:*

$$
\dot{\mu}=\mathrm{ad}^*_\Omega\,\mu=\{\mu,\,H\}\,, \tag{5.64}
$$

where

$$
\frac{\delta l}{\delta\Omega}=\mu\,,\quad g^{-1}\dot{g}=\Omega\,,\quad \Omega=\frac{\partial H}{\partial\mu}\,. \tag{5.65}
$$

These are Hamiltonian with the standard Lie–Poisson bracket defined on the dual Lie algebra $so(3,\mathbb{C}^3)^*$. *Because of the properties of the trace norm, we may take* $\mu=\mathrm{skew}\,\mathbb{I}\Omega$. *(Alternatively, we may set the preserved symmetric part of* μ *initially to zero.) Hence,* μ *may be taken as a skew-symmetric complex matrix, which again may be identified with the components of a complex three-vector* $\mathbf{z}$ *as* $(\mu)_{jk}=-z^i\epsilon_{ijk}$. *On making this identification, Euler's Equations (5.51) emerge for* $\mathbf{z}\in\mathbb{C}^3$, *with real* $\mathbb{I}$. *The* $\mathcal{PT}$*-symmetric initial conditions on the real level sets of the preserved complex quantities* C *and* H *form an invariant manifold of this system of three complex ordinary differential equations. On this invariant manifold, the complex angular motion is completely integrable. By following the approach established by Manakov [Man1976] this reasoning may also extend to the rigid body on* $SO(n,\mathbb{C})$.

Exercise. The Bers left behind a toy monopole. This is a rigid body that rotates by complex angles and whose three moments of inertia are the complex cube roots of unity. What are the equations of motion for this toy monopole? For a hint, take a look at [Iv2006]. ★

Generalised rigid body.
Recall the following definitions for the left action of a Lie group G on the cotangent bundle T^*Q of a manifold Q:

- The diamond operation $\diamond : T^*Q \to \mathfrak{g}^*$ is defined by

$$\Big\langle p \diamond q \,,\, \xi \Big\rangle = \langle\langle p,\, -\pounds_\xi q \rangle\rangle \,,$$

with Lie derivative $\pounds_\xi q$ given by the infinitesimal generator of the action of the Lie algebra element ξ on the coordinate, q, and pairings $\langle \cdot\,,\,\cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ and $\langle\langle\, \cdot\,,\cdot\,\rangle\rangle : TQ^* \times TQ \to \mathbb{R}$.
- The cotangent-lift momentum map for this action is given by

$$J = -\, p \diamond q : T^*Q \to \mathfrak{g}^*$$

for canonical variables $(q,p) \in T^*Q$ satisfying $\{q,p\} = Id$.

Let the Hamiltonian H_{grb} for a generalized 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 $(\,\cdot\,,\,\cdot\,) : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ obtained, e.g., from the Killing form K on $\mathfrak{g}$ (which is symmetric and nondegenerate).

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.

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

★

5.4. **Two times and the continuum spin chain.**

Scenario 5.44. *The Bichrons are an alien life form who use two-dimensional time $u = (s,t) \in \mathbb{R}^2$ for time travel. To decide whether we are an intelligent life form, they require us to define spatial and body angular velocity for free rigid rotation in their two time dimensions. What should we tell them?*

Answer. **(Bichrons)** Following the Euler-Poincaré approach to rotating motion, let's define a trajectory of a moving point $\mathbf{x} = \mathbf{r}(u) \in \mathbb{R}^3$ as a smooth invertible map $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ with "time" $u \in \mathbb{R}^2$, so that $u = (s,t)$. Suppose the components of the trajectory are given in terms of a fixed and a moving orthonormal frame by

$$\begin{aligned} \mathbf{r}(u) &= r_0^A(u)\mathbf{e}_A(0) && \text{fixed frame,} \\ &= r^a\mathbf{e}_a(u) && \text{moving frame,} \end{aligned}$$

with constant $r^a \in \mathbb{R}^3$ and moving orthonormal frame defined by $O : \mathbb{R}^2 \to SO(3)$, so that

$$\mathbf{e}_a(u) = O(u)\mathbf{e}_a(0)\,.$$

Here $O(u)$ is a *map* $\mathbb{R}^2 \to SO(3)$ parameterised by the two times $u = (s,t) \in \mathbb{R}^2$. The exterior derivative[3] of the moving frame relation above yields the infinitesimal spatial displacement,

$$d\mathbf{r}(u) = r^a d\mathbf{e}_a(u) = r^a dOO^{-1}(u)\mathbf{e}_a(u) = \widehat{\omega}(u)\mathbf{r}\,,$$

in which $\widehat{\omega}(u) = dOO^{-1}(u) \in \mathfrak{so}(3)$ is the one-form for spatial angular displacement. One denotes

$$dO = O'\,ds + \dot{O}\,dt\,,$$

so that the spatial angular displacement is the right-invariant $\mathfrak{so}(3)$-valued one-form

$$\widehat{\omega}(u) = dOO^{-1} = (O'\,ds + \dot{O}\,dt)O^{-1}\,.$$

[3]This subsection uses the notation of differential forms and wedge products. Readers unfamiliar with it may regard this subsection as cultural background.

Likewise, the body angular displacement is the left-invariant $\mathfrak{so}(3)$-valued one-form

$$\begin{aligned} \widehat{\Omega}(u) &= \mathrm{Ad}_{O^{-1}}\widehat{\omega}(u) = O^{-1}\widehat{\omega}(u)O = O^{-1}dO \\ &= O^{-1}(O'\,ds + \dot{O}\,dt) =: \widehat{\Omega}_s ds + \widehat{\Omega}_t dt\,, \end{aligned} \tag{5.66}$$

and $\widehat{\Omega}_s$ and $\widehat{\Omega}_t$ are its *two body angular velocities.*

This is the answer the Bichrons wanted: For them, free rotation takes place on a surface in $SO(3)$ parameterised by $u = (s,t) \in \mathbb{R}^2$ and it has two body angular velocities because such a surface has two independent tangent vectors. ▲

Scenario 5.45. *What would the Bichrons do with this information?*

Answer. To give an idea of what the Bichrons might do with our answer, let us define the ***coframe*** at position $\mathbf{x} = \mathbf{r}(u)$ as the infinitesimal displacement in *body coordinates*,

$$\boldsymbol{\Xi} = O^{-1}d\mathbf{r}\,. \tag{5.67}$$

Taking its exterior derivative gives the two-form,

$$d\boldsymbol{\Xi} = -O^{-1}dO \wedge O^{-1}d\mathbf{r} = -\widehat{\Omega} \wedge \boldsymbol{\Xi}\,, \tag{5.68}$$

in which the left-invariant $\mathfrak{so}(3)$-valued one-form $\widehat{\Omega} = O^{-1}dO$ encodes the exterior derivative of the coframe as a rotation by the body angular displacement. In differential geometry, $\widehat{\Omega}$ is called the ***connection form*** and Equation (5.68) is called ***Cartan's first structure equation*** for a moving orthonormal frame [Fl1963, Da1994]. Taking another exterior derivative gives zero (because $d^2 = 0$) in the form of

$$0 = d^2\boldsymbol{\Xi} = -d\widehat{\Omega} \wedge \boldsymbol{\Xi} - \widehat{\Omega} \wedge d\boldsymbol{\Xi} = -(d\widehat{\Omega} + \widehat{\Omega} \wedge \widehat{\Omega}) \wedge \boldsymbol{\Xi}\,.$$

Hence we have ***Cartan's second structure equation***,

$$d\widehat{\Omega} + \widehat{\Omega} \wedge \widehat{\Omega} = 0\,. \tag{5.69}$$

The left-hand side of this equation is called the ***curvature two-form*** associated with the connection form $\widehat{\Omega}$. The interpretation of (5.69) is that the connection form $\widehat{\Omega} = O^{-1}dO$ has zero curvature. This makes sense because the rotating motion takes place in Euclidean space, $\mathbb{R}^3$, which is flat.

Of course, one may also prove the ***zero curvature relation*** (5.69) directly from the definition $\widehat{\Omega} = O^{-1}dO$ by computing

$$d\widehat{\Omega} = d(O^{-1}dO) = -O^{-1}dO \wedge O^{-1}dO = -\widehat{\Omega} \wedge \widehat{\Omega}\,.$$

Expanding this out using the two angular velocities $\widehat{\Omega}_s = O^{-1}O'$ and $\widehat{\Omega}_t = O^{-1}\dot{O}$ gives (by using antisymmetry of the wedge product, $ds \wedge dt = -dt \wedge ds$)

$$\begin{aligned} d\widehat{\Omega}(u) &= d(\widehat{\Omega}_s ds + \widehat{\Omega}_t dt) \\ &= -\Big(\widehat{\Omega}_s ds + \widehat{\Omega}_t dt\Big) \wedge \Big(\widehat{\Omega}_s ds + \widehat{\Omega}_t dt\Big) \\ &= -\widehat{\Omega} \wedge \widehat{\Omega} \\ &= \frac{\partial \widehat{\Omega}_s}{\partial t}\, dt \wedge ds + \frac{\partial \widehat{\Omega}_t}{\partial s}\, ds \wedge dt \\ &= -\widehat{\Omega}_s\widehat{\Omega}_t\, ds \wedge dt - \widehat{\Omega}_t\widehat{\Omega}_s\, dt \wedge ds \\ &= \left(\frac{\partial \widehat{\Omega}_t}{\partial s} - \frac{\partial \widehat{\Omega}_s}{\partial t}\right) ds \wedge dt \\ &= \big(\widehat{\Omega}_t\widehat{\Omega}_s - \widehat{\Omega}_s\widehat{\Omega}_t\big)\, ds \wedge dt \\ &=: \big[\,\widehat{\Omega}_t\,,\,\widehat{\Omega}_s\,\big]\, ds \wedge dt\,. \end{aligned}$$

Since $ds \wedge dt \neq 0$, this equality implies that the coefficients are equal. In other words, this calculation proves the following.

Proposition

5.46. *The zero curvature relation (5.69) may be expressed equivalently as*

$$\frac{\partial\widehat{\Omega}_t}{\partial s} - \frac{\partial\widehat{\Omega}_s}{\partial t} = \widehat{\Omega}_t\widehat{\Omega}_s - \widehat{\Omega}_s\widehat{\Omega}_t = \big[\,\widehat{\Omega}_t\,,\,\widehat{\Omega}_s\,\big]\,, \tag{5.70}$$

in terms of the two angular velocities, $\widehat{\Omega}_s = O^{-1}O'$ *and* $\widehat{\Omega}_t = O^{-1}\dot{O}$.

▲

Exercise. Why would $\widehat{\Omega}$ be called a connection form? ★

Answer. Consider the one-form Equation (5.67) written in components as

$$\Xi^j = \Xi^j_\alpha(r)dr^\alpha\,, \tag{5.71}$$

in which the matrix $\Xi^j_\alpha(r)$ depends on spatial location, and it need not be orthogonal. In the basis $\Xi^j(r)$, a one-form v may be expanded in components as

$$v = v_j\Xi^j\,. \tag{5.72}$$

Its differential is computed in this basis as

$$\begin{aligned} dv &= d(v_j\Xi^j) \\ &= dv_j\wedge\Xi^j + v_jd\Xi^j\,. \end{aligned}$$

Substituting Equation (5.68) in components as

$$d\,\Xi^j = -\,\widehat{\Omega}^j_k\wedge\Xi^k \tag{5.73}$$

then yields the differential two-form,

$$\begin{aligned} dv &= (dv_k - v_j\widehat{\Omega}^j_k)\wedge\Xi^k \\ &=: (dv_k - v_j\Gamma^j_{kl}\Xi^l)\wedge\Xi^k \\ &=: Dv_k\wedge\Xi^k\,. \end{aligned} \tag{5.74}$$

The last equation defines the covariant exterior derivative operation D in the basis of one-form displacements $\boldsymbol{\Xi}(r)$. The previous equation introduces the quantities Γ^j_{kl} defined as

$$\widehat{\Omega}^j_k = \Gamma^j_{kl}\Xi^l\,. \tag{5.75}$$

Γ^j_{kl} are the Christoffel coefficients in the ***local coframe*** given by Equation (5.71). These are the standard connection coefficients for curvilinear geometry. ▲

Exercise. Prove from their definition in formula (5.74) that the Christoffel coefficients are symmetric under the exchange of indices, $\Gamma^j_{kl} = \Gamma^j_{lk}$. ★

Definition

5.47 (Body covariant derivative)**.** *The relation in Equation (5.74)*

$$Dv_k := dv_k - v_j\widehat{\Omega}^j_k = dv_k - v_j\Gamma^j_{kl}\Xi^l \tag{5.76}$$

defines the components of the ***covariant derivative*** *of the one-form v in the body frame; that is, in the $\boldsymbol{\Xi}$-basis. That is, $\widehat{\Omega}$ is a connection form in the standard sense of differential geometry* [Fl1963, Da1994, doCa1976].

Remark

5.48 (Metric tensors)**.** *The metric tensors in the two bases of infinitesimal displacements* $d\mathbf{r}$ *and* $\boldsymbol{\Xi}$ *are related by requiring that the element of length measured in either basis must be the same. That is,*

$$ds^2 = g_{\alpha\beta}\, dr^\alpha \otimes dr^\beta = \delta_{jk}\, \Xi^j \otimes \Xi^k\,, \tag{5.77}$$

where $\otimes$ *is the symmetric tensor product. This implies a relation between the metrics,*

$$g_{\alpha\beta} = \delta_{jk}\, \Xi^j_\alpha \Xi^k_\beta\,, \tag{5.78}$$

which, in turn, implies

$$\Gamma^\nu_{\beta\mu}(r) = \frac{1}{2} g^{\nu\alpha} \left[\frac{\partial g_{\alpha\mu}(r)}{\partial r^\beta} + \frac{\partial g_{\alpha\beta}(r)}{\partial r^\mu} - \frac{\partial g_{\beta\mu}(r)}{\partial r^\alpha} \right]. \tag{5.79}$$

This equation identifies $\Gamma^\nu_{\beta\mu}(r)$ *as the Christoffel coefficients in the* ***spatial basis****. Note that the spatial Christoffel coefficients are symmetric under the exchange of indices,* $\Gamma^\nu_{\beta\mu}(r) = \Gamma^\nu_{\mu\beta}(r)$.

Definition

5.49 (Spatial covariant derivative)**.** *For the spatial metric* $g_{\alpha\mu}$*, the* ***covariant derivative*** *of the one-form* $v = v_\beta dr^\beta$ *in the spatial coordinate basis* dr^β *is defined by the standard formula, cf. Equation (5.76),*

$$Dv_\beta = dv_\beta - v_\nu \Gamma^\nu_{\alpha\beta} dr^\alpha\,,$$

or, in components,

$$\nabla_\alpha v_\beta = \partial_\alpha v_\beta - v_\nu \Gamma^\nu_{\alpha\beta}\,.$$

Remark

5.50 (What the Bichrons knew, and we found out)**.**

- *In differential geometry, the connection one-form (5.75) in the local coframe encodes the Riemannian Christoffel coefficients for the spatial coordinates, via the equivalence of metric length (5.77) as measured in either set of coordinates.*
- *The left-invariant* $\mathfrak{so}(3)$*-valued one-form* $\widehat{\Omega} = O^{-1}dO$ *that the Bichrons need for keeping track of the higher-dimensional time components of their rotations in body coordinates in (5.66) plays the same role for their two-time surfaces in* $SO(3)$ *as the connection one-form does for taking covariant derivatives in a local coframe. For more discussion of connection one-forms and their role in differential geometry, see, e.g.,* [Fl1963, Da1994].

Exercise. Write the two-time version of the Euler–Poincaré equation for a left-invariant Lagrangian defined on $\mathfrak{so}(3)$. ★

Answer. This exercise is worked out in the next section. ▲

5.4.1. *Induced Riemannian geometry.* Let's explore two more implications of the Riemannian geometry induced by the interpretation of body angular displacement $\widehat{\Omega} = O^{-1}dO$ as a connection one-form. Let's begin by expanding the first structure Equation (5.73) in the spatial coordinate basis with $\Xi^j = \Xi^j_\beta dr^\beta$ and $\widehat{\Omega}^j_k = \widehat{\Omega}^j_{k\alpha} dr^\alpha$ to find

$$d\,\Xi^j + \widehat{\Omega}^j_k \wedge \Xi^k = \Big(\partial_\alpha \Xi^j_\beta + \widehat{\Omega}^j_{k\alpha}\, \Xi^k_\beta \Big) dr^\alpha \wedge dr^\beta = 0\,. \tag{5.80}$$

This equation implies that the term in parentheses with *mixed spatial and body indices* must be symmetric in the spatial indices $(\alpha\beta)$. We may write this in suggestive notation as,

$$\partial_\alpha \Xi^j_\beta + \widehat{\Omega}^j_{k\alpha}\, \Xi^k_\beta = \Gamma^j_{\beta\alpha} = \Gamma^j_{\alpha\beta}\,. \tag{5.81}$$

To transform properly, the quantity $\Gamma^j_{\alpha\beta}$ defined in Equation (5.81) must be linearly related to both the coframe matrix Ξ^j_ν (with mixed indices) and the Christoffel coefficients in the spatial basis $\Gamma^\nu_{\alpha\beta}$ (with only spatial indices). This requires,

$$\Gamma^j_{\beta\alpha} = \Xi^j_\nu \Gamma^\nu_{\alpha\beta}\,. \tag{5.82}$$

Equations (5.81) and (5.82) define the following covariant derivative relation for the coframe matrix,

$$D_\alpha \Xi^j_\beta := \partial_\alpha \Xi^j_\beta(r) + \widehat{\Omega}^j_{k\alpha}\,\Xi^k_\beta - \Xi^j_\mu \Gamma^\mu_{\alpha\beta} = 0\,. \tag{5.83}$$

This is vanishing of the spatial covariant derivative of the coframe matrix $\Xi^j_\beta(r)$ in Equation (5.71). Taking the spatial covariant derivative of the metric relation (5.78) using (5.83) now implies

$$D_\mu g_{\alpha\beta} = D_\mu\big(\delta_{jk}\,\Xi^j_\alpha \Xi^k_\beta\big) = 0\,, \tag{5.84}$$

which is the required relation for the covariant derivative of a Riemannian metric [doCa1976, ?]. Thus, the equivalence of length (5.77) as measured by infinitesimal displacements in either space and body coordinates and the transformation property (5.82) combine to induce a Riemannian metric $g_{\alpha\beta}(r)$ in the corresponding spatial coordinates through the coframe relation (5.71) between infinitesimal displacements.

6. Hamiltonian and Lagrangian formulations of $SO(3)$-strands

Now we will begin thinking in terms of Hamiltonian partial differential equations (PDEs) 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 $\mathfrak{so}(3)$-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 for this problem will be derived via the Legendre Transformation by following calculations similar to those done previously for the rigid body in Section 4. 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.

6.1. Formulating the continuum spin chain equations.

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

$$S = \int_a^b\!\!\int_{-\infty}^{\infty} \ell(\Omega,\Xi)\,dx\,dt\,, \tag{6.1}$$

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

$$\Omega(t,x) = g^{-1}\partial_t g(t,x) \quad\text{and}\quad \Xi(t,x) = g^{-1}\partial_x g(t,x)\,, \tag{6.2}$$

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

$$
\begin{aligned}
\ell &= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{Tr}\Big(\big[g^{-1}\partial_t g, g^{-1}\partial_x g\big]^2\Big)\,dx \\
&= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{Tr}\Big(\big[\Omega,\,\Xi\big]^2\Big)\,dx\,,
\end{aligned}
\tag{6.3}
$$

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 $\mathbb{R}^3$.

6.2. **Euler–Poincaré equations.** The Euler–Poincaré procedure systematically produces the following results.

Auxiliary equations. By definition, $\Omega(t,x) = g^{-1}\partial_t g(t,x)$ and $\Xi(t,x) = g^{-1}\partial_x g(t,x)$ are Lie-algebra-valued functions over $\mathbb{R}\times\mathbb{R}$. The evolution of 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,

$$
\partial_t \Xi(t,x) - \partial_x \Omega(t,x) = \Xi\,\Omega - \Omega\,\Xi = [\Xi,\,\Omega] =: -\,\mathrm{ad}_\Omega \Xi\,. \tag{6.4}
$$

This is the auxiliary equation for $\Xi(t,x)$, cf. equation (5.70). In differential geometry, this relation is called a ***zero curvature relation***, because it implies that the curvature vanishes for the Lie-algebra-valued connection one-form $A = \Omega dt + \Xi dx$ [doCa1976].

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

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

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

$$
\frac{\partial}{\partial t}\frac{\delta\ell}{\delta\Omega} = \mathrm{ad}^*_\Omega\frac{\delta\ell}{\delta\Omega} - \frac{\partial}{\partial x}\frac{\delta\ell}{\delta\Xi} + \mathrm{ad}^*_\Xi\frac{\delta\ell}{\delta\Xi}\,. \tag{6.5}
$$

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

Exercise. Use Remark 3.13 to show that the Euler–Poincaré Equation (6.5) is a ***conservation law*** for the spin angular momentum $\Pi = \delta\ell/\delta\Omega$. That is, show

$$\frac{\partial}{\partial t}\left(\mathrm{Ad}^*_{g(t,x)^{-1}}\frac{\delta l}{\delta\Omega}\right) = -\frac{\partial}{\partial x}\left(\mathrm{Ad}^*_{g(t,x)^{-1}}\frac{\delta l}{\delta\Xi}\right). \tag{6.6}$$

★

6.3. **Hamiltonian formulation.**

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

$$h(\Pi,\Xi) = \Big\langle \Pi\,,\,\Omega\Big\rangle - \ell(\Omega,\Xi)\,. \tag{6.7}$$

Differentiating the Hamiltonian determines its partial derivatives:

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

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

$$\begin{aligned}
\partial_t\Pi &= \mathrm{ad}^*_{\delta h/\delta\Pi}\,\Pi + \partial_x\frac{\delta h}{\delta\Xi} - \mathrm{ad}^*_{\Xi}\,\frac{\delta h}{\delta\Xi}\,, \\
\partial_t\Xi &= \partial_x\frac{\delta h}{\delta\Pi} - \mathrm{ad}_{\delta h/\delta\Pi}\,\Xi\,.
\end{aligned} \tag{6.8}$$

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,

$$\frac{\partial}{\partial t}\begin{bmatrix}\Pi\\ \Xi\end{bmatrix} = \begin{bmatrix}\mathrm{ad}^*_{\Box}\Pi & \mathrm{div} - \mathrm{ad}^*_{\Xi}\\ \mathrm{grad} + \mathrm{ad}_{\Xi} & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta\Pi\\ \delta h/\delta\Xi\end{bmatrix} \tag{6.9}$$

The box $\Box$ in Equation (6.9) indicates how the ad and ad* operations are applied in the matrix multiplication. For example,

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

so each matrix entry acts on its corresponding vector component.[4]

[4] 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].

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 (6.9). 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 (6.9) 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,

$$\frac{\partial}{\partial t}\begin{bmatrix}\Pi_\alpha\\ \Xi^\alpha_i\end{bmatrix} = B_{\alpha\beta}\begin{bmatrix}\delta h/\delta\Pi_\beta\\ \delta h/\delta\Xi^\beta_j\end{bmatrix}, \tag{6.10}$$

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

$$B_{\alpha\beta} = \begin{bmatrix}-\Pi_\kappa\, t^{\kappa}_{\alpha\beta} & \delta^\beta_\alpha\partial_j + t^{\beta}_{\alpha\kappa}\Xi^\kappa_j\\ \delta^\alpha_\beta\partial_i - t^{\alpha}_{\beta\kappa}\Xi^\kappa_i & 0\end{bmatrix}. \tag{6.11}$$

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 (3.7), the Lie–Poisson Hamiltonian matrix in (6.11) may be rewritten for the $so(3)$ case in $\mathbb{R}^3$ vector form as

$$\frac{\partial}{\partial t}\begin{bmatrix}\boldsymbol{\Pi}\\ \boldsymbol{\Xi}_i\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Pi}\times & \partial_j + \boldsymbol{\Xi}_j\times\\ \partial_i + \boldsymbol{\Xi}_i\times & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta\boldsymbol{\Pi}\\ \delta h/\delta\boldsymbol{\Xi}_j\end{bmatrix}. \tag{6.12}$$

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 $\boldsymbol{\Lambda} = \delta h/\delta\boldsymbol{\Xi}$,

$$\frac{d}{dx}\boldsymbol{\Lambda} = -\,\boldsymbol{\Xi}\times\boldsymbol{\Lambda} - \boldsymbol{\Pi}\times\boldsymbol{\Omega} \quad\text{and}\quad \frac{d}{dx}\boldsymbol{\Omega} = -\,\boldsymbol{\Xi}\times\boldsymbol{\Omega}\,.$$

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 (6.5) becomes

$$\frac{\partial}{\partial t}\frac{\delta\ell}{\delta\boldsymbol{\Omega}} = -\,\boldsymbol{\Omega}\times\frac{\delta\ell}{\delta\boldsymbol{\Omega}} - \frac{\partial}{\partial x}\frac{\delta\ell}{\delta\boldsymbol{\Xi}} - \boldsymbol{\Xi}\times\frac{\delta\ell}{\delta\boldsymbol{\Xi}}\,. \tag{6.13}$$

Choices for the Lagrangian.

- Interesting choices for the Lagrangian include those symmetric under exchange of $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}$, such as

$$\ell_\perp = |\boldsymbol{\Omega}\times\boldsymbol{\Xi}|^2/2 \quad\text{and}\quad \ell_\| = (\boldsymbol{\Omega}\cdot\boldsymbol{\Xi})^2/2\,,$$

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,

$$\frac{\delta\ell_\|}{\delta\boldsymbol{\Omega}} = (\boldsymbol{\Omega}\cdot\boldsymbol{\Xi})\boldsymbol{\Xi} =: |\boldsymbol{\Xi}|^2\boldsymbol{\Omega}_\|\,,$$
$$\frac{\delta\ell_\|}{\delta\boldsymbol{\Xi}} = (\boldsymbol{\Omega}\cdot\boldsymbol{\Xi})\boldsymbol{\Omega} =: |\boldsymbol{\Omega}|^2\boldsymbol{\Xi}_\|\,,$$

for $\ell_\|$. With either of these choices, $\ell_\perp$ or $\ell_\|$, Equation (6.13) 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*** [Sk1961], a nonlinear topological model of pions in nuclear physics.

- Another interesting choice for $G = SO(3)$ and the Lagrangian $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ 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 (6.5) becomes
$$\frac{\partial}{\partial t}\mathbb{A}\boldsymbol{\Omega} + \boldsymbol{\Omega}\times\mathbb{A}\boldsymbol{\Omega} + \frac{\partial}{\partial x}\mathbb{B}\boldsymbol{\Xi} + \boldsymbol{\Xi}\times\mathbb{B}\boldsymbol{\Xi} = 0\,. \tag{6.14}$$
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 (6.14) recovers the ***chiral model***, or ***sigma model***, which is completely integrable.

Hamiltonian structures. The Hamiltonian structures of these equations on $\mathfrak{so}(3)^*$ 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 (6.5) becomes
$$\frac{\partial}{\partial t}\boldsymbol{\Pi} = \boldsymbol{\Pi}\times\frac{\delta h}{\delta\boldsymbol{\Pi}} + \frac{\partial}{\partial x}\frac{\delta h}{\delta\boldsymbol{\Xi}} + \boldsymbol{\Xi}\times\frac{\delta h}{\delta\boldsymbol{\Xi}}\,, \tag{6.15}$$
and the auxiliary Equation (6.4) becomes
$$\frac{\partial}{\partial t}\boldsymbol{\Xi} = \frac{\partial}{\partial x}\frac{\delta h}{\delta\boldsymbol{\Pi}} + \boldsymbol{\Xi}\times\frac{\delta h}{\delta\boldsymbol{\Pi}}\,, \tag{6.16}$$
which recovers the Lie–Poisson structure in Equation (6.12).

Finally, the reconstruction equations may be expressed using the hat map as
$$\partial_t O(t,x) = O(t,x)\widehat{\Omega}(t,x) \quad\text{and}\quad \partial_x O(t,x) = O(t,x)\widehat{\Xi}(t,x)\,. \tag{6.17}$$

Remark

6.1. *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. Let the set of 2×2 matrices M_i with $i=1,2,3$ satisfy the defining relation for the symplectic Lie group $Sp(2)$,
$$M_i J M_i^T = J \quad\text{with}\quad J = \begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}. \tag{6.18}$$
The corresponding elements of its Lie algebra $m_i = \dot{M}_i M_i^{-1} \in \mathfrak{sp}(2)$ 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
$$\mathsf{X}' = \dot{\mathsf{Y}} + [\mathsf{X},\mathsf{Y}]_J\,, \tag{6.19}$$
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 (6.19) on $\mathfrak{sp}(2)$. ★

Exercise. Write the Bichron equations when $SO(3)$ is replaced by $\mathrm{Diff}(\mathbb{R})$.

Hint: one could simply replace the ad- and ad*-operations in the Hamiltonian form (6.9) by their vector field equivalents and see what ensues. For details and further developments, consult [HoIvPe2012]. ★

7. More variations on the rigid body theme

7.1. $\mathbb{C}^2$ oscillators & Hopf fibration.

Example

7.1 (The 1 : 1 resonance [Ku1978]). *This example computes the momentum map for $\mathbb{C}^2 \mapsto \mathfrak{u}(2)^*$ and explains the relation of this momentum map to the Poincaré sphere and Hopf fibration.*

A unitary 2×2 matrix $U(s)$ acts on a complex two-vector $\mathbf{a}\in\mathbb{C}^2$ by matrix multiplication as

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

in which $i\xi = U'U^{-1}|_{s=0}$ is a 2×2 skew-Hermitian matrix. Therefore, the infinitesimal generator $\xi(\mathbf{a})\in\mathbb{C}^2$ may be expressed as a linear transformation,

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

in which the matrix $\xi^\dagger = \xi$ is Hermitian.

7.1.1. *Momentum map $J:\mathbb{C}^2\mapsto\mathfrak{u}(2)^*$.* The ***momentum map*** $J(\mathbf{a}) : \mathbb{C}^2 \mapsto \mathfrak{u}(2)^*$ for the matrix action of $U(2)$ on $\mathbb{C}^2$ is defined by

$$\begin{aligned} J^\xi(\mathbf{a}) &:= \big\langle J(\mathbf{a}),\,\boldsymbol{\xi}\big\rangle_{\mathfrak{u}(2)} = \frac{i}{2}\big\langle\!\big\langle \mathbf{a},\,\xi(\mathbf{a})\big\rangle\!\big\rangle_{\mathbb{C}^2} \\ &= \frac{1}{2}\omega(\mathbf{a},\xi(\mathbf{a})) \quad\text{with}\quad \xi(\mathbf{a}) = i\xi\mathbf{a}, \end{aligned} \tag{7.1}$$

and $\xi^\dagger=\xi$. The $\mathbb{C}^2$ pairing $\langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle_{\mathbb{C}^2}$ in this map is the Hermitian pairing, which for skew-Hermitian $\xi(\mathbf{a})^\dagger = -\xi(\mathbf{a})$ is also the canonical symplectic form, $\omega(\mathbf{a},\mathbf{b}) = \mathrm{Im}(\mathbf{a}^*\cdot\mathbf{b})$ on $\mathbb{C}^2$, as discussed in [MaRa1994]. Thus,

$$\begin{aligned} 2J^\xi(\mathbf{a}) &:= \omega(\mathbf{a},\xi(\mathbf{a})) = \omega(\mathbf{a}, i\xi\mathbf{a}) \\ &= \mathrm{Im}(a_k^*(i\xi)_{kl}a_l) \\ &= a_k^*\xi_{kl}a_l \\ &= \mathrm{tr}\big((\mathbf{a}\otimes\mathbf{a}^*)\,\xi\big) \\ &= \mathrm{tr}\big(J^\dagger(\mathbf{a}^*,\mathbf{a})\,\xi\big)\,. \end{aligned} \tag{7.2}$$

Consequently, the momentum map $J:\mathbb{C}^2\mapsto\mathfrak{u}(2)^*$ is given by the Hermitian expression

$$J(\mathbf{a}) = \frac{1}{2}\mathbf{a}\otimes\mathbf{a}^*. \tag{7.3}$$

This conclusion may be checked by computing the differential of the Hamiltonian $dJ^\xi(\mathbf{a})$ for the momentum map, which should be canonically related to its Hamiltonian vector field $X_{J^\xi(\mathbf{a})} = \{\,\cdot\,, J^\xi(\mathbf{a})\}$.

As the infinitesimal generator $\xi(\mathbf{a}) = i\xi\mathbf{a}$ is linear, we have

$$
\begin{aligned}
dJ^\xi(\mathbf{a}) &= d\big\langle J(\mathbf{a}), \boldsymbol{\xi} \big\rangle_{u(2)} = \frac{i}{2}\big\langle\!\big\langle \mathbf{a}, \xi(d\mathbf{a}) \big\rangle\!\big\rangle_{\mathbb{C}^2} + \frac{i}{2}\big\langle\!\big\langle\, d\mathbf{a}, \xi(\mathbf{a}) \big\rangle\!\big\rangle_{\mathbb{C}^2} \\
&= \Im\big\langle\!\big\langle \xi(\mathbf{a}),\, d\mathbf{a}\, \big\rangle\!\big\rangle_{\mathbb{C}^2} = \omega(\xi(\mathbf{a}), \,\cdot\,) = X_{J^\xi(\mathbf{a})} \,\lrcorner\, \omega ,
\end{aligned}
$$

which is the desired canonical relation.

7.1.2. *The Poincaré sphere $S^2 \subset S^3$.* We expand the Hermitian matrix $J = \frac{1}{2}\mathbf{a}\otimes\mathbf{a}^*$ in (7.3) in a basis of four 2×2 unit Hermitian matrices $(\sigma_0\,, \boldsymbol{\sigma})$, with $\boldsymbol{\sigma} = (\sigma_1,\, \sigma_2,\, \sigma_3)$ given by

$$
\begin{aligned}
\sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\
\sigma_2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.
\end{aligned} \tag{7.4}
$$

The result is the decomposition

$$
J = \frac{1}{4}\big(R\sigma_0 + \mathbf{Y}\cdot\boldsymbol{\sigma} \big)\,. \tag{7.5}
$$

Here we denote $R := \mathrm{tr}(J\,\sigma_0) = |a_1|^2 + |a_2|^2$ and

$$
\mathbf{Y} = \mathrm{tr}(J\,\boldsymbol{\sigma}) = a_k^*\boldsymbol{\sigma}_{kl}a_l\,, \tag{7.6}
$$

with vector notation $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$. In components, one finds

$$
J = \frac{1}{2}\begin{bmatrix} a_1^*a_1 & a_1^*a_2 \\ a_2^*a_1 & a_2^*a_2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} R + Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & R - Y_3 \end{bmatrix}, \tag{7.7}
$$

with trace $\mathrm{tr}\, J = R$. Thus, the decomposition (7.5) splits the momentum map into its trace part $R \in \mathbb{R}$ and its traceless part $\mathbf{Y} \in \mathbb{R}^3$, given by

$$
\mathbf{Y} = J - \frac{1}{2}(\mathrm{tr}\, J)\,\mathrm{Id} \in \mathfrak{su}(2)^* \cong \mathbb{R}^3\,. \tag{7.8}
$$

This formula is the $SU(2)$ momentum map for the Poincaré sphere.

Definition

7.2 (Poincaré sphere)**.** *The coefficients $R \in \mathbb{R}$ and $\mathbf{Y} \in \mathbb{R}^3$ in the expansion of the matrix J in (7.5) comprise the four real quadratic quantities,*

$$
\begin{aligned}
R &= \frac{1}{2}\big(|a_1|^2 + |a_2|^2 \big)\,, \\
Y_3 &= \frac{1}{2}\big(|a_1|^2 - |a_2|^2 \big) \quad \textit{and} \\
Y_1 - i\,Y_2 &= a_1^* a_2\,.
\end{aligned} \tag{7.9}
$$

These quantities are all invariant under the action $\mathbf{a} \to e^{i\phi}\mathbf{a}$ of $\phi \in S^1$ on $\mathbf{a} \in \mathbb{C}^2$. The S^1-invariant coefficients in the expansion of the momentum map $J = \mathbf{a}\otimes\mathbf{a}^*$ (7.3) in the basis of sigma matrices (7.4) satisfy the relation

$$
4\,\det\, J = R^2 - |\mathbf{Y}|^2 = 0\,, \quad \textit{with} \quad |\mathbf{Y}|^2 \equiv Y_1^2 + Y_2^2 + Y_3^2\,. \tag{7.10}
$$

*This relation defines the **Poincaré sphere** $S^2 \in S^3$ of radius R which, in turn, is related to the Hopf fibration $\mathbb{C}^2/S^1 \simeq S^3$. For more information about the Poincaré sphere and the Hopf fibration, consult, e.g.,* [Ho2011] *and references therein.*

7.1.3. *The $U(2)$ Lie group structure.* The Lie group $U(2) = S^1 \times SU(2)$ is the direct product of its centre,

$$
Z(U(2)) = \{ zI \text{ with } |z| = 1 \} \equiv S^1,
$$

and the special unitary group in two dimensions,

$$
SU(2) = \left\{ \begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix} \text{ with } |\alpha|^2 + |\beta|^2 = 1 \right\}.
$$

As a consequence, the momentum map $J(\mathbf{a}) = \frac{1}{2}\mathbf{a}\otimes\mathbf{a}^*$ in (7.3) for the action $U(2)\times\mathbb{C}^2\to\mathbb{C}^2$ decomposes into two momentum maps obtained by separating J into its trace part $J_{S^1} = R\in\mathbb{R}$ and its traceless part $\mathbf{J}_{SU(2)} = \mathbf{Y}\in\mathbb{R}^3$. This decomposition may be sketched, as follows.

$$
\begin{tikzcd}
 & \mathbb{C}^2 \arrow[dl, "\mathbf{J}_{SU(2)} = \mathbf{Y}"'] \arrow[dr, "J_{S^1} = R"] & \\
\mathfrak{su}(2)^* & & \mathbb{R}
\end{tikzcd}
$$

The target spaces $\mathfrak{su}(2)^*$ and $\mathbb{R}$ of the left and right legs of this pair of momentum maps are each Poisson manifolds, with coordinates $\mathbf{Y}\in\mathfrak{su}(2)^*$ and $R\in\mathbb{R}$, respectively. The corresponding Poisson brackets are given in tabular form as

(7.11)

$\{\cdot\,,\,\cdot\}$	Y_1	Y_2	Y_3	R
Y_1	0	Y_3	$-Y_2$	0
Y_2	$-Y_3$	0	Y_1	0
Y_3	Y_2	$-Y_1$	0	0
R	0	0	0	0

.

In index notation, these Poisson brackets are given as

$$
\{Y_k\,,\,Y_l\} = \epsilon_{klm}Y_m \quad\text{and}\quad \{Y_k\,,\,R\} = 0\,. \tag{7.12}
$$

The last Poisson bracket relation means that the spaces with coordinates $\mathbf{Y}\in\mathfrak{su}(2)^*$ and $R\in\mathbb{R}$ are *symplectically orthogonal* in $\mathfrak{u}(2)^* = \mathfrak{su}(2)^*\times\mathbb{R}$.

Equations (7.12) prove the following.

Theorem

7.3 (Momentum map (7.3) is Poisson). *The direct-product structure of $U(2) = S^1\times SU(2)$ decomposes the momentum map J in Equation (7.3) into two other momentum maps, $J_{SU(2)}:\mathbb{C}^2\mapsto\mathfrak{su}(2)^*$ and $J_{S^1}:\mathbb{C}^2\mapsto\mathbb{R}$. These other momentum maps are also Poisson maps. That is, they each satisfy the Poisson property for smooth functions F and H,*

$$
\big\{F\circ J\,,\,H\circ J\big\} = \big\{F\,,\,H\big\}\circ J\,. \tag{7.13}
$$

This relation defines a Lie–Poisson bracket on $\mathfrak{su}(2)^*$ that inherits the defining properties of a Poisson bracket from the canonical relations

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

for the canonical symplectic form, $\omega = \Im\,(da_j\wedge da_j^*)$.

Remark

7.4. *The Poisson bracket table in (7.11) is the $\mathfrak{so}(3)^*$ Lie–Poisson bracket table for angular momentum in the* spatial *frame. It differs by an overall sign from the $\mathfrak{so}(3)^*$ Lie–Poisson bracket table for angular momentum in the* body *frame, see (4.11).*

Definition

7.5 (Dual pairs). *Let (M,ω) be a symplectic manifold and let P_1, P_2 be two Poisson manifolds. A pair of Poisson mappings*

$$
P_1 \xleftarrow{J_1} (M,\omega) \xrightarrow{J_2} P_2
$$

*is called a **dual pair** [We1983b] if $\ker TJ_1$ and $\ker TJ_2$ are symplectic orthogonal complements of one another, that is,*

$$
(\ker TJ_1)^\omega = \ker TJ_2. \tag{7.14}
$$

A systematic treatment of dual pairs can be found in Chapter 11 of [OrRa2004]. *The infinite-dimensional case is treated in* [GaVi2010].

Remark

7.6 (Summary). In the pair of momentum maps

$$\mathfrak{su}(2)^* \equiv \mathbb{R}^3 \xleftarrow{\mathbf{Y}} (\mathbb{C}^2, \omega) \xrightarrow{R} \mathbb{R}, \tag{7.15}$$

$\mathbf{Y}$ *maps the fibres of R, which are three-spheres, into two-spheres, that are coadjoint orbits of $SU(2)$. The restriction of* $\mathbf{Y}$ *to these three-spheres is the Hopf fibration. Further pursuit of the theory of dual pairs is beyond our present scope. See* [HoVi2010] *for a recent discussion of dual pairs for resonant oscillators from the present viewpoint.*

Exercise. For $\mathbf{a} \in \mathbb{C}^3$ one may write the 3×3 Hermitian matrix $Q = \mathbf{a}\otimes\mathbf{a}^*$ as the sum $Q = S + iA$ of a 3×3 real symmetric matrix S plus i times a 3×3 real antisymmetric matrix A:

$$\begin{aligned} Q &= \begin{bmatrix} M_1 & N_3 - iL_3 & N_2 + iL_2 \\ N_3 + iL_3 & M_2 & N_1 - iL_1 \\ N_2 - iL_2 & N_1 + iL_1 & M_3 \end{bmatrix} \\ &= \begin{bmatrix} M_1 & N_3 & N_2 \\ N_3 & M_2 & N_1 \\ N_2 & N_1 & M_3 \end{bmatrix} + i \begin{bmatrix} 0 & -L_3 & L_2 \\ L_3 & 0 & -L_1 \\ -L_2 & L_1 & 0 \end{bmatrix}. \end{aligned}$$

(i) Compute the Poisson brackets of the L's, M's and N's among themselves, given that $\{a_j,\, a_k^*\} = -2i\delta_{jk}$ for $j,k = 1,2,3$.

(ii) Transform into a rotating frame in which the real symmetric part of Q is diagonal. Write the Hamiltonian equations for the L's, M's and N's in that rotating frame for a rotationally invariant Hamiltonian.

★

7.2. $\mathbb{C}^3$ **oscillators. Answer. (Oscillator variables in three dimensions)** The nine elements of Q are the S^1-invariants

$$Q_{jk} = a_j a_k^* = S_{jk} + iA_{jk}\,, \quad j,k = 1,2,3.$$

The Poisson brackets among these variables are evaluated from the canonical relation,

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

by using the Leibniz property (product rule) for Poisson brackets to find

$$\{Q_{jk}\,,\, Q_{lm}\} \quad = \quad 2i\,(\delta_{kl}Q_{jm} - \delta_{jm}Q_{kl})\,, \quad j,k,l,m = 1,2,3\,.$$

Remark

7.7. *The quadratic S^1-invariant quantities in $\mathbb{C}^3$ Poisson commute among themselves. This property of* ***closure*** *is to be expected for a simple reason. The Poisson bracket between two homogeneous polynomials of weights w_1 and w_2 produces a homogeneous polynomial of weight $w = w_1 + w_2 - 2$ and $2+2-2 = 2$; so the quadratic homogeneous polynomials Poisson-commute among themselves.*

The result is also a simple example of Poisson reduction by symmetry, obtained by ***transforming to quantities that are invariant*** *under the action of a Lie group. The action in this case is the (diagonal) S^1 phase shift $a_j \to a_j e^{i\phi}$ for $j = 1,2,3$.*

(i) One defines $L_a := -\frac{1}{2}\,\epsilon_{ajk}A_{jk} = (\mathbf{p}\times\mathbf{q})_a$ and finds the Poisson bracket relations,

$$\begin{aligned} \{L_a\,,\, L_b\} &= A_{ab} - A_{ba} = \epsilon_{abc}L_c\,, \\ \{L_a\,,\, Q_{jk}\} &= \frac{1}{2}\left[\epsilon_{ajc}Q_{ck} - \epsilon_{akc}Q_{jc}\right]. \end{aligned}$$

Thus, perhaps not unexpectedly, the Poisson bracket for quadratic S^1-invariant quantities in $\mathbb{C}^3$ contains the angular momentum Poisson bracket among the variables L_a with $a = 1, 2, 3$. This could be expected, because the 3×3 form of Q contains the 2×2 form, which we know admits the Hopf fibration into quantities which satisfy Poisson bracket relations dual to the Lie algebra $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$. Moreover, the imaginary part $\operatorname{Im} Q = \mathbf{L} \cdot \widehat{\mathbf{J}} = L_a \widehat{J}_a$, where $\widehat{J}_a$ with $a = 1, 2, 3$ is a basis set for $\mathfrak{so}(3)$ as represented by the 3×3 skew-symmetric real matrices.

Another interesting set of Poisson bracket relations among the M's, N's and L's may be found. These relations are

$$\begin{aligned}
\{N_a - iL_a\,,\, N_b - iL_b\} &= 2i\,\epsilon_{abc}(N_c + iL_c)\,,\\
\{M_a\,,\, M_b\} &= 0\,,\\
\{M_a\,,\, N_b - iL_b\} &= 2i\,\mathrm{sgn}(b-a)(-1)^{a+b}\\
&\qquad\qquad (N_b - iL_b)\,,
\end{aligned}$$

where $\mathrm{sgn}(b-a)$ is the sign of the difference $(b-a)$, which vanishes when $b = a$.

Additional Poisson bracket relations may also be read off from the Poisson commutators of the real and imaginary components of $Q = S + iA$ among themselves as

$$\begin{aligned}
\{S_{jk}\,,\, S_{lm}\} &= \delta_{jl}A_{mk} + \delta_{kl}A_{mj} - \delta_{jm}A_{kl} - \delta_{km}A_{jl}\,,\\
\{S_{jk}\,,\, A_{lm}\} &= \delta_{jl}S_{mk} + \delta_{kl}S_{mj} - \delta_{jm}S_{kl} - \delta_{km}S_{jl}\,,\\
\{A_{jk}\,,\, A_{lm}\} &= \delta_{jl}A_{mk} - \delta_{kl}A_{mj} + \delta_{jm}A_{kl} - \delta_{km}A_{jl}\,.
\end{aligned} \tag{7.16}$$

These relations produce the following five tables of Poisson brackets in addition to $\{M_a\,,\, M_b\} = 0$:

$\{\cdot\,,\,\cdot\}$	L_1	L_2	L_3
L_1	0	L_3	$-L_2$
L_2	$-L_3$	0	L_1
L_3	L_2	$-L_1$	0

$\{\cdot\,,\,\cdot\}$	N_1	N_2	N_3
N_1	0	$-L_3$	L_2
N_2	L_3	0	$-L_1$
N_3	$-L_2$	L_1	0

$\{\cdot\,,\,\cdot\}$	L_1	L_2	L_3
M_1	0	$2N_2$	$-2N_3$
M_2	$-2N_1$	0	$2N_3$
M_3	$2N_1$	$-2N_2$	0

$\{\cdot\,,\,\cdot\}$	N_1	N_2	N_3
M_1	0	$-2L_2$	$2L_3$
M_2	$2L_1$	0	$-2L_3$
M_3	$-2L_1$	$2L_2$	0

$\{\cdot\,,\,\cdot\}$	L_1	L_2	L_3
N_1	$M_2 - M_3$	$-N_3$	N_2
N_2	N_3	$M_3 - M_1$	$-N_1$
N_3	$-N_2$	N_1	$M_1 - M_2$

As expected, the system is closed and it has the angular momentum Poisson bracket table as a closed subset. Next, we will come to understand that this is because the Lie algebra $\mathfrak{su}(2)$ is a subalgebra of $su(3)$.

(ii) The rotation group $SO(3)$ is a subgroup of $SU(3)$. An element $Q \in su(3)^*$ transforms under $SO(3)$ by the coAdjoint action

$$\mathrm{Ad}^*_R\, Q = R^{-1}QR = R^{-1}SR + iR^{-1}AR\,.$$

Choose $R \in SO(3)$ so that $R^{-1}SR = D = \mathrm{diag}(d_1, d_2, d_3)$ is diagonal. (That is, rotate into principal axis coordinates for S.) The eigenvalues are unique up to their order, which one may fix as, say, $d_1 \geq d_2 \geq d_3$. While it diagonalises the symmetric part of Q, the rotation R takes the antisymmetric part from the spatial frame to the body frame, where S is diagonal. At the same time the spatial angular momentum matrix A is transformed to $B = R^{-1}AR$, which is the body angular momentum. Thus,

$$\mathrm{Ad}^*_R\, Q = R^{-1}SR + iR^{-1}AR =: D + iB\,.$$

Define the body angular velocity $\Omega = R^{-1}\dot{R} \in \mathfrak{so}(3)$, which is left-invariant. The Hamiltonian dynamical system obeys

$$\dot{Q} = \{Q,\, H(Q)\}\,.$$

For $B = R^{-1}AR$, this implies

$$\dot{B} + [\Omega,\, B] = R^{-1}\dot{A}R = R^{-1}\{A,\, H(Q)\}R\,.$$

However, $H(Q)$ being rotationally symmetric means the spatial angular momentum A will be time-independent $\dot{A} = \{A,\, H(Q)\} = 0$. Hence,

$$\dot{B} + [\Omega,\, B] = 0\,.$$

Thus, the equation for the body angular momentum B is formally identical to Euler's equations for rigid-body motion. Physically, this represents conservation of spatial angular momentum, because of the rotational symmetry of the Hamiltonian.

Likewise, for $D = R^{-1}SR$, one finds

$$\dot{D} + [\Omega,\, D] = R^{-1}\dot{S}R = R^{-1}\{S,\, H(Q)\}R \neq 0\,.$$

The body angular momentum B satisfies Euler's rigid-body equations, but this body is not rigid! While the rotational degrees of freedom satisfy spatial angular momentum conservation, the shape of the body depends on the value of the Poisson bracket $R^{-1}\{S,\, H(Q)\}R$ which is likely to be highly nontrivial! For example, the Hamiltonian $H(Q)$ may be chosen to be a function of the following three rotationally invariant quantities:

$$\begin{aligned}
\mathrm{tr}(A^TA) &= \mathrm{tr}(B^TB)\,,\\
\mathrm{tr}(A^TSA) &= \mathrm{tr}(B^TDB)\,,\\
\mathrm{tr}(A^TS^2A) &= \mathrm{tr}(B^TD^2B)\,.
\end{aligned}$$

Dependence of the Hamiltonian on these quantities will bring the complications of the Poisson bracket relations in (7.16) into the dynamics of the triaxial ellipsoidal shape represented by D.

▲

Remark

7.8. *The quantity*

$$\widetilde{Q} = \mathbf{a}\otimes\mathbf{a}^* - \frac{1}{3}\,\mathrm{Id}|\mathbf{a}|^2 \,:\, \mathbb{C}^3 \mapsto su(3)^*$$

corresponds for the action of $SU(3)$ *on* $\mathbb{C}^3$ *to the momentum map* $J:\,\mathbb{C}^2\mapsto\mathfrak{su}(2)^*$ *in Example 7.1.1 for the action of* $SU(2)$ *on* $\mathbb{C}^2$.

7.3. **Motion on the symplectic Lie group** $Sp(2)$**.** Let the set of 2×2 matrices M_i with $i = 1,2,3$ satisfy the defining relation for the symplectic Lie group $Sp(2)$,

$$M_iJM_i^T = J \quad\text{with}\quad J = \begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}.$$

The corresponding elements of its Lie algebra $m_i = \dot{M}_iM_i^{-1} \in \mathfrak{sp}(2)$ 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. For definiteness, we may choose a basis given by

$$\mathsf{X}_1 = Jm_1 = \begin{pmatrix} 2 & 0\\ 0 & 0\end{pmatrix},\ \mathsf{X}_2 = Jm_2 = \begin{pmatrix} 0 & 0\\ 0 & 2\end{pmatrix},\ \mathsf{X}_3 = Jm_3 = \begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}.$$

This basis corresponds to the momentum map $\mathbb{R}^6 \to \mathfrak{sp}(2)^*$ of quadratic phase-space functions $\mathbf{X} = (|\mathbf{q}|^2, |\mathbf{p}|^2, \mathbf{q}\cdot\mathbf{p})^T$. One sees this by using the symmetric matrices $\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3$ above to compute the following three quadratic forms defined using $\mathbf{z} = (\mathbf{q},\mathbf{p})^T$:

$$\frac{1}{2}\mathbf{z}^T\mathsf{X}_1\mathbf{z} = |\mathbf{q}|^2 = X_1,\ \frac{1}{2}\mathbf{z}^T\mathsf{X}_2\mathbf{z} = |\mathbf{p}|^2 = X_2,\ \frac{1}{2}\mathbf{z}^T\mathsf{X}_3\mathbf{z} = \mathbf{q}\cdot\mathbf{p} = X_3\,.$$

Exercise. (The Lie bracket) For $\mathsf{X} = Jm$ and $\mathsf{Y} = Jn \in sym(2)$ with $m, n \in \mathfrak{sp}(2)$, prove

$$[\mathsf{X}, \mathsf{Y}]_J := \mathsf{X}J\mathsf{Y} - \mathsf{Y}J\mathsf{X} = -J(mn - nm) = -J[m, n].$$

Use this equality to prove the Jacobi identity for the J-bracket $[\mathsf{X}, \mathsf{Y}]_J$. ★

Answer. The first part is a straightforward calculation using $J^2 = -\mathrm{Id}_{2\times 2}$ with the definitions of X and Y. The second part follows from the Jacobi identity for the symplectic Lie algebra and linearity in the definitions of $\mathsf{X}, \mathsf{Y} \in sym(2)$ in terms of $m, n \in \mathfrak{sp}(2)$. ▲

Exercise. (A variational identity) If $\mathsf{X} = J\dot{M}M^{-1}$ for derivative $\dot{M} = \partial M(s,\sigma)/\partial s|_{\sigma=0}$ and $\mathsf{Y} = JM'M^{-1}$ for variational derivative $\delta M = M' = \partial M(s,\sigma)/\partial\sigma|_{\sigma=0}$, show that equality of cross derivatives in s and σ implies the relation

$$\delta \mathsf{X} = \mathsf{X}' = \dot{\mathsf{Y}} + [\mathsf{X}, \mathsf{Y}]_J\,.$$

★

Answer. This relation follows from an important standard calculation in geometric mechanics, performed earlier in deriving Equation (3.8). It begins by computing the time derivative of $MM^{-1} = Id$ along the curve $M(s)$ to find $(MM^{-1})^{\cdot} = 0$, so that

$$(M^{-1})^{\cdot} = -\,M^{-1}\dot{M}M^{-1}\,.$$

Next, one defines $m = \dot{M}M^{-1}$ and $n = M'M^{-1}$. Then the previous relation yields

$$\begin{aligned} m' &= \dot{M}'M^{-1} - \dot{M}M^{-1}M'M^{-1} \\ \dot{n} &= \dot{M}'M^{-1} - M'M^{-1}\dot{M}M^{-1} \end{aligned}$$

so that subtraction yields the relation

$$m' - \dot{n} = nm - mn =: -[m, n].$$

Then, upon substituting the definitions of X and Y, one finds

$$\begin{aligned} \mathsf{X}' = Jm' &= J\dot{n} - J[m, n] \\ &= \dot{\mathsf{Y}} + [\mathsf{X}, \mathsf{Y}]_J = \dot{\mathsf{Y}} + 2\mathrm{sym}(\mathsf{X}J\mathsf{Y}). \end{aligned}$$

▲

Exercise. (Hamilton's principle for $\mathfrak{sp}(2)$) Use the previous relation to compute the Euler–Poincaré equation for evolution resulting from Hamilton's principle,

$$0 = \delta S = \delta \int \ell(\mathsf{X}(s))\,ds = \int \mathrm{tr}\left(\frac{\partial \ell}{\partial \mathsf{X}}\,\delta\mathsf{X}\right) ds.$$

★

Answer. Integrate by parts and rearrange as follows:

$$\begin{aligned} 0 = \delta S &= \int \mathrm{tr}\left(\frac{\partial \ell}{\partial \mathsf{X}}\,\mathsf{X}'\right) ds \\ &= \int \mathrm{tr}\left(\frac{\partial \ell}{\partial \mathsf{X}}\,(\dot{\mathsf{Y}} - \mathsf{Y}J\mathsf{X} + \mathsf{X}J\mathsf{Y})\right) ds \\ &= \int \mathrm{tr}\left(\left(-\,\frac{d}{ds}\frac{\partial \ell}{\partial \mathsf{X}} - J\mathsf{X}\frac{\partial \ell}{\partial \mathsf{X}} + \frac{\partial \ell}{\partial \mathsf{X}}\mathsf{X}J\right)\mathsf{Y}\right) ds \\ &= \int \mathrm{tr}\left(\left(-\,\frac{d}{ds}\frac{\partial \ell}{\partial \mathsf{X}} - 2\mathrm{sym}\left(J\mathsf{X}\frac{\partial \ell}{\partial \mathsf{X}}\right)\right)\mathsf{Y}\right) ds\,, \end{aligned}$$

upon setting the boundary term $\mathrm{tr}(\frac{\partial \ell}{\partial \mathsf{X}} Y)|_{s_0}^{s_1}$ equal to zero. This results in the Euler–Poincaré equation,

$$\frac{d}{ds}\frac{\partial \ell}{\partial \mathsf{X}} = -2\mathrm{sym}\Big(J\mathsf{X}\frac{\partial \ell}{\partial \mathsf{X}}\Big) = 2\mathrm{sym}\Big(\frac{\partial \ell}{\partial \mathsf{X}}\mathsf{X}J\Big)\,. \tag{7.17}$$

▲

Exercise. (Geodesic motion on $\mathfrak{sp}(2)^*$) Specialise this evolution equation to the case that $\ell(\mathsf{X}) = \frac{1}{2}\mathrm{tr}(\mathsf{X}^2)$, where tr denotes the trace of a matrix. (This is geodesic motion on the matrix Lie group $Sp(2)$ with respect to the trace norm of matrices.)
★

Answer. When $\ell(\mathsf{X}) = \frac{1}{2}\mathrm{tr}(\mathsf{X}^2)$ we have $\partial\ell/\partial\mathsf{X} = \mathsf{X}$, so the Euler–Poincaré Equation (7.17) becomes

$$\dot{\mathsf{X}} = -2\mathrm{sym}\big(J\mathsf{X}^2\big) = \mathsf{X}^2 J - J\mathsf{X}^2 = [\mathsf{X}^2,\, J]\,. \tag{7.18}$$

This is called a ***Bloch–Iserles equation*** [BlIs2006]. ▲

Exercise. (Lie–Poisson Hamiltonian formulation) Write the Hamiltonian form of the Euler–Poincaré equation on $SP(2)$ and identify the associated Lie–Poisson bracket.
★

Answer. The Hamiltonian form of the Euler–Poincaré Equation (7.17) is found from the Legendre transform via the dual relations

$$\mu = \frac{\partial \ell}{\partial \mathsf{X}} \quad \text{and} \quad \mathsf{X} = \frac{\partial h}{\partial \mu} \quad \text{with} \quad h(\mu) = \mathrm{tr}(\mu\mathsf{X}) - \ell(\mathsf{X})\,.$$

Thus,

$$\dot{\mu} = -2\mathrm{sym}\Big(J\frac{\partial h}{\partial \mu}\mu\Big) = -J\frac{\partial h}{\partial \mu}\mu + \mu\frac{\partial h}{\partial \mu}J\,.$$

The Lie–Poisson bracket is obtained from

$$\begin{aligned}
\frac{d}{ds}f(\mu) &= \mathrm{tr}\Big(\frac{\partial f}{\partial \mu}\frac{d\mu}{ds}\Big) \\
&= -2\mathrm{tr}\Big(\mu\,\mathrm{sym}\Big(\frac{\partial f}{\partial \mu}J\frac{\partial h}{\partial \mu}\Big)\Big) \\
&= -\mathrm{tr}\Big(\mu\Big[\frac{\partial f}{\partial \mu}\,,\,\frac{\partial h}{\partial \mu}\Big]_J\Big) \\
&=: \Big\{f\,,h\Big\}_J\,.
\end{aligned}$$

The Jacobi identity for this Lie–Poisson bracket follows from that of the J-bracket discussed earlier.

The geodesic Bloch–Iserles Equation (7.18) is recovered when the Hamiltonian is chosen as $h = \frac{1}{2}\mathrm{tr}(\mu^2)$ and one sets $\mu \to \mathsf{X}$. ▲

Exercise. (A second Bloch–Iserles Poisson bracket) Show that the geodesic Bloch–Iserles Equation (7.18) may also be written in Hamiltonian form with Hamiltonian $h = \frac{1}{3}\mathrm{tr}(\mu^3)$. ★

Answer. Equation (7.18) may also be written as

$$\dot{\mu} = -2\mathrm{sym}\Big(J\frac{\partial h}{\partial \mu}\mu\Big) = -J\frac{\partial h}{\partial \mu} + \frac{\partial h}{\partial \mu}J$$

with Hamiltonian $h = \frac{1}{3}\mathrm{tr}(\mu^3)$. The corresponding Poisson bracket has constant coefficients,

$$\begin{aligned}\frac{d}{ds}f(\mu) &= \mathrm{tr}\Big(\frac{\partial f}{\partial \mu}\frac{d\mu}{ds}\Big)\\ &= -2\mathrm{tr}\Big(\mathrm{sym}\Big(\frac{\partial f}{\partial \mu}J\frac{\partial h}{\partial \mu}\Big)\Big)\\ &= -\mathrm{tr}\Big(\Big[\frac{\partial f}{\partial \mu}\,,\frac{\partial h}{\partial \mu}\Big]_J\Big)\\ &=: \Big\{f\,,h\Big\}_{J2}.\end{aligned}$$

▲

Exercise. (A parallel with the rigid body) The geodesic Bloch–Iserles Equation (7.18) may be written in a form reminiscent of the rigid body, as

$$\frac{d}{dt}X = [X,\,\Omega] \quad \text{with} \quad \Omega = JX + XJ = -\Omega^T.$$

This suggests the Manakov form

$$\frac{d}{dt}(X + \lambda J) = [X + \lambda J,\, JX + XJ + \lambda^2 J^2].$$

This seems dual to the Manakov form (3.39) for the rigid body, because the symmetric and antisymmetric matrices exchange roles.

Verify these equations and explain what the Manakov form means in determining the conservation laws for this problem. ★

Exercise. (The Bloch–Iserles G-strand) Write the two-time version of the Euler–Poincaré equation for a left-invariant Lagrangian defined on $\mathfrak{sp}(2)$. ★

7.4. **Two coupled rigid bodies.** In the centre of mass frame, the Lagrangian for the problem of two coupled rigid bodies may be written as depending only on the angular velocities of the two bodies $\Omega_1 = A_1^{-1}\dot A_1(t)$, $\Omega_2 = A_2^{-1}\dot A_2(t)$ and the relative angle $A = A_1^{-1}A_2$ between the bodies [GrKrMa1988],

$$l(\Omega_1,\Omega_2,A) : \mathfrak{so}(3)\times\mathfrak{so}(3)\times SO(3)\to\mathbb{R},$$

which we write as

$$l(\Omega_1,\Omega_2,A) = \frac12\begin{pmatrix}\Omega_1\\ \Omega_2\end{pmatrix}^T\cdot M(A)\begin{pmatrix}\Omega_1\\ \Omega_2\end{pmatrix},$$

where $M(A)$ is a 6×6 block matrix containing both A and the two inertia tensors of the bodies.

Upon identifying $\mathbb{R}_3$ with $\mathfrak{so}(3) = T_eSO(3)$ by the hat map, this Lagrangian becomes

$$l = l(\widehat{\boldsymbol{\Omega}}_1,\widehat{\boldsymbol{\Omega}}_2,A)$$

and we may identify $SO(3)$ with its dual $SO^*(3)$ through the matrix pairing $SO(3)\times SO^*(3)\to\mathbb{R}$.

The Lagrangian is then a function

$$l : \mathfrak{so}(3)\times\mathfrak{so}(3)\times SO^*(3)\to\mathbb{R}$$

which may be written as

$$l(\Omega,A) = \frac12\big\langle M(A)\,\Omega,\,\Omega\big\rangle =: \frac12\big\langle \Pi,\Omega\big\rangle\,,$$

where a nondegenerate matrix trace pairing is defined in components by

$$\big\langle \Pi,\Omega\big\rangle := \mathrm{Tr}\left[\begin{pmatrix}\widehat{\boldsymbol{\Omega}}_1\\ \widehat{\boldsymbol{\Omega}}_2\end{pmatrix}^T\cdot\begin{pmatrix}\widetilde{\boldsymbol{\Pi}}_1\\ \widetilde{\boldsymbol{\Pi}}_2\end{pmatrix}\right]$$

for all $\Omega = (\widehat{\boldsymbol{\Omega}}_1,\widehat{\boldsymbol{\Omega}}_2)\in\mathfrak{so}(3)\times\mathfrak{so}(3)$, $\Pi = (\widetilde{\boldsymbol{\Pi}}_1,\widetilde{\boldsymbol{\Pi}}_2)\in\mathfrak{so}^*(3)\times\mathfrak{so}^*(3)$.

The Euler–Poincaré theory has been developed to treat Lagrangians of the form

$$l : \mathfrak{g} \times V^* \to \mathbb{R}$$

where V is a vector space on which the Lie algebra acts.

Exercise. Formulate the Euler–Poincaré equations for the problem of two coupled rigid bodies. ★

Answer. The direct-product Lie algebra

$$\mathfrak{g} = \mathfrak{so}(3) \times \mathfrak{so}(3)$$

is endowed with the product Lie bracket

$$\mathrm{ad}_\Omega \Xi = \big[\Omega, \Xi\big] = \left[\begin{pmatrix} \widehat{\mathbf{\Omega}}_1 \\ \widehat{\mathbf{\Omega}}_2 \end{pmatrix}, \begin{pmatrix} \widehat{\mathbf{\Xi}}_1 \\ \widehat{\mathbf{\Xi}}_2 \end{pmatrix} \right] = \begin{pmatrix} [\![\widehat{\mathbf{\Omega}}_1, \widehat{\mathbf{\Xi}}_1]\!] \\ [\![\widehat{\mathbf{\Omega}}_2, \widehat{\mathbf{\Xi}}_2]\!] \end{pmatrix} \tag{7.19}$$

where $[\![\cdot\,, \cdot]\!]$ indicates the standard $\mathfrak{so}(3)$ matrix commutator.

Formulating the Euler–Poincaré theorem for this problem will require a Lie algebra action of $\mathfrak{so}(3) \times \mathfrak{so}(3)$ on $SO^*(3)$, which fortunately is readily available. Indeed, from the definitions of the two body angular velocities and relative angle $A = A_1^{-1} A_2$, one finds

$$\frac{dA}{dt} = -\widehat{\mathbf{\Omega}}_1 A + A \widehat{\mathbf{\Omega}}_2\,, \tag{7.20}$$

which is the Lie algebra action we seek, abbreviated as

$$\frac{dA}{dt} = -\,\Omega(A). \tag{7.21}$$

The Euler–Poincaré variational principle is then $\delta S = 0$, for

$$\begin{aligned} \delta \int_{t_0}^{t_1} l(\Omega, A)\, dt &= \int_{t_0}^{t_1} \left\langle \frac{\delta l}{\delta \Omega},\, \delta\Omega \right\rangle + \left\langle \frac{\delta l}{\delta A},\, \delta A \right\rangle dt \\ &= \int_{t_0}^{t_1} \left\langle \frac{\delta l}{\delta \Omega},\, \frac{d\Xi}{dt} + \mathrm{ad}_\Omega \Xi \right\rangle + \left\langle \frac{\delta l}{\delta A},\, -\,\Xi(A) \right\rangle dt \\ &= \int_{t_0}^{t_1} \left\langle -\,\frac{d}{dt}\frac{\delta l}{\delta \Omega} + \mathrm{ad}^*_\Omega \frac{\delta l}{\delta \Omega} + \frac{\delta l}{\delta A} \diamond A,\, \Xi \right\rangle dt \end{aligned}$$

with $\delta\Omega = \dot{\Xi} + \mathrm{ad}_\Omega \Xi$ and $\delta A = -\,\Xi(A)$. As a result, the (left-invariant) Euler–Poincaré equations may be written as

$$\frac{d}{dt}\frac{\delta l}{\delta \Omega} = \mathrm{ad}^*_\Omega \frac{\delta l}{\delta \Omega} + \frac{\delta l}{\delta A} \diamond A\,. \tag{7.22}$$

This, of course, is the general form of the Euler–Poincaré equations with advected quantities.

The Euler–Poincaré equations for the present problem of coupled rigid bodies will take their final form, once we have computed the diamond operation $(\,\diamond\,)$,

$$\diamond : SO(3) \times SO(3)^* \to \mathfrak{so}(3)^*\,. \tag{7.23}$$

The Lie algebra action (7.21) yields the following definition of diamond for our case,

$$\begin{aligned} \left\langle \frac{\delta l}{\delta A} \diamond A,\, \Xi \right\rangle &:= -\left\langle A,\, \Xi\Big(\frac{\delta l}{\delta A}\Big) \right\rangle \\ &= -\left\langle A,\, \Big(\widehat{\mathbf{\Xi}}_1 \frac{\delta l}{\delta A} - \frac{\delta l}{\delta A}\widehat{\mathbf{\Xi}}_2\Big) \right\rangle \\ &= -\left\langle A\,\frac{\delta l}{\delta A},\, \widehat{\mathbf{\Xi}}_1 \right\rangle + \left\langle \frac{\delta l}{\delta A}\,A,\, \widehat{\mathbf{\Xi}}_2 \right\rangle, \end{aligned}$$

where the last step is justified by the cyclic property of the trace. Consequently, the components of the diamond operation are given by

$$\frac{\delta l}{\delta A} \diamond A = \Big(-A\frac{\delta l}{\delta A},\, \frac{\delta l}{\delta A} A \Big)$$

and substituting them into the general form of the Euler–Poincaré equations in (7.22) gives the equations of motion of our problem:

$$\frac{d}{dt}\widetilde{\mathbf{\Pi}}_1 = \mathrm{ad}^*_{\widehat{\mathbf{\Omega}}_1}\widetilde{\mathbf{\Pi}}_1 - A\frac{\delta l}{\delta A}\,,$$
$$\frac{d}{dt}\widetilde{\mathbf{\Pi}}_2 = \mathrm{ad}^*_{\widehat{\mathbf{\Omega}}_2}\widetilde{\mathbf{\Pi}}_2 + \frac{\delta l}{\delta A}A\,.$$

These along with the auxiliary Equation (7.20) comprise the Euler–Poincaré form of the Lie–Poisson equations that are derived for the motion of two coupled rigid bodies in [GrKrMa1988]. The corresponding Lie–Poisson equations in [GrKrMa1988] may be derived from the Euler–Poincaré equations here by applying a symmetry-reduced Legendre transform. ▲

8. Symmetry breaking by potential energy: the heavy top

8.1. **Heavy top: Introduction and definitions.** A top is a rigid body of mass m rotating with a fixed point of support in a constant gravitational field of acceleration $-g\hat{\mathbf{z}}$ pointing vertically downward. The orientation of the body relative to the vertical axis $\hat{\mathbf{z}}$ is defined by the unit vector $\mathbf{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$ for a curve $\mathbf{R}(t) \in SO(3)$. According to its definition, the unit vector $\mathbf{\Gamma}$ represents the motion of the vertical direction as seen from the rotating body. Consequently, it satisfies the auxiliary motion equation,

$$\dot{\mathbf{\Gamma}} = -\,\mathbf{R}^{-1}\dot{\mathbf{R}}(t)\mathbf{\Gamma} = -\,\widehat{\mathbf{\Omega}}(t)\mathbf{\Gamma} = \mathbf{\Gamma}\times\mathbf{\Omega}\,. \tag{8.1}$$

Here the rotation matrix $\mathbf{R}(t) \in SO(3)$, the skew matrix $\widehat{\mathbf{\Omega}} = \mathbf{R}^{-1}\dot{\mathbf{R}} \in so(3)$ and the body angular frequency vector $\mathbf{\Omega} \in \mathbb{R}^3$ are related by the hat map, $\mathbf{\Omega} = \big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)\widehat{\ }$, where

$$\text{hat map,}\quad \widehat{\ }: (so(3), [\cdot,\cdot]) \to (\mathbb{R}^3, \times)\,,$$

with $\widehat{\mathbf{\Omega}}\mathbf{v} = \mathbf{\Omega}\times\mathbf{v}$ for any $\mathbf{v} \in \mathbb{R}^3$.

The motion of a top is determined from Euler's equations in vector form,

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega}\times\mathbf{\Omega} + mg\,\mathbf{\Gamma}\times\boldsymbol{\chi}\,, \tag{8.2}$$
$$\dot{\mathbf{\Gamma}} = \mathbf{\Gamma}\times\mathbf{\Omega}\,, \tag{8.3}$$

where $\mathbf{\Omega}$, $\mathbf{\Gamma}$, $\boldsymbol{\chi} \in \mathbb{R}^3$ are vectors in the rotating body frame. Here

- $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is the body angular velocity vector.
- $\mathbb{I} = \mathrm{diag}(I_1, I_2, I_3)$ is the moment of inertia tensor, diagonalised in the body principal axes.
- $\mathbf{\Gamma} = R^{-1}(t)\hat{\mathbf{z}}$ represents the motion of the unit vector along the vertical axis, as seen from the body.
- $\boldsymbol{\chi}$ is the constant vector in the body from the point of support to the body's centre of mass.
- m is the total mass of the body and g is the constant acceleration of gravity.

8.2. **Heavy-top action principle.**

Proposition

8.1. *The heavy-top motion equation (8.2) is equivalent to the* ***heavy-top action principle*** $\delta S_{\mathrm{red}} = 0$ *for a* ***reduced action****,*

$$S_{\mathrm{red}} = \int_a^b l(\mathbf{\Omega}, \mathbf{\Gamma})\,dt = \int_a^b \frac{1}{2}\Big\langle \mathbb{I}\mathbf{\Omega}\,,\,\mathbf{\Omega}\Big\rangle - \Big\langle mg\,\boldsymbol{\chi}\,,\,\mathbf{\Gamma}\Big\rangle\,dt\,, \tag{8.4}$$

where variations of vectors $\mathbf{\Omega}$ *and* $\mathbf{\Gamma}$ *are restricted to be of the form*

$$\delta\mathbf{\Omega} = \dot{\mathbf{\Sigma}} + \mathbf{\Omega}\times\mathbf{\Sigma} \quad \textit{and} \quad \delta\mathbf{\Gamma} = \mathbf{\Gamma}\times\mathbf{\Sigma}\,, \tag{8.5}$$

arising from variations of the corresponding definitions $\widehat{\mathbf{\Omega}} = \mathbf{R}^{-1}\dot{\mathbf{R}}$ *and* $\mathbf{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$ *in which* $\widehat{\mathbf{\Sigma}}(t) = \mathbf{R}^{-1}\delta\mathbf{R}$ *is a curve in* $\mathbb{R}^3$ *that vanishes at the endpoints in time.*

Proof. Since $\mathbb{I}$ is symmetric and $\boldsymbol{\chi}$ is constant, one finds the variation,

$$\begin{aligned}
\delta \int_a^b l(\boldsymbol{\Omega},\boldsymbol{\Gamma})\,dt &= \int_a^b \Big\langle\, \mathbb{I}\boldsymbol{\Omega}\,,\,\delta\boldsymbol{\Omega}\Big\rangle - \Big\langle mg\,\boldsymbol{\chi}\,,\,\delta\boldsymbol{\Gamma}\Big\rangle\,dt \\
&= \int_a^b \Big\langle\, \mathbb{I}\boldsymbol{\Omega}\,,\,\dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega}\times\boldsymbol{\Sigma}\Big\rangle - \Big\langle mg\,\boldsymbol{\chi}\,,\,\boldsymbol{\Gamma}\times\boldsymbol{\Sigma}\Big\rangle\,dt \\
&= \int_a^b \Big\langle -\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega}\,,\,\boldsymbol{\Sigma}\Big\rangle + \Big\langle\, \mathbb{I}\boldsymbol{\Omega}\,,\,\boldsymbol{\Omega}\times\boldsymbol{\Sigma}\Big\rangle - \Big\langle mg\,\boldsymbol{\chi}\,,\,\boldsymbol{\Gamma}\times\boldsymbol{\Sigma}\Big\rangle dt \\
&= \int_a^b \Big\langle -\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega} + \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega} + mg\,\boldsymbol{\Gamma}\times\boldsymbol{\chi}\,,\,\boldsymbol{\Sigma}\Big\rangle\,dt,
\end{aligned}$$

upon integrating by parts and using the endpoint conditions, $\boldsymbol{\Sigma}(b) = \boldsymbol{\Sigma}(a) = 0$. Since $\boldsymbol{\Sigma}$ is otherwise arbitrary, (8.4) is equivalent to

$$-\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega} + \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega} + mg\,\boldsymbol{\Gamma}\times\boldsymbol{\chi} = 0\,,$$

which is Euler's motion equation for the heavy top (8.2). This motion equation is completed by the auxiliary equation $\dot{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma}\times\boldsymbol{\Omega}$ in (8.3) arising from the definition of $\boldsymbol{\Gamma}$. $\square$

The Legendre transformation for $l(\boldsymbol{\Omega},\boldsymbol{\Gamma})$ gives the body angular momentum

$$\boldsymbol{\Pi} = \frac{\partial l}{\partial\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega}\,.$$

The well-known energy Hamiltonian for the heavy top then emerges as

$$h(\boldsymbol{\Pi},\boldsymbol{\Gamma}) = \boldsymbol{\Pi}\cdot\boldsymbol{\Omega} - l(\boldsymbol{\Omega},\boldsymbol{\Gamma}) = \frac{1}{2}\langle\boldsymbol{\Pi}\,,\,\mathbb{I}^{-1}\boldsymbol{\Pi}\rangle + \langle mg\,\boldsymbol{\chi}\,,\,\boldsymbol{\Gamma}\,\rangle\,, \tag{8.6}$$

which is the sum of the kinetic and potential energies of the top.

8.3. **Lie–Poisson brackets.**

Definition

8.2. *Let $f,h : \mathfrak{g}^* \to \mathbb{R}$ be real-valued functions on the dual space $\mathfrak{g}^*$. Denoting elements of $\mathfrak{g}^*$ by μ, the functional derivative of f at μ is defined as the unique element $\delta f/\delta\mu$ of $\mathfrak{g}$ defined by*

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[f(\mu+\varepsilon\delta\mu) - f(\mu)] = \left\langle \delta\mu, \frac{\delta f}{\delta\mu}\right\rangle, \tag{8.7}$$

for all $\delta\mu \in \mathfrak{g}^*$, where $\langle\cdot\,,\,\cdot\rangle$ denotes the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$.

Definition

8.3 (Lie–Poisson equations)**.** *The* $(\pm)$ ***Lie–Poisson brackets*** *are defined by*

$$\{f,h\}_{\pm}(\mu) = \pm\left\langle \mu, \left[\frac{\delta f}{\delta\mu}, \frac{\delta h}{\delta\mu}\right]\right\rangle = \mp\left\langle \mu, \mathrm{ad}_{\delta h/\delta\mu}\frac{\delta f}{\delta\mu}\right\rangle. \tag{8.8}$$

The corresponding ***Lie–Poisson equations****, determined by $\dot f = \{f,h\}$, read*

$$\dot\mu = \{\mu,h\} = \mp\,\mathrm{ad}^*_{\delta h/\delta\mu}\,\mu\,, \tag{8.9}$$

where one defines the ad^* operation in terms of the pairing $\langle\cdot\,,\,\cdot\rangle$, by

$$\{f,h\} = \left\langle \mu, \mathrm{ad}_{\delta h/\delta\mu}\frac{\delta f}{\delta\mu}\right\rangle = \left\langle \mathrm{ad}^*_{\delta h/\delta\mu}\,\mu, \frac{\delta f}{\delta\mu}\right\rangle.$$

Remark

8.4. *The Lie–Poisson setting of mechanics is a special case of the general theory of systems on Poisson manifolds, for which there is now extensive theoretical development. (See* [MaRa1994] *for a start on this literature.)*

8.4. **Lie–Poisson brackets and momentum maps.** An important feature of the rigid-body bracket carries over to general Lie algebras. Namely, *Lie–Poisson brackets on* $\mathfrak{g}^*$ *arise from canonical brackets on the cotangent bundle (phase space)* T^*G associated with a Lie group G which has $\mathfrak{g}$ as its associated Lie algebra. Thus, the process by which the Lie–Poisson brackets arise is the momentum map

$$T^*G \mapsto \mathfrak{g}^* \,.$$

For example, a rigid body is free to rotate about its centre of mass and G is the (proper) rotation group $SO(3)$. The choice of T^*G as the primitive phase space is made according to the classical procedures of mechanics described earlier. For the description using Lagrangian mechanics, one forms the velocity phase space TG. The Hamiltonian description on T^*G is then obtained by standard procedures: Legendre transforms, etc.

The passage from T^*G to the space of $\mathbf{\Pi}$'s (body angular momentum space) is determined by *left* translation on the group. This mapping is an example of a ***momentum map***; that is, a mapping whose components are the "Noether quantities" associated with a symmetry group. That the map from T^*G to $\mathfrak{g}^*$ is a Poisson map *is a general fact about momentum maps*. The Hamiltonian point of view of all this is a standard subject reviewed, for example, in [MaRa1994].

Remark

8.5 (Lie–Poisson description of the heavy top)**.**
As it turns out, the underlying Lie algebra for the Lie–Poisson description of the heavy top consists of the Lie algebra $se(3,\mathbb{R})$ *of infinitesimal Euclidean motions in* $\mathbb{R}^3$*. This is a bit surprising, because heavy-top motion itself does* not *actually arise through spatial translations by the Euclidean group; in fact, the body has a fixed point! Instead, the Lie algebra* $se(3,\mathbb{R})$ *arises for another reason associated with the breaking of the* $SO(3)$ *isotropy by the presence of the gravitational field. This symmetry breaking introduces a semidirect-product Lie–Poisson structure which happens to coincide with the dual of the Lie algebra* $se(3,\mathbb{R})$ *in the case of the heavy top.*

8.5. **Lie–Poisson brackets for the heavy top.** The Lie algebra of the special Euclidean group in three dimensions is $se(3)=\mathbb{R}^3\times\mathbb{R}^3$ with the Lie bracket

$$[(\boldsymbol{\xi},\mathbf{u}),(\boldsymbol{\eta},\mathbf{v})]=(\boldsymbol{\xi}\times\boldsymbol{\eta},\boldsymbol{\xi}\times\mathbf{v}-\boldsymbol{\eta}\times\mathbf{u})\,. \tag{8.10}$$

We identify the dual space with pairs $(\mathbf{\Pi},\mathbf{\Gamma})$; the corresponding $(-)$ Lie–Poisson bracket called the ***heavy-top bracket*** is

$$\{f\,,\,h\}(\mathbf{\Pi},\mathbf{\Gamma})=-\,\mathbf{\Pi}\cdot\frac{\partial f}{\partial\mathbf{\Pi}}\times\frac{\partial h}{\partial\mathbf{\Pi}}-\mathbf{\Gamma}\cdot\Big(\frac{\partial f}{\partial\mathbf{\Pi}}\times\frac{\partial h}{\partial\mathbf{\Gamma}}-\frac{\partial h}{\partial\mathbf{\Pi}}\times\frac{\partial f}{\partial\mathbf{\Gamma}}\Big)\,.$$

This Lie–Poisson bracket and the Hamiltonian (8.6) recover Equations (8.2) and (8.3) for the heavy top, as

$$\begin{aligned}\dot{\mathbf{\Pi}}=\{\mathbf{\Pi}\,,\,h\}\quad&=\quad\mathbf{\Pi}\times\frac{\partial h}{\partial\mathbf{\Pi}}+\mathbf{\Gamma}\times\frac{\partial h}{\partial\mathbf{\Gamma}}\\&=\quad\mathbf{\Pi}\times\mathbb{I}^{-1}\mathbf{\Pi}+\mathbf{\Gamma}\times mg\,\boldsymbol{\chi}\,,\\[1em]\dot{\mathbf{\Gamma}}=\{\mathbf{\Gamma}\,,\,h\}\quad&=\quad\mathbf{\Gamma}\times\frac{\partial h}{\partial\mathbf{\Pi}}=\mathbf{\Gamma}\times\mathbb{I}^{-1}\mathbf{\Pi}\,.\end{aligned}$$

Remark

8.6 (Semidirect products and symmetry breaking)**.** *The Lie algebra of the Euclidean group has a structure which is a special case of what is called a* ***semidirect product****. Here, it is the semidirect-product action* $so(3)\,\circledS\,\mathbb{R}^3$ *of the Lie algebra of rotations* $so(3)$ *acting on the infinitesimal translations* $\mathbb{R}^3$*, which happens to coincide with* $se(3,\mathbb{R})$*.*

In general, the Lie bracket for semidirect-product action $\mathfrak{g}\,\circledS\,V$ *of a Lie algebra* $\mathfrak{g}$ *on a vector space* V *is given by*

$$\Big[(X,a),(\overline{X},\overline{a})\Big]=\Big([X,\overline{X}\,],\overline{X}(a)-X(\overline{a})\Big),$$

in which $X, \overline{X} \in \mathfrak{g}$ and $a, \overline{a} \in V$. Here, the action of the Lie algebra on the vector space is denoted, e.g., $X(\overline{a})$. Usually, this action would be the Lie derivative.

Lie–Poisson brackets defined on the dual spaces of semidirect-product Lie algebras tend to occur under rather general circumstances when the symmetry in T^*G is broken, e.g., reduced to an isotropy subgroup of a set of parameters. In particular, there are similarities in structure between the Poisson bracket for compressible flow and that for the heavy top. In the latter case, the vertical direction of gravity breaks the isotropy of $\mathbb{R}^3$ from $SO(3)$ to $SO(2)$. The general theory for semidirect products is reviewed in a variety of places, including [MaRaWe1984a, MaRaWe1984b].

Many interesting examples of Lie–Poisson brackets on semidirect products exist for fluid dynamics. These semidirect-product Lie–Poisson Hamiltonian theories range from simple fluids, to charged fluid plasmas, to magnetised fluids, to multiphase fluids, to super fluids, to Yang–Mills fluids, relativistic or not, and to liquid crystals. Many of these theories are discussed from the Euler–Poincaré viewpoint in [HoMaRa1998] and [Ho2002].

8.6. **Heavy top: Clebsch action principle.**

Proposition

8.7 (Clebsch heavy-top action principle)**.** *The heavy-top Equations (8.2) and (8.3) follow from a Clebsch constrained action principle, $\delta S = 0$, with*

$$S = \int_a^b \frac{1}{2}\Big\langle \mathbb{I}\boldsymbol{\Omega}\,,\,\boldsymbol{\Omega}\Big\rangle - \Big\langle mg\,\boldsymbol{\chi}\,,\,\boldsymbol{\Gamma}\Big\rangle + \Big\langle \boldsymbol{\Xi}\,,\,\dot{\boldsymbol{\Gamma}} + \boldsymbol{\Omega}\times\boldsymbol{\Gamma}\Big\rangle\,dt\,. \tag{8.11}$$

Remark

8.8. *The last term in this action is the **Clebsch constraint** for the auxiliary equation satisfied by the unit vector $\boldsymbol{\Gamma}$. From its definition $\boldsymbol{\Gamma} = \mathbf{R}^{-1}(t)\hat{\mathbf{z}}$ and the definition of the body angular velocity $\boldsymbol{\Omega} = \mathbf{R}^{-1}(t)\dot{\mathbf{R}}$, this unit vector must satisfy*

$$\dot{\boldsymbol{\Gamma}} = -\,\mathbf{R}^{-1}\dot{\mathbf{R}}(t)\boldsymbol{\Gamma} = -\,\widehat{\boldsymbol{\Omega}}(t)\boldsymbol{\Gamma} = -\,\boldsymbol{\Omega}\times\boldsymbol{\Gamma}\,.$$

(The third equality invokes the hat map.) According to the Clebsch construction, the Lagrange multiplier $\boldsymbol{\Xi}$ enforcing the auxiliary Equation (8.11) will become the momentum canonically conjugate to the auxiliary variable $\boldsymbol{\Gamma}$.

Proof. The stationary variations of the constrained action (8.11) yield the following three ***Clebsch relations***, cf. Equations (4.21) for the rigid body,

$$\begin{aligned} \delta\boldsymbol{\Omega}: &\quad \mathbb{I}\boldsymbol{\Omega} + \boldsymbol{\Gamma}\times\boldsymbol{\Xi} = 0\,,\\ \delta\boldsymbol{\Xi}: &\quad \dot{\boldsymbol{\Gamma}} + \boldsymbol{\Omega}\times\boldsymbol{\Gamma} = 0\,,\\ \delta\boldsymbol{\Gamma}: &\quad \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega}\times\boldsymbol{\Xi} + mg\,\boldsymbol{\chi} = 0\,. \end{aligned}$$

The first Clebsch relation defines the momentum map $T^*\mathbb{R}^3 \to so(3)^*$ for the body angular momentum $\mathbb{I}\boldsymbol{\Omega}$. From the other two Clebsch relations, the equation of motion for the body angular momentum may be computed as

$$\begin{aligned} \mathbb{I}\dot{\boldsymbol{\Omega}} &= -\,\dot{\boldsymbol{\Gamma}}\times\boldsymbol{\Xi} - \boldsymbol{\Gamma}\times\dot{\boldsymbol{\Xi}}\\ &= (\boldsymbol{\Omega}\times\boldsymbol{\Gamma})\times\boldsymbol{\Xi} + \boldsymbol{\Gamma}\times(\boldsymbol{\Omega}\times\boldsymbol{\Xi} + mg\,\boldsymbol{\chi})\\ &= \boldsymbol{\Omega}\times(\boldsymbol{\Gamma}\times\boldsymbol{\Xi}) + \boldsymbol{\Gamma}\times mg\,\boldsymbol{\chi}\\ &= -\,\boldsymbol{\Omega}\times(\mathbb{I}\boldsymbol{\Omega}) + mg\,\boldsymbol{\Gamma}\times\boldsymbol{\chi}\,, \end{aligned}$$

which recovers Euler's motion Equation (8.2) for the heavy top. $\square$

8.7. **Heavy top: Kaluza–Klein construction.** The Lagrangian in the heavy-top action principle (8.4) may be transformed into quadratic form. This is accomplished by suspending the system in a higher-dimensional space via the ***Kaluza–Klein construction***. This construction proceeds for the heavy top as a slight modification of the well-known Kaluza–Klein construction for a charged particle in a prescribed magnetic field.

Let Q_{KK} be the manifold $SO(3)\times\mathbb{R}^3$ with variables $(\mathbf{R},\mathbf{q})$. On Q_{KK} introduce the ***Kaluza–Klein Lagrangian***

$$L_{KK}: TQ_{KK}\simeq TSO(3)\times T\mathbb{R}^3\mapsto\mathbb{R}\,,$$

as

$$\begin{aligned} L_{KK}(\mathbf{R},\mathbf{q},\dot{\mathbf{R}},\dot{\mathbf{q}};\hat{\mathbf{z}}) &= L_{KK}(\boldsymbol{\Omega},\boldsymbol{\Gamma},\mathbf{q},\dot{\mathbf{q}}) \\ &= \frac{1}{2}\langle\,\mathbb{I}\boldsymbol{\Omega}\,,\,\boldsymbol{\Omega}\,\rangle+\frac{1}{2}|\boldsymbol{\Gamma}+\dot{\mathbf{q}}|^2\,, \end{aligned}\tag{8.12}$$

with $\boldsymbol{\Omega}=\big(\mathbf{R}^{-1}\dot{\mathbf{R}}\big)^{\widehat{}}$ and $\boldsymbol{\Gamma}=\mathbf{R}^{-1}\hat{\mathbf{z}}$. The Lagrangian L_{KK} is positive-definite in $(\boldsymbol{\Omega},\boldsymbol{\Gamma},\dot{\mathbf{q}})$; so it may be regarded as a kinetic energy which defines a metric, the ***Kaluza–Klein metric*** on TQ_{KK}.

The Legendre transformation for L_{KK} gives the momenta

$$\boldsymbol{\Pi}=\mathbb{I}\boldsymbol{\Omega}\qquad\text{and}\qquad\mathbf{p}=\boldsymbol{\Gamma}+\dot{\mathbf{q}}\,.\tag{8.13}$$

Since L_{KK} does not depend on $\mathbf{q}$, the Euler–Lagrange equation

$$\frac{d}{dt}\frac{\partial L_{KK}}{\partial\dot{\mathbf{q}}}=\frac{\partial L_{KK}}{\partial\mathbf{q}}=0$$

shows that $\mathbf{p}=\partial L_{KK}/\partial\dot{\mathbf{q}}$ is conserved. The ***constant vector*** $\mathbf{p}$ is now identified as the vector in the body,

$$\mathbf{p}=\boldsymbol{\Gamma}+\dot{\mathbf{q}}=-\,mg\,\boldsymbol{\chi}\,.$$

After this identification, the heavy-top action principle in Proposition 8.1 with the Kaluza–Klein Lagrangian returns Euler's motion equation for the heavy top (8.2).

The Hamiltonian H_{KK} associated with L_{KK} by the Legendre transformation (8.13) is

$$\begin{aligned} H_{KK}(\boldsymbol{\Pi},\boldsymbol{\Gamma},\mathbf{q},\mathbf{p}) &= \boldsymbol{\Pi}\cdot\boldsymbol{\Omega}+\mathbf{p}\cdot\dot{\mathbf{q}}-L_{KK}(\boldsymbol{\Omega},\boldsymbol{\Gamma},\mathbf{q},\dot{\mathbf{q}}) \\ &= \frac{1}{2}\boldsymbol{\Pi}\cdot\mathbb{I}^{-1}\boldsymbol{\Pi}-\mathbf{p}\cdot\boldsymbol{\Gamma}+\frac{1}{2}|\mathbf{p}|^2 \\ &= \frac{1}{2}\boldsymbol{\Pi}\cdot\mathbb{I}^{-1}\boldsymbol{\Pi}+\frac{1}{2}|\mathbf{p}-\boldsymbol{\Gamma}|^2-\frac{1}{2}|\boldsymbol{\Gamma}|^2\,. \end{aligned}$$

Recall that $\boldsymbol{\Gamma}$ is a unit vector. On the constant level set $|\boldsymbol{\Gamma}|^2=1$, the Kaluza–Klein Hamiltonian H_{KK} is a positive quadratic function, shifted by a constant. Likewise, on the constant level set $\mathbf{p}=-\,mg\,\boldsymbol{\chi}$, the Kaluza–Klein Hamiltonian H_{KK} is a function of only the variables $(\boldsymbol{\Pi},\boldsymbol{\Gamma})$ and is equal to the Hamiltonian (8.6) for the heavy top up to an additive constant. As a result we have the following.

Proposition

8.9. *The Lie–Poisson equations for the Kaluza–Klein Hamiltonian H_{KK} recover Euler's equations for the heavy top, (8.2) and (8.3).*

Proof. The Lie–Poisson bracket may be written in matrix form explicitly as

$$\{\,f\,,\,h\,\}=\begin{bmatrix}\partial f/\partial\boldsymbol{\Pi}\\ \partial f/\partial\boldsymbol{\Gamma}\\ \partial f/\partial\mathbf{q}\\ \partial f/\partial\mathbf{p}\end{bmatrix}^T\begin{bmatrix}\boldsymbol{\Pi}\times & \boldsymbol{\Gamma}\times & 0 & 0\\ \boldsymbol{\Gamma}\times & 0 & 0 & 0\\ 0 & 0 & 0 & Id\\ 0 & 0 & -\,Id & 0\end{bmatrix}\begin{bmatrix}\partial h/\partial\boldsymbol{\Pi}\\ \partial h/\partial\boldsymbol{\Gamma}\\ \partial h/\partial\mathbf{q}\\ \partial h/\partial\mathbf{p}\end{bmatrix}.\tag{8.14}$$

Consequently, one obtains the following Hamiltonian equations for $h=H_{KK}(\boldsymbol{\Pi},\boldsymbol{\Gamma},\mathbf{q},\mathbf{p})$,

$$\begin{bmatrix}\dot{\boldsymbol{\Pi}}\\ \dot{\boldsymbol{\Gamma}}\\ \dot{\mathbf{q}}\\ \dot{\mathbf{p}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Pi}\times & \boldsymbol{\Gamma}\times & 0 & 0\\ \boldsymbol{\Gamma}\times & 0 & 0 & 0\\ 0 & 0 & 0 & Id\\ 0 & 0 & -\,Id & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}\\ -\,\mathbf{p}\\ 0\\ \mathbf{p}-\boldsymbol{\Gamma}\end{bmatrix}.\tag{8.15}$$

These recover the heavy-top Equations (8.2) and (8.3) upon evaluating $\mathbf{p} = -mg\boldsymbol{\chi}$. □

Exercise. In an attempt to mimic Manakov's beautiful idea for showing the integrability of the rigid body on $SO(n)$, one might imagine writing the three-dimensional heavy-top Equations (8.2) and (8.3) by inserting a spectral parameter λ as

$$\frac{d}{dt}\big(\boldsymbol{\Gamma} + \lambda\boldsymbol{\Pi} + \lambda^2\boldsymbol{J}\big) = \big(\boldsymbol{\Gamma} + \lambda\boldsymbol{\Pi} + \lambda^2\boldsymbol{J}\big) \times \big(\boldsymbol{\Omega} + \lambda\boldsymbol{K}\big)\,,$$

with constant vectors $\boldsymbol{J}$ and $\boldsymbol{K}$ in $\mathbb{R}^3$. Does this formulation provide enough constants of motion to show the integrability of the heavy-top equations for some values of $\boldsymbol{\chi}$ and $\mathbb{I}$? If so, which types of tops may be shown to be integrable this way? ★

Answer. The polynomial equation above implies the following relations, for powers of λ:

$$\begin{aligned}
&\lambda^3: \ \boldsymbol{J}\times\boldsymbol{K} = \mathbf{0} \implies \boldsymbol{J} \parallel \boldsymbol{K}, \implies \boldsymbol{J} = \alpha\boldsymbol{K}, \ \alpha = const.\\
&\lambda^2: \ \dot{\boldsymbol{J}} = \mathbf{0} = \boldsymbol{\Pi}\times\boldsymbol{K} + \boldsymbol{J}\times\boldsymbol{\Omega}, \implies (\mathbb{I}\boldsymbol{\Omega} - \alpha\boldsymbol{\Omega})\times\boldsymbol{K} = \mathbf{0}.\\
&\lambda^1: \ \dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi}\times\boldsymbol{\Omega} + \boldsymbol{\Gamma}\times\boldsymbol{K}, \implies \boldsymbol{K} = mg\boldsymbol{\chi}\,.\\
&\lambda^0: \ \dot{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma}\times\boldsymbol{\Omega}.
\end{aligned}$$

These relationships hold, provided the moment of inertia $\mathbb{I}$ is either proportional to the identity (Euler top), or has two equal entries that make it cylindrically symmetric about the vector $\boldsymbol{\chi}$ (Lagrange top).

This system conserves each of the coefficients of the powers of λ in $|\boldsymbol{\Gamma} + \lambda\boldsymbol{\Pi} + \lambda^2\boldsymbol{J}|^2$. That is, besides the kinematic constant $|\boldsymbol{J}|^2$, it conserves

$$|\boldsymbol{\Gamma}|^2, \quad \boldsymbol{\Gamma}\cdot\boldsymbol{\Pi}\,, \quad \frac{1}{2\alpha}|\boldsymbol{\Pi}|^2 + mg\boldsymbol{\Gamma}\cdot\boldsymbol{\chi}\,, \quad \boldsymbol{\Pi}\cdot\boldsymbol{\chi}\,.$$

The first two are the Casimirs of the Lie–Poisson bracket in (8.11), the third is the Hamiltonian and the last is the $\boldsymbol{\chi}$-component of the angular momentum, which is conserved when the moment of inertia $\mathbb{I}$ is cylindrically symmetric about the vector $\boldsymbol{\chi}$.

Cylindrical symmetry holds for the Euler top and the Lagrange top, which are indeed known to be integrable. For in-depth discussions of this approach to heavy-top dynamics, see [Ra1982, RaVM1982]. ▲

Exercise. Manakov's approach for the heavy top in the vector notation of the previous exercise suggests a similar application to the $n\times n$ matrix commutator equation

$$\frac{d}{dt}\big(\Gamma + \lambda\Pi + \lambda^2 J\big) = \big[\Gamma + \lambda\Pi + \lambda^2 J,\, \Omega + \lambda K\big]$$

with skew-symmetric $(\Gamma, \Pi, \Omega, J, K)$ with constant (J, K). Determine whether this approach could be used to extend Manakov's treatment of the rigid body in n dimensions to the n-dimensional versions of the Euler top and the Lagrange top. ★

Exercise. Extend the Manakov approach even further by computing the system of $n\times n$ matrix equations for

$$\frac{d}{dt}\big(\Gamma + \lambda M + \lambda^2 N + \lambda^3 J\big) = \big[\Gamma + \lambda M + \lambda^2 N + \lambda^3 J,\, \Omega + \lambda\omega + \lambda^2 K\big]$$

Is this extended matrix system Hamiltonian? If so, what is its Lie–Poisson bracket?
★

9. Euler–Poincaré reduction for continua

As discussed in Theorem 2.3, Euler–Poincaré reduction starts with a G-invariant Lagrangian

$$L : TG \to \mathbb{R}$$

defined on the tangent bundle of a Lie group G.

Definition

9.1. *A Lagrangian $L : TG \to \mathbb{R}$ is said to be right G-invariant if $L(TR_h(v)) = L(v)$, for all $v \in T_gG$ and for all $g, h \in G$. In shorter notation,* ***right invariance*** *of the Lagrangian may be written as*

$$L(g(t)h, \dot{g}(t)h) = L(g(t), \dot{g}(t))\,,$$

for all $h \in G$.

For a G-invariant Lagrangian defined on TG, reduction by symmetry takes Hamilton's principle from TG to $TG/G \simeq \mathfrak{g}$. Stationarity of the symmetry-reduced Hamilton's principle yields the Euler–Poincaré equations on $\mathfrak{g}^*$ discussed in Section 2. As we have seen, the corresponding reduced Legendre transformation yields the now-standard ***Lie–Poisson bracket*** for the Hamiltonian formulation of these equations.

Theorem

9.2 (Euler–Poincaré reduction)**.**
Let G be a Lie group and $L : TG \to \mathbb{R}$ be a ***right invariant Lagrangian****. Let $\ell := L|_{\mathfrak{g}} : \mathfrak{g} \to \mathbb{R}$ be its restriction to $\mathfrak{g}$. For a curve $g(t) \in G$, let*

$$u(t) = \dot{g}(t)\cdot g(t)^{-1} := T_{g(t)}R_{g(t)^{-1}}\dot{g}(t) \in \mathfrak{g}\,.$$

Then the following four statements are equivalent:

(i) *$g(t)$ satisfies the* ***Euler–Lagrange equations*** *for Lagrangian L defined on G.*

(ii) *The variational principle*

$$\delta \int_a^b L(g(t), \dot{g}(t))\mathrm{d}t = 0\,, \tag{9.1}$$

holds, for variations with fixed endpoints.

(iii) *The (right invariant)* ***Euler–Poincaré equations*** *hold:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta \ell}{\delta u} = -\,\mathrm{ad}^*_u \frac{\delta \ell}{\delta u}\,. \tag{9.2}$$

(iv) *The variational principle*

$$\delta \int_a^b \ell(u(t))\,\mathrm{d}t = 0\,, \tag{9.3}$$

holds on $\mathfrak{g}$, using variations of the form

$$\delta u = \dot{v} + [u, v]\,, \tag{9.4}$$

where $u(t)$ is an arbitrary path in $\mathfrak{g}$ that vanishes at the endpoints, i.e. $u(a) = u(b) = 0$.

We identify the Lie group G with the smooth invertible maps with smooth inverses; that is, we identify G with $\mathrm{Diff}(\mathcal{D})$ the group of diffeomorphisms acting on the domain $\mathcal{D}$. We will forego any analytical technicalities that may arise in making this identification. These issues for the diffeomorphism group remain an active field of current research.

The adjoint action of $\mathfrak{X}(\mathcal{D})$ on itself is

$$\mathrm{ad}_u v = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\Phi_u(t))_* v = -\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\Phi_u(t))^* v = -\mathcal{L}_u v = -[u, v], \tag{9.5}$$

where the bracket on the right is the standard ***Jacobi–Lie bracket*** of the vector fields,

$$\begin{aligned} -\,(\mathrm{ad}_u v)^i &= [u, v]^i = u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j}\,, \\ \text{or}\quad -\,\mathrm{ad}_{\mathbf{u}}\mathbf{v} &= [\mathbf{u}, \mathbf{v}] = \mathbf{u}\cdot\nabla\mathbf{v} - \mathbf{v}\cdot\nabla\mathbf{u}\,. \end{aligned} \tag{9.6}$$

Thus, the Lie bracket on $\mathfrak{X}(\mathcal{D})$, considered as the Lie algebra of $\mathrm{Diff}(\mathcal{D})$, is minus the standard Jacobi-Lie bracket.

10. EPDiff: Euler–Poincaré equation on the diffeomorphisms

10.1. **The n-dimensional EPDiff equation.** Eulerian geodesic motion of a fluid in n dimensions is generated as an EP equation via Hamilton's principle, when the Lagrangian is given by the kinetic energy. The kinetic energy defines a norm $\|\mathbf{u}\|^2$ for the Eulerian fluid velocity, represented by the contravariant vector function $\mathbf{u}(\mathbf{x},t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$. The choice of the kinetic energy as a positive functional of fluid velocity $\mathbf{u}$ is a modelling step that depends upon the physics of the problem being studied. We shall choose the kinetic-energy Lagrangian,

$$\ell = L_{\mathfrak{g}} = \frac{1}{2}\|u\|^2_{Q_{op}} = \frac{1}{2}\int \mathbf{u}\cdot\mathbf{m}\,\mathrm{d}V \quad \text{with} \quad \mathbf{m} := Q_{op}\mathbf{u}\,. \tag{10.1}$$

This Lagrangian may also be expressed as the L^2 pairing,

$$\ell = \frac{1}{2}\Big\langle u\,,\,m\Big\rangle = \frac{1}{2}\int \mathbf{u}\cdot Q_{op}\mathbf{u}\,\mathrm{d}V\,, \tag{10.2}$$

where, in a coordinate basis, the components of the vector field u and the 1-form density m are defined by

$$u = u^j\frac{\partial}{\partial x^j} = \mathbf{u}\cdot\nabla \quad \text{and} \quad m = m_i \mathrm{d}x^i \otimes \mathrm{d}V = \mathbf{m}\cdot\mathrm{d}\mathbf{x}\otimes\mathrm{d}V\,.$$

We use the same font for a quantity and its dual. In particular, italic font denotes vector field u and 1-form density m, and bold denotes vector $\mathbf{u}$ and covector $\mathbf{m}$. In eqns (10.1) and (10.2), the positive-definite, symmetric operator Q_{op} defines the norm $\|\mathbf{u}\|$, for appropriate (homogeneous, say, or periodic) boundary conditions. Conversely, the spatial velocity vector $\mathbf{u}$ is obtained by convolution of the momentum covector $\mathbf{m}$ with the ***Green's function*** for the operator Q_{op}. This Green's function G is defined by the vector equation

$$Q_{op}G = \delta(\mathbf{x})\,,$$

in which $\delta(\mathbf{x})$ is the Dirac measure and G satisfies appropriate boundary conditions. Consequently,

$$\mathbf{u}(\mathbf{x}) = (G * \mathbf{m})(\mathbf{x}) = \int G(\mathbf{x},\mathbf{x}')\mathbf{m}(\mathbf{x}')\,\mathrm{d}\mathbf{x}'\,. \tag{10.3}$$

For more discussion of Green's functions for linear differential operators, see [Tay96].

Remark

10.1. *An analogy exists between the kinetic energy in eqn (10.1) based on the norm $\|u\|_{Q_{op}}$ and the kinetic energy for the rigid body. In this analogy, the spatial velocity vector field u corresponds to body angular velocity, the operator Q_{op} to moment of inertia, and G to its inverse.*

Remark

10.2. *As defined earlier, the **EPDiff equation** is the Euler–Poincaré equation (9.2) for the Eulerian geodesic motion of a fluid with respect to norm $\|\mathbf{u}\|$. Its explicit form is given in the notation of Hamilton's principle by*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\ell}{\delta u} + \mathrm{ad}^*_u\frac{\delta\ell}{\delta u} = 0\,, \quad \textit{in which} \quad \ell[u] = \frac{1}{2}\|\mathbf{u}\|^2\,. \tag{10.4}$$

Definition

10.3. *The variational derivative of ℓ is defined by using the L^2 **pairing** between vector fields and 1-form densities as*

$$\delta\ell[u] = \left\langle \frac{\delta\ell}{\delta u}\,,\,\delta u\right\rangle = \int \frac{\delta\ell}{\delta\mathbf{u}}\cdot\delta\mathbf{u}\,\mathrm{d}V\,. \tag{10.5}$$

Consequently, the ***variational derivative*** with respect to the vector field u is the ***one-form density*** of momentum given as in eqn (10.1),

$$\frac{\delta \ell}{\delta u} = \frac{\delta \ell}{\delta \mathbf{u}} \cdot d\mathbf{x} \otimes \mathrm{d}V = m\,, \tag{10.6}$$

which has vector components given by

$$\frac{\delta \ell}{\delta \mathbf{u}} = Q_{op}\mathbf{u} = \mathbf{m}\,. \tag{10.7}$$

In addition, ad^* is the dual of the vector-field ad-operation (minus the vector-field commutator) with respect to the L^2 pairing,

$$\langle \mathrm{ad}^*_u\, m\,,\, v\rangle = \langle m\,,\, \mathrm{ad}_u v\rangle\,, \tag{10.8}$$

where u and v are vector fields. The notation $\mathrm{ad}_u\, v$ denotes the adjoint action of the ***right Lie algebra*** of $\mathrm{Diff}(\mathcal{D})$ on itself, given by

$$\mathrm{ad}_u\, v = -\,[u, v] \tag{10.9}$$

where $[u, v]$ is the commutator of vector fields, defined by

$$[u, v] := u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j}\,. \tag{10.10}$$

The pairing in eqn (10.8) is the L^2 pairing. Hence, upon integration by parts, one finds

$$\begin{aligned} \langle \mathrm{ad}^*_u\, m\,,\, v\rangle &= \langle m\,,\, \mathrm{ad}_u v\rangle \\ &= -\int m_i \Big(u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j}\Big) \mathrm{d}V \\ &= \int \Big(\frac{\partial}{\partial x^j}(m_i u^j) + m_j \frac{\partial u^j}{\partial x^i}\Big) v^i\, \mathrm{d}V\,, \end{aligned}$$

for homogeneous boundary conditions. In a coordinate basis, the preceding formula for $\mathrm{ad}^*_u\, m$ has the ***coordinate expression*** in $\mathbb{R}^n$,

$$\Big(\mathrm{ad}^*_u\, m\Big)_i \mathrm{d}x^i \otimes \mathrm{d}V = \Big(\frac{\partial}{\partial x^j}(m_i u^j) + m_j \frac{\partial u^j}{\partial x^i}\Big)\, \mathrm{d}x^i \otimes \mathrm{d}V\,. \tag{10.11}$$

In this notation, the abstract EPDiff equation (10.4) may be written explicitly in Euclidean coordinates as a partial differential equation for a covector function $\mathbf{m}(\mathbf{x}, t) : R^n \times R^1 \to R^n$. Namely, the ***EPDiff equation*** is given explicitly in Euclidean coordinates as

$$\boxed{\frac{\partial}{\partial t}\mathbf{m} \;+\; \underbrace{\mathbf{u}\cdot\nabla\mathbf{m}}_{\text{Convection}} \;+\; \underbrace{(\nabla\mathbf{u})^T\cdot\mathbf{m}}_{\text{Stretching}} \;+\; \underbrace{\mathbf{m}(\mathrm{div}\,\mathbf{u})}_{\text{Expansion}} = 0\,.} \tag{10.12}$$

Here, one denotes $(\nabla\mathbf{u})^T\cdot\mathbf{m} = \sum_j m_j \nabla u^j$. To explain the terms in underbraces, we rewrite EPDiff as preservation of the one-form density of momentum along the characteristic curves of the velocity. In vector coordinates, this is

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{m}\cdot \mathrm{d}\mathbf{x}\otimes \mathrm{d}V\Big) = 0 \quad\text{along}\quad \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{u} = G * \mathbf{m}\,. \tag{10.13}$$

This form of the EPDiff equation also emphasizes its non-locality, since the velocity is obtained from the momentum density by convolution against the Green's function G of the operator Q_{op}, as in eqn (10.3). One may check that the ***characteristic form*** of EPDiff in eqn (10.13) recovers its Eulerian form by computing directly the result that

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{m}\cdot\mathrm{d}\mathbf{x}\otimes\mathrm{d}V\Big) \\ &\quad= \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\cdot\mathrm{d}\mathbf{x}\otimes\mathrm{d}V + \mathbf{m}\cdot\mathrm{d}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\otimes\mathrm{d}V + \mathbf{m}\cdot\mathrm{d}\mathbf{x}\otimes\Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}V\Big) \\ &\quad= \Big(\frac{\partial}{\partial t}\mathbf{m} + \mathbf{u}\cdot\nabla\mathbf{m} \;+ \nabla\mathbf{u}^T\cdot\mathbf{m} + \mathbf{m}(\mathrm{div}\,\mathbf{u})\Big)\cdot\mathrm{d}\mathbf{x}\otimes\mathrm{d}V = 0\,, \end{aligned} \tag{10.14}$$

along

$$\frac{d\mathbf{x}}{\mathrm{d}t} = \mathbf{u} = G * \mathbf{m}\,.$$

This calculation explains the terms convection, stretching and expansion in the under-braces in eqn (10.12).

Remark

10.4. *In 2D and 3D, the EPDiff equation (10.12) may also be written equivalently in terms of the operators div, grad and curl as,*

$$\frac{\partial}{\partial t}\mathbf{m} - \mathbf{u} \times \operatorname{curl}\mathbf{m} + \nabla(\mathbf{u}\cdot\mathbf{m}) + \mathbf{m}(\operatorname{div}\mathbf{u}) = 0\,. \tag{10.15}$$

Thus, for example, the numerical solution of EPDiff would require an algorithm that has the capability to deal with the distinctions and relationships among the operators div, grad and curl.

10.2. **Variational derivation of EPDiff.** The EPDiff equation (10.12) may be derived by the following direct calculation for the present right invariant case in the continuum notation,

$$\begin{aligned}
\delta\int_a^b l(u)\mathrm{d}t &= \int_a^b \left\langle \frac{\delta l}{\delta u}, \delta u\right\rangle \mathrm{d}t = \int_a^b \left\langle \frac{\delta l}{\delta u}\,,\, \frac{\mathrm{d}v}{\mathrm{d}t} - \mathrm{ad}_u v \right\rangle \mathrm{d}t \\
&= \int_a^b \left\langle \frac{\delta l}{\delta u}\,,\, \frac{\mathrm{d}v}{\mathrm{d}t}\right\rangle \mathrm{d}t - \int_a^b \left\langle \frac{\delta l}{\delta u}\,,\, \mathrm{ad}_u v\right\rangle \mathrm{d}t \\
&= -\int_a^b \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta u} + \mathrm{ad}_u^*\frac{\delta l}{\delta u}\,,\, v\right\rangle \mathrm{d}t\,,
\end{aligned}$$

where, using (10.9) in (9.4), we have set

$$\delta u = \frac{\mathrm{d}v}{\mathrm{d}t} - \mathrm{ad}_u v\,, \tag{10.16}$$

for the variation of the right invariant vector field u. The angle brackets $\langle\cdot\,,\,\cdot\rangle$ denote the pairing between elements of the Lie algebra and its dual. In our case, this is the L^2 pairing between vector fields and 1-form densities in eqn (10.5), written in components as

$$\left\langle \frac{\delta l}{\delta u}, \delta u\right\rangle = \int \frac{\delta l}{\delta u^i}\,\delta u^i\,\mathrm{d}V\,.$$

This L^2 pairing yields the component form of the EPDiff equation explicitly, as

$$\begin{aligned}
\int_a^b \left\langle \frac{\delta l}{\delta u}, \delta u\right\rangle \mathrm{d}t &= \int_a^b \mathrm{d}t \int \frac{\delta l}{\delta u^i}\Big(\frac{\partial v^i}{\partial t} + u^j\frac{\partial v^i}{\partial x^j} - v^j\frac{\partial u^i}{\partial x^j}\Big)\,\mathrm{d}V \\
&= -\int_a^b \mathrm{d}t \int \Big\{\frac{\partial}{\partial t}\frac{\delta l}{\delta u^i} + \frac{\partial}{\partial x^j}\Big(\frac{\delta l}{\delta u^i}u^j\Big) + \frac{\delta l}{\delta u^j}\frac{\partial u^j}{\partial x^i}\Big\}v^i\,\mathrm{d}V \\
&\quad + \int_a^b \mathrm{d}t \int \Big\{\frac{\partial}{\partial t}\Big(\frac{\delta l}{\delta u^i}v^i\Big) + \frac{\partial}{\partial x^j}\Big(\frac{\delta l}{\delta u^i}v^iu^j\Big)\Big\}\mathrm{d}V\,.
\end{aligned} \tag{10.17}$$

Invoking $v^i = 0$ at the endpoints in time and taking the fluid velocity vector $\mathbf{u}$ to be tangent to the (fixed) boundary in space, then substituting the definition $m = \delta l/\delta u$ recovers the coordinate forms in Euclidean components for the coadjoint action of vector fields in eqn (10.11) and the EPDiff equation itself in eqn (10.12). When $\ell[u] = \frac{1}{2}\|u\|^2$, EPDiff describes geodesic motion on the diffeomorphisms with respect to the norm $\|u\|$.

10.3. **Noether's theorem for EPDiff.** Noether's theorem associates conservation laws to continuous symmetries of a Lagrangian. See, e.g., [Olv00] for a clear discussion of the classical theory. Momentum and energy conservation for the EPDiff equation in eqn (10.12) readily emerge from Noether's theorem, since the Lagrangian in eqn (10.1) admits space and time translations. That is, the action for EPDiff,

$$S = \int \ell[\mathbf{u}]\mathrm{d}t = \int \frac{1}{2}\|\mathbf{u}\|^2\mathrm{d}t\,,$$

is invariant under the following transformations,

$$x^j \to x'^j = x^j + c^j \quad \text{and} \quad t \to t' = t + \tau\,, \tag{10.18}$$

for constants τ and c^j, with $j = 1, 2, 3$. Noether's theorem then implies conservation of corresponding momentum components m_j, with $j = 1, 2, 3$, and energy E of the expected forms,

$$m_j = \frac{\delta \ell}{\delta u^j} \quad \text{and} \quad E = \frac{\delta \ell}{\delta u^j} u^j - \ell[\mathbf{u}]\,, \tag{10.19}$$

which may be readily verified.

Exercise. Show that the EPDiff equation (10.4) may be written as

$$\Big(\frac{\partial}{\partial t} + \mathcal{L}_{\mathbf{u}}\Big)\Big(\mathbf{m} \cdot \mathrm{d}\mathbf{x} \otimes \mathrm{d}V\Big) = 0\,, \tag{10.20}$$

where $\mathcal{L}_{\mathbf{u}}$ is the Lie derivative with respect to the vector field with components $\mathbf{u} = G * \mathbf{m}$. How does the Lie-derivative form of EPDiff in eqn (10.20) differ from its characteristic form (10.13)? Hint: compare the coordinate expression obtained from the dynamical definition of the Lie derivative with the corresponding expression obtained from its definition via Cartan's formula. ★

Exercise. Show that EPDiff in 1D may be written as

$$m_t + u m_x + 2 m u_x = 0\,. \tag{10.21}$$

How does the factor of 2 arise in this equation? Hint: Take a look at eqn (10.12). ★

Exercise. Write the EPDiff equation in coordinate form (10.12) for (a) the L^2 norm and (b) the H^1 norm (L^2 norm of the gradient) of the spatial fluid velocity. ★

Exercise. Verify that the EPDiff equation (10.12) conserves the spatially integrated momentum and energy in eqn (10.19). Hint: for momentum conservation look at eqn (10.17) when $v^j = c^j$ for spatial translations. ★

10.4. **Fluids background for EPDiff.** The configuration space $\mathrm{Diff}(\mathcal{D})$ is a group, with the group operation being composition and the group identity being the identity map. This group acts on $\mathcal{D}$ in the obvious way: $g \cdot X := g(X)$, where we are using the 'dot' notation for the group action.

Definition

10.5. *During a motion g_t or $g(t)$, the particle labelled by X describes a path in $\mathcal{D}$ along a locus of points*

$$x(X, t) := g_t(X) = g(t) \cdot X\,, \tag{10.22}$$

which are called the ***Eulerian*** *or* ***spatial points*** *of the path. This locus of points in $\mathbb{R}^n$ is also called the* ***Lagrangian****, or* ***material, trajectory****, because a Lagrangian fluid parcel follows this path in space.*

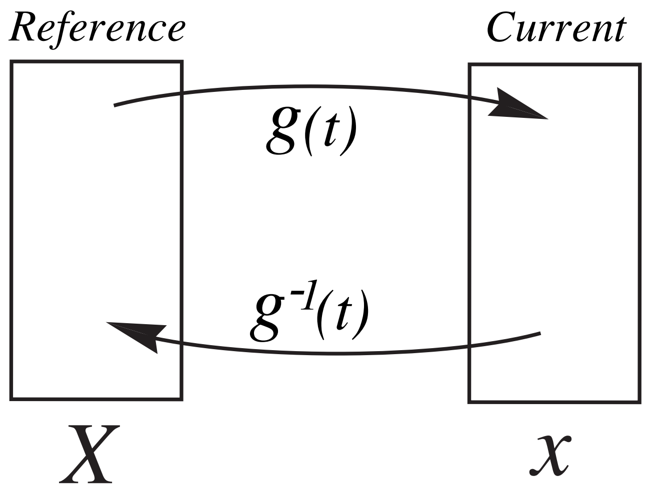


FIGURE 4. The map from Lagrange reference coordinates X in the fluid to the current Eulerian spatial position x is performed by the time-dependent diffeomorphism $g(t)$, so that $x(t, X) = g(t) \cdot X$.

Definition

10.6. *The* ***Lagrangian****, or* ***material****,* ***velocity*** *U of the system along the motion g_t or $g(t)$ is defined by taking the time derivative of the Lagrangian trajectory (10.22) keeping the particle labels X fixed:*

$$U(X, t) := \frac{\partial}{\partial t} g_t \cdot X = \frac{\partial}{\partial t} x(X, t)\,. \tag{10.23}$$

Thus $U(X, t)$ is the velocity of the particle with label X at time t.

The ***Eulerian****, or* ***spatial****,* ***velocity*** *u of the system is velocity expressed as a function of spatial position and time, meaning that if $x = x(X, t) = g_t(X)$ then*

$$u(x, t) := U(X, t) = U(g_t^{-1}(x), t)\,. \tag{10.24}$$

Thus, $u(x, t)$ is the velocity at time t of the particle currently in position x.

The Eulerian velocity u can also be regarded as a time-dependent vector field $u_t \in \mathfrak{X}(\mathcal{D})$, where $u_t(x) := u(x, t)$. It follows from eqn (10.24) that

$$U_t = u_t \circ g_t\,. \tag{10.25}$$

In this sense, the Lagrangian velocity field at a particular time is a *right translation* of the Eulerian velocity field. This observation leads to consideration of the Lie-group structure of Diff$(\mathcal{D})$.

11. Clebsch action principle for EPDiff(Emb$(S, \mathbb{R}^n)$)

To set the notation, let the domain $\mathcal{D} = \mathbb{R}^n$, fix a k-dimensional manifold S with a given volume element and whose points are denoted $s \in S$. Let Emb$(S, \mathbb{R}^n)$ denote the set of smooth embeddings $\mathbf{Q} : S \to \mathbb{R}^n$. (If the EPDiff equations are taken on a manifold M, replace $\mathbb{R}^n$ with M.) Under appropriate technical conditions, which we shall just treat formally here, Emb$(S, \mathbb{R}^n)$ is a smooth manifold. (See, for example, [EM70] for a discussion and references.)

The tangent space $T_{\mathbf{Q}}\,\mathrm{Emb}(S, \mathbb{R}^n)$ to Emb$(S, \mathbb{R}^n)$ at the point $\mathbf{Q} \in \mathrm{Emb}(S, \mathbb{R}^n)$ is given by the space of ***material velocity fields***, namely the linear space of maps $\mathbf{V} : S \to \mathbb{R}^n$ that are vector fields over the map $\mathbf{Q}$. The dual space to this space will be identified with the space of one-form densities over $\mathbf{Q}$, which we shall regard as maps $\mathbf{P} : S \to (\mathbb{R}^n)^*$. In summary, the cotangent bundle $T^*\,\mathrm{Emb}(S, \mathbb{R}^n)$ is identified with the space of pairs of maps $(\mathbf{Q}, \mathbf{P})$.

Definition

11.1 (Clebsch action principle for EPDiff). *For a given functional $l : \mathfrak{X}(\mathcal{D}) \times T^*M$ with $M := \mathrm{Emb}(S, \mathbb{R}^n)$, the smooth manifold of embeddings, the Clebsch action principle for EPDiff is*

$$\delta S[\boldsymbol{u}(t); \boldsymbol{Q}(t), \boldsymbol{P}(t)] = \delta \int_{t_1}^{t_2} l[\boldsymbol{u}(t)] + \Big\langle \boldsymbol{P}(t), \dot{\boldsymbol{Q}}(t) - \boldsymbol{u}(\boldsymbol{Q}, t) \Big\rangle_{T^*M} \,\mathrm{d}t = 0, \tag{11.1}$$

where $\boldsymbol{u} \in \mathfrak{X}(\mathcal{D})$ is the Eulerian velocity, $\mathbf{Q}(t)$ is the map $S \to \mathbb{R}^n$, $\boldsymbol{P}(t)$ is a Lagrange multiplier in $T^*_{\boldsymbol{Q}}M$ and $\langle \cdot, \cdot \rangle_{T^*M}$ is the standard inner product on T^*M, which contains an implied sum and integral (denoted simply as $\sum$) over the connected components of $M := \mathrm{Emb}(S, \mathbb{R}^n)$.

This Lagrangian leads to the following Clebsch variational equations.

Lemma

11.2 (Clebsch equations).
The action principle (11.1) has stationary points when the following Clebsch variational equations are satisfied:

$$\frac{\delta l}{\delta \boldsymbol{u}} = \sum \boldsymbol{P}(t)\, \delta(\boldsymbol{x} - \boldsymbol{Q}(t)), \qquad \dot{\boldsymbol{Q}} = \boldsymbol{u}(\boldsymbol{Q}, t), \qquad \dot{\boldsymbol{P}} = -\left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{Q}}\right)^T \cdot \boldsymbol{P} \tag{11.2}$$

Proof.

$$\begin{aligned}
0 &= \delta \int_{t_1}^{t_2} l[\boldsymbol{u}] + \Big\langle \boldsymbol{P}, \dot{\boldsymbol{Q}} - \boldsymbol{u}(\boldsymbol{Q}, t) \Big\rangle_{T^*M} \mathrm{d}t \\
&= \delta \int_{t_1}^{t_2} l[\boldsymbol{u}] + \Big\langle \boldsymbol{P}, \dot{\boldsymbol{Q}} - \int \boldsymbol{u}(\boldsymbol{x}, t)\delta(\boldsymbol{x} - \boldsymbol{Q}(t))\, \mathrm{d}x \Big\rangle_{T^*M} + \delta \int_{t_1}^{t_2} \Big\langle \boldsymbol{P}, \dot{\boldsymbol{Q}} \Big\rangle_{T^*M} \mathrm{d}t \\
&= \int_{t_1}^{t_2} \Big\langle \frac{\delta l}{\delta \boldsymbol{u}} - \sum \boldsymbol{P}\, \delta(\boldsymbol{x} - \boldsymbol{Q}(t)),\, \delta \boldsymbol{u} \Big\rangle_{\mathfrak{X}^*} + \Big\langle \delta \boldsymbol{P}, \dot{\boldsymbol{Q}} - \boldsymbol{u}(\boldsymbol{Q}, t) \Big\rangle_{T^*M} + \Big\langle \boldsymbol{P}, \delta \dot{\boldsymbol{Q}} - \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{Q}} \cdot \delta \boldsymbol{Q} \Big\rangle_{T^*M} \mathrm{d}t \\
&= \\
&= \int_{t_1}^{t_2} \Big\langle \frac{\delta l}{\delta \boldsymbol{u}} - \sum \boldsymbol{P}\, \delta(\boldsymbol{x} - \boldsymbol{Q}(t)),\, \delta \boldsymbol{u} \Big\rangle_{\mathfrak{X}^*} \\
&\qquad + \Big\langle \delta \boldsymbol{P}, \dot{\boldsymbol{Q}} - \boldsymbol{u}(\boldsymbol{Q}, t) \Big\rangle_{T^*M} - \Big\langle \dot{\boldsymbol{P}} + \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{Q}}\right)^T \cdot \boldsymbol{P},\, \delta \boldsymbol{Q} \Big\rangle_{T^*M} \mathrm{d}t
\end{aligned}$$

and the result follows since $\delta \boldsymbol{u}$, $\delta \boldsymbol{P}$, and $\delta \boldsymbol{Q}$ are arbitrary. □

Corollary

11.3. *The momentum equation arising from the Clebsch action principle (11.2)*

$$\boldsymbol{m} := \frac{\delta l}{\delta \boldsymbol{u}} = \sum \boldsymbol{P}(t)\, \delta(\boldsymbol{x} - \boldsymbol{Q}(t))$$

for measure-valued solutions of the EPDiff *equation* (9.2)*, defines an equivariant momentum map*

$$\mathbf{J}_{\mathrm{Sing}} : T^* \mathrm{Emb}(S, \mathbb{R}^n) \to \mathfrak{X}(\mathbb{R}^n)^*,$$

called the ***singular solution momentum map****. For discussions of this momentum map from various viewpoints, see* [HoMa2004].

In particular, expressing $\mathbf{m} \in \mathfrak{X}(\mathbb{R}^n)^*$ in terms of $\mathbf{Q}, \mathbf{P} \in T^* \mathrm{Emb}(S, \mathbb{R}^n)$ (which are functions of coordinates s on the connected components of the embedded manifold S) is a momentum map from the space of $(\mathbf{Q}(s), \mathbf{P}(s))$ to the space of $\mathbf{m}(\boldsymbol{x})$. Its equivariance follows because it is a cotangent lift. Consequently, $\boldsymbol{m} := \frac{\delta l}{\delta \boldsymbol{u}}$ satisfies the (right invariant) ***Euler–Poincaré equations*** (9.2).

Lemma

11.4 (Legendre transform). *The Clebsch equations (11.2) for $\boldsymbol{Q}$, $\boldsymbol{P}$ and the momentum map are canonical for the Hamiltonian (Routhian) given by the Legendre transform,*

$$H(\boldsymbol{Q}, \boldsymbol{P}) = \langle \boldsymbol{P}(t), \boldsymbol{u}(\boldsymbol{Q}, t) \rangle_{T^*M} - l[\boldsymbol{u}(t)].$$

Proof. The result follows directly by calculating the canonical equations for this Hamiltonian, from

$$\delta H = \Big\langle \delta \boldsymbol{P}, \boldsymbol{u}(\boldsymbol{Q},t) \Big\rangle_{T^*M} + \left\langle \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{Q}} \right)^T \cdot \boldsymbol{P}, \delta \boldsymbol{Q} \right\rangle_{T^*M} - \left\langle \frac{\delta l}{\delta \boldsymbol{u}} - \sum \boldsymbol{P}\, \delta(\boldsymbol{x} - \boldsymbol{Q}(t)),\, \delta \boldsymbol{u} \right\rangle_{\mathfrak{X}^*}$$

□

12. EPDiff solution behaviour

> I shall speak of things . . . so singular in their oddity as in some manner to instruct, or at least entertain, without wearying.
>
> – Lorenzo da Ponte (1749–1838), [LDP2000]

This section discusses the coherent particle-like properties of the unidirectional singular solutions of the EPDiff equation (10.21). These singular solutions emerge from any smooth spatially confined initial velocity profile $u(x, 0)$. After emerging, they dominate the evolution in interacting fully nonlinearly by exchanging momentum in elastic collisions. The mechanism for their emergence is shown to be pulse steepening due to nonlinearity. Several examples of the dynamics among singular solutions are given.

12.1. **Introduction.** Consider the following particular case of the EPDiff equation (10.21) in one spatial dimension,

$$m_t + um_x + 2mu_x = 0 \quad \text{with} \quad m = (1 - \alpha^2 \partial_x^2)u\,, \tag{12.1}$$

in which the fluid velocity u is a function of position x on the real line and time t. This equation governs geodesic motion on the smooth invertible maps (diffeomorphisms) of the real line with respect to the metric associated with the H^1 Sobolev norm of the fluid velocity given by

$$\|u\|_{H^1}^2 = \int (u^2 + \alpha^2 u_x^2)\, \mathrm{d}x\,. \tag{12.2}$$

The ***peakon*** is the solitary travelling wave solution for the EPDIff equation (12.1),

$$u(x, t) = c\, \mathrm{e}^{-|x-ct|/\alpha}\,. \tag{12.3}$$

The peakon travelling wave moves at a speed equal to its maximum height, at which it has a sharp peak (jump in derivative). The spatial velocity profile $\mathrm{e}^{-|x|/\alpha}$ is the ***Green's function*** for the Helmholtz operator $(1-\alpha^2\partial_x^2)$ on the real line with vanishing boundary conditions at spatial infinity. In particular, it satisfies

$$(1 - \alpha^2 \partial_x^2)\mathrm{e}^{-|x-ct|/\alpha} = 2\alpha\delta(x - ct)\,. \tag{12.4}$$

A novel feature of the EPDiff equation (12.1) is that it admits solutions representing a ***wave train of peakons***

$$u(x, t) = \sum_{a=1}^{N} p_a(t)\mathrm{e}^{-|x-q_a(t)|/\alpha}\,. \tag{12.5}$$

By eqn (12.4), this corresponds to a sum over delta functions representing the ***singular solution*** in momentum,

$$m(x, t) = 2\alpha \sum_{a=1}^{N} p_a(t)\, \delta(x - q_a(t))\,, \tag{12.6}$$

in which the ***delta function*** $\delta(x - q)$ is defined by

$$f(q) = \int f(x)\delta(x - q)\, \mathrm{d}x\,, \tag{12.7}$$

for an arbitrary smooth function f. Such a sum is an *exact solution* of the EPDiff equation (12.1) provided the time-dependent parameters $\{p_a\}$ and $\{q_a\}$, $a = 1, \ldots, N$, satisfy certain canonical Hamiltonian equations that will be discussed later.

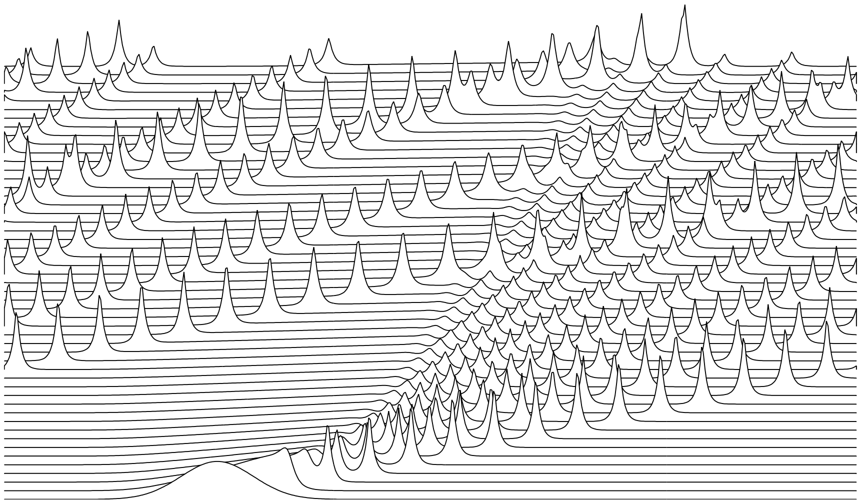


FIGURE 5. Under the evolution of the EPDiff equation (12.1), an ordered ***wave train of peakons*** emerges from a smooth localized initial condition (a Gaussian). The spatial profiles at successive times are offset in the vertical to show the evolution. The peakon wave train eventually wraps around 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 [CH93].

Remark

12.1. *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 (12.1) for essentially any spatially confined initial condition. An example of the emergence of a wave train of peakons from a Gaussian initial condition is shown in Figure 5.*

12.2. **Steepening lemma: the peakon-formation mechanism.** We may understand the mechanism for the emergent formation of the peakons seen in Figure 5, by showing that initial conditions exist for which the solution of the EPDiff equation (13.11) 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,

Lemma

12.2 (Steepening lemma [CH93])**.**
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 H^1 Sobolev norm of the fluid velocity in eqn (12.2) 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 equation (12.1), rewritten as,

$$(1 - \alpha^2\partial^2)(u_t + uu_x) = -\,\partial\Big(u^2 + \frac{\alpha^2}{2}u_x^2\Big)\,, \tag{12.8}$$

yields an equation for the evolution of s. Namely, using $u_{xx}(\overline{x}(t), t) = 0$ leads to

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\,\frac{1}{2}s^2 + \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{sgn}(\overline{x} - y)\mathrm{e}^{-|\overline{x}-y|}\partial_y\left(u^2 + \frac{1}{2}u_y^2\right)\mathrm{d}y\,. \tag{12.9}$$

Integrating by parts and using the inequality $A^2 + B^2 \geq 2AB$, for any two real numbers A and B, leads to

$$\begin{aligned} \frac{\mathrm{d}s}{\mathrm{d}t} &= -\frac{1}{2}s^2 - \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{e}^{-|\overline{x}-y|}\left(u^2 + \frac{1}{2}u_y^2\right)\mathrm{d}y + u^2(\overline{x}(t),t) \\ &\leq -\frac{1}{2}s^2 + 2u^2(\overline{x}(t),t)\,. \end{aligned} \tag{12.10}$$

Then, provided $u^2(\overline{x}(t),t)$ remains finite, say less than a number $M/4$, we have

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\frac{1}{2}s^2 + \frac{M}{2}\,, \tag{12.11}$$

which implies, for negative slope initially $s \leq -\sqrt{M}$, that

$$s \leq \sqrt{M}\coth\left(\sigma + \frac{t}{2}\sqrt{M}\right)\,, \tag{12.12}$$

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 eqn (12.11) exists is verified in general by a Sobolev inequality. In fact, $M = 8H_1$, since

$$\max_{x\in\mathbb{R}} u^2(x,t) \leq \int_{-\infty}^{\infty}(u^2 + u_x^2)\,\mathrm{d}x = 2H_1 = const\,. \tag{12.13}$$

$\square$

Remark

12.3. *Suppose the initial condition is anti-symmetric, so the inflection point at $u = 0$ is fixed and $\mathrm{d}\overline{x}/\mathrm{d}t = 0$, due to the symmetry $(u,x) \to (-u,-x)$ admitted by eqn (14.1). 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.*

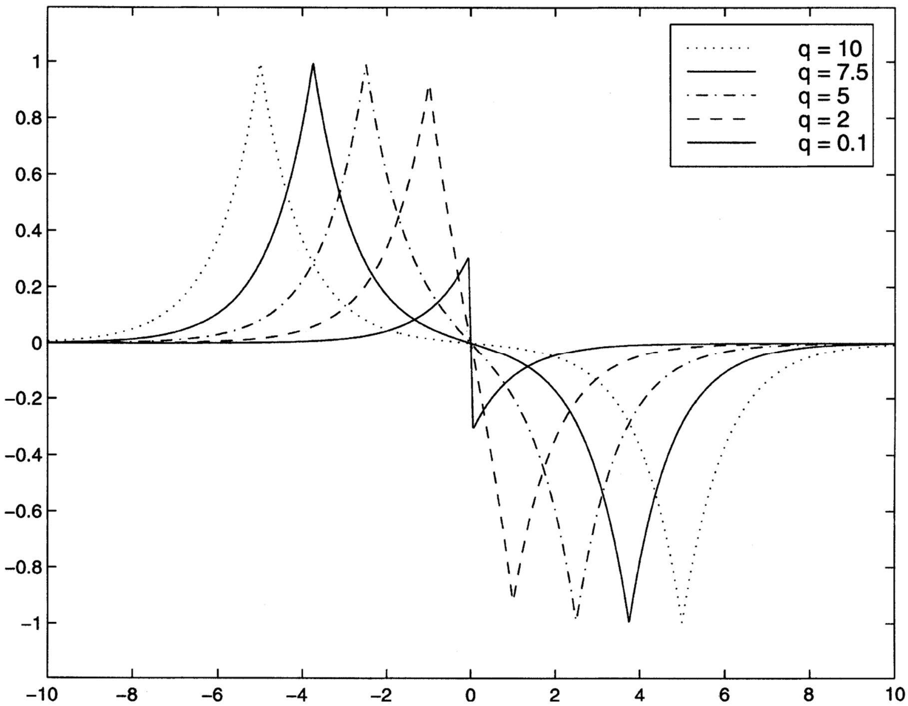


FIGURE 6. This is the velocity profile (13.35) for the peakon-antipeakon head-on collision as a function of separation between the peaks [FH01].

Remark

12.4 (Implications of the steepening lemma)**.**

• *The steepening lemma indicates that travelling wave solutions of the EPDiff equation (12.1) cannot have the* sech^2 *shape that appears for KdV solitons, 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 of negative slope in any confined initial velocity distribution triggers the steepening lemma as the* ***mechanism*** *for the formation of the peakons.*

• *Namely. the initial (positive) velocity profile "leans" to the right and steepens, then produces a peakon that is taller than the initial profile, so it propagates away to the right.*

• *This process leaves a profile behind with an inflection point of negative slope; so it repeats, thereby producing a wave train of peakons with the tallest and fastest ones moving rightward in order of height.*

• *Remarkably, this recurrent process produces only peakons.*

The EPDiff equation (12.1) arises from a shallow water wave equation in the limit of zero linear dispersion in one dimension. As we shall see, the peakon solutions (12.6) for EPDiff generalize to higher dimensions and other kinetic energy norms.

Exercise. Verify that the EPDiff equation (12.1) preserves the H^1 norm (12.2). ★

Exercise. Verify that the peakon formula (12.3) provides the solitary travelling wave solution for the EPDiff equation (12.1). ★

Exercise. Verify formula (12.4) for the Green's function. Why is this formula useful in representing the travelling-wave solution of the EPDIff equation (12.1)? ★

13. Shallow-water background for peakons

The EPDiff equation (12.1) whose solutions admit peakon wave trains (12.5) may be derived by taking the zero-dispersion limit of another equation obtained from Euler's fluid equations by using asymptotic expansions for shallow water waves [CH93]. Euler's equations for irrotational incompressible ideal fluid motion under gravity with a free surface have an asymptotic expansion for shallow water waves that involves two small parameters, ϵ and δ^2, with ordering $\epsilon \geq \delta^2$. These small parameters are $\epsilon = a/h_0$ (the ratio of wave amplitude to mean depth) and $\delta^2 = (h_0/l_x)^2$ (the squared ratio of mean depth to horizontal length, or wavelength).

In one spatial dimension, EPDiff is the zero-dispersion limit of the Camassa–Holm (CH) equation for shallow water waves, which is the $b = 2$ case of the following ***b-equation***, that results from the asymptotic expansion for shallow water waves,

$$m_t + c_0 u_x + u m_x + b\, m u_x - \gamma u_{xxx} = 0\,. \tag{13.1}$$

Here, $m = u - \alpha^2 u_{xx}$ is the momentum variable, and the constants α^2 and γ/c_0 are squares of length scales. At *linear* order in the asymptotic expansion for shallow water waves in terms of the small parameters ϵ and δ^2, one finds $\alpha^2 \to 0$, so that $m \to u$ in (13.1). In this case, the famous ***Korteweg–de Vries*** (KdV) soliton equation is recovered for $b = 2$,

$$u_t + 3uu_x = -c_0 u_x + \gamma u_{xxx}\,. \tag{13.2}$$

Any value of the parameter *except* $b = -1$ may be achieved in eqn (13.1) by an appropriate near-identity (normal form) transformation of the solution [DGH04]. The value $b = -1$ is disallowed in (13.1) because it cancels the leading-order nonlinearity and, thus, breaks the asymptotic ordering.

Because of the relation $m = u - \alpha^2 u_{xx}$, the b-equation (13.1) is ***non-local***. In other words, it is an integral-partial differential equation. In fact, after writing eqn (13.1) equivalently as,

$$(1 - \alpha^2 \partial_x^2)(u_t + uu_x) = -\,\partial_x \left(\frac{b}{2} u^2 + \frac{3-b}{2}\, \alpha^2 u_x^2 \right) - c_0 u_x + \gamma u_{xxx}\,. \tag{13.3}$$

The b-equation may be expressed in ***hydrodynamic form*** as

$$u_t + uu_x = -\,p_x\,, \tag{13.4}$$

with a 'pressure' p given by

$$p = G * \left(\frac{b}{2}u^2 + \frac{3-b}{2}\,\alpha^2 u_x^2 + c_0 u - \gamma u_{xx}\right), \tag{13.5}$$

in which the convolution kernel is the Green's function $G(x,y) = (2\alpha)^{-1}\mathrm{e}^{-|x-y|/\alpha}$ for the Helmholtz operator $(1-\alpha^2\partial_x^2)$.

One sees the interplay between local and non-local linear dispersion in the b-equation by linearizing eqn (13.3) around $u = 0$ to find its phase-velocity relation,

$$\frac{\omega}{k} = \frac{c_0 + \gamma\, k^2}{1 + \alpha^{\,2} k^2}\,, \tag{13.6}$$

obtained for waves with frequency ω and wave number k. For $\gamma/c_0 > 0$, short waves and long waves travel in the same direction. Long waves travel faster than short ones (as required in shallow water) provided $\gamma/c_0 < \alpha^2$. Then, the phase velocity lies in the interval $\omega/k \in (\gamma/\alpha^{\,2}, c_0]$. The parameters c_0 and γ represent linear wave dispersion, which modifies and may eventually balance the tendency for nonlinear waves to steepen and break. The parameter α, which introduces non-locality, also allows a balance leading to a stable wave shape, even in the absence of c_0 and γ.

The nonlinear effects of the parameter b on the solutions of eqn (13.1) were investigated in Holm and Staley [HS03], where b was treated as a bifurcation parameter. In the limiting case when the linear dispersion coefficients are absent, peakon solutions of eqn (13.1) are allowed theoretically for any value of b. However, they were found numerically to be stable only for $b > 1$. These coherent solutions are allowed, because the two nonlinear terms in eqn (13.1) may balance each other, even in the *absence* of linear dispersion. However, the instability of the peakons found numerically for $b < 1$ indicates that the relative strengths of the two nonlinearities will determine whether this balance can be maintained.

Proposition

13.1. *A solution u of the b-equation (13.1) with $c_0 = 0$ and $\gamma = 0$ vanishing at spatial infinity blows up in H^1 if and only if its first-order derivative blows up, that is, if wave breaking occurs.*

Proof. This result is implied by Exercise 13.2. □

Lemma

13.2 (Steepening lemma for the b-equation with $b > 1$)**.**
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. Assume that the velocity at the inflection point remains finite. Then, the negative slope at the inflection point will become vertical in finite time, provided $b > 1$.

Proof. Consider the evolution of the slope at the inflection point $x = \overline{x}(t)$. Define $s = u_x(\overline{x}(t), t)$. Then, the b-equation (13.1) with $c_0 = 0$ and $\gamma = 0$ may be rewritten in hydrodynamic form as, cf. eqn (13.4),

$$u_t + uu_x = -\,\partial_x G * \left(\frac{b}{2}u^2 + \frac{3-b}{2}\,\alpha^2 u_x^2\right). \tag{13.7}$$

The spatial derivative of this yields an equation for the evolution of s. Namely, using $u_{xx}(\overline{x}(t),t)=0$ leads to

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t}+s^2 &= -\partial_x^2(G*p) \quad \text{with} \quad p:=\left(\frac{b}{2}u^2(\overline{x}(t),t)+\frac{3-b}{2}\,\alpha^2 s^2\right)\\ &= \frac{1}{\alpha^2}(1-\alpha^2\partial_x^2)G*p-\frac{1}{\alpha^2}G*p\\ &= \frac{1}{\alpha^2}p-\frac{1}{\alpha^2}G*p\,. \end{aligned} \tag{13.8}$$

This calculation implies

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t} &= \frac{1-b}{2}s^2-\frac{1}{2\alpha}\int_{-\infty}^{\infty}\mathrm{e}^{-|\overline{x}-y|/\alpha}\left(\frac{b}{2}u^2+\frac{3-b}{2}\alpha^2u_y^2\right)\mathrm{d}y+\frac{b}{2\alpha^2}u^2(\overline{x}(t),t)\\ &\le \frac{1-b}{2}s^2+\frac{b}{2\alpha^2}u^2(\overline{x}(t),t)\,, \end{aligned} \tag{13.9}$$

where we have dropped the negative middle term in the last step. Then, provided $u^2(\overline{x}(t),t)$ remains finite, say less than a number M, we have

$$\frac{\mathrm{d}s}{\mathrm{d}t} \le \frac{1-b}{2}s^2+\frac{bM}{2\alpha^2}\,, \tag{13.10}$$

which implies, for negative slope initially and $b>1$, that the slope remains negative and becomes vertical in finite time. This proves the steepening lemma for the b-equation and identifies $b=1$ as a special value. □

Remark

13.3. *One might wonder whether the dispersionless CH equation is the only shallow water b-equation that both possesses peakon solutions and is completely integrable as a Hamiltonian system. Mikhailov and Novikov* [MN02] *showed that among the b-equations only the cases $b=2$ and $b=3$ are completely integrable as Hamiltonian systems. The case $b=3$ is the Degasperis–Processi equation, whose peakon solutions are studied in* [DHH03].

Remark

13.4. *Hereafter, we specialize the b-equation (13.1) to the case $b=2$. If, in addition, $c_0=0$ and $\gamma=0$, then the b-equation specializes to EPDiff.*

13.1. **Hamiltonian dynamics of EPDiff peakons.** Upon substituting the peakon solution expressions (12.5) for velocity u and eqn (12.6) for momentum m into the EPDiff equation,

$$m_t+um_x+2mu_x=0\,,\quad \text{with} \quad m=u-\alpha^2u_{xx}\,, \tag{13.11}$$

one finds ***Hamilton's canonical equations*** for the dynamics of the discrete set of peakon parameters $p_a(t)$ and $q_a(t)$. Namely,

$$\dot q_a(t)=\frac{\partial H_N}{\partial p_a} \quad \text{and} \quad \dot p_a(t)=-\,\frac{\partial H_N}{\partial q_a}\,, \tag{13.12}$$

for $a=1,2,\ldots,N$, with Hamiltonian given by [CH93],

$$H_N=\tfrac{1}{2}\sum_{a,b=1}^{N}p_a\,p_b\,\mathrm{e}^{-|q_a-q_b|/\alpha}\,. \tag{13.13}$$

The first canonical equation in eqn (13.12) implies that the peaks at the positions $x=q^a(t)$ in the peakon-train solution (12.5) 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 momentum solution ansatz (12.6) is the map from Lagrangian coordinates to Eulerian coordinates (that is, the ***Lagrange-to-Euler map***) for the momentum.

Remark

13.5. *The peakon wave train (12.6) forms a finite-dimensional invariant manifold of solutions of the EPDiff equation. On this invariant manifold of solutions for the partial differential equation (13.11), the dynamics turns out to be canonically Hamiltonian as in eqn (13.12). This canonical Hamiltonian structure of the peakon solutions arises because the solution ansatz (12.6) for momentum m is a momentum map* [HoMa2004].

13.2. **Pulsons: Singular solutions of EPDiff for other Green's functions.** The Hamiltonian H_N in eqn (13.13) depends on G, the Green's function for the relation $u = G * m$ between velocity u and momentum m. For the Helmholtz operator on the real line this Green's function is given by eqn (12.4) as $G(x) = \mathrm{e}^{-|x|/\alpha}/2\alpha$. However, the singular momentum solution ansatz (12.6) is *independent* of this Green's function. Thus, we may conclude the following [FH01].

Proposition

13.6. *The singular momentum solution ansatz*

$$m(x,t) = \sum_{a=1}^{N} p_a(t)\,\delta(x - q_a(t))\,, \tag{13.14}$$

for EPDiff,

$$m_t + u m_x + 2 m u_x = 0\,, \quad \textit{with} \quad u = G * m\,, \tag{13.15}$$

provides a finite-dimensional invariant manifold of solutions governed by canonical Hamiltonian dynamics, for any choice *of the Green's function $G(x)$ relating velocity u and momentum m.*

Proof. The singular momentum solution ansatz (13.14) is *independent* of the Green's function G. □

Remark

13.7. *The pulson singular solutions (13.14) of the EPDiff equation (13.15) form an N-dimensional invariant symplectic manifold, on which the EPDiff solution dynamics is governed by a canonical Hamiltonian system for the conjugate pairs of variables $(q_a,\, p_a)$ with $a = 1, 2, \ldots, N$.*

Perhaps surprisingly, these singular solutions will turn out to emerge from any smooth confined initial distribution of momentum.

The fluid velocity solutions corresponding to the singular momentum ansatz (13.14) for eqn (13.15) are the ***pulsons***. A pulson wave train is defined by the sum over N velocity profiles determined by the Green's function G, as

$$u(x,t) = \sum_{a=1}^{N} p_a(t) G\big(x, q_a(t)\big)\,. \tag{13.16}$$

A solitary travelling wave solution for the pulson is given by

$$u(x,t) = cG(x,ct) = cG(x - ct) \quad \text{with} \quad G(0) = 1\,, \tag{13.17}$$

where one finds $G(x,ct) = G(x-ct)$, provided the Green's function G is translation-invariant.

For EPDiff (13.15) with any choice of the Green's function G, the singular momentum solution ansatz (13.14) results in a finite-dimensional invariant manifold of exact solutions. The $2N$ parameters $p_a(t)$ and $q_a(t)$ in these pulson-train solutions of EPDiff satisfy Hamilton's canonical equations

$$\frac{\mathrm{d}q_a}{\mathrm{d}t} = \frac{\partial H_N}{\partial p_a} \quad \text{and} \quad \frac{\mathrm{d}p_a}{\mathrm{d}t} = -\,\frac{\partial H_N}{\partial q_a}\,, \tag{13.18}$$

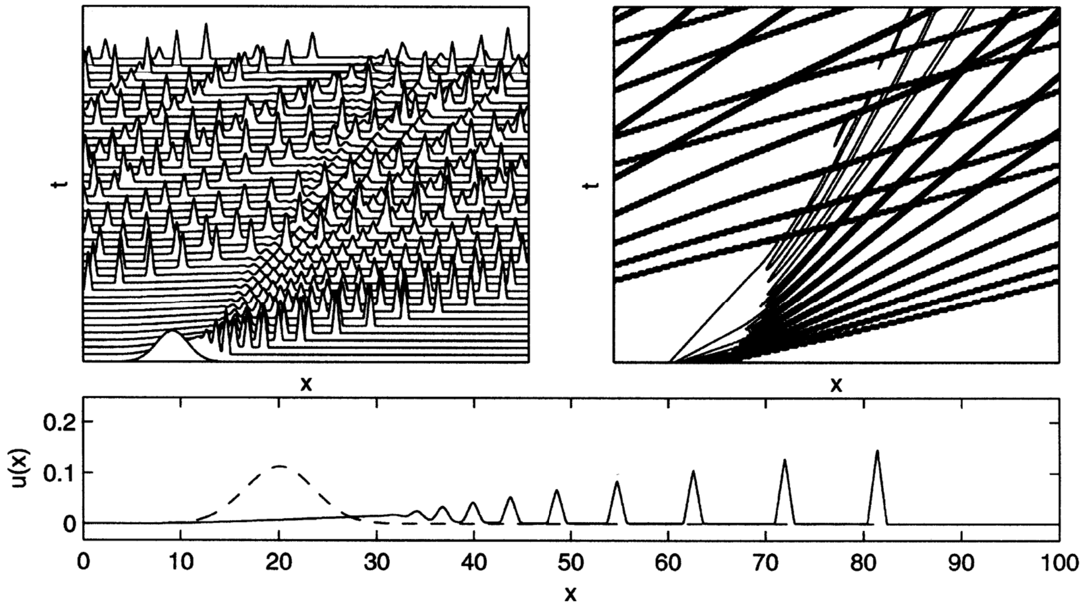


FIGURE 7. When the Green's function G has a triangular profile, a train of triangular pulsons emerges from a Gaussian initial velocity distribution as it evolves under the EPDiff equation (12.1). The upper panels show the collisions that occur as the faster triangular pulsons overtake the slower ones as they cross and re-cross the periodic domain. The upper left panel shows the progress of the pulsons by by showing offsets of the velocity profile at equal time intervals. The upper right panel shows the pulson paths obtained by plotting their elevation topography.

with N-particle Hamiltonian,

$$H_N = \frac{1}{2}\sum_{a,b=1}^{N} p_a\, p_b\, G(q_a, q_b)\,. \tag{13.19}$$

The canonical equations for the parameters in the pulson train define an invariant manifold of singular momentum solutions and provide a phase-space description of geodesic motion with respect to the cometric (inverse metric) given by the Green's function G. Mathematical analysis and numerical results for the dynamics of these pulson solutions are given in [FH01] whose results show how the results of collisions of pulsons (13.16) depend upon the *shape* of their travelling wave profile. The effects of the travelling-wave pulse shape

$$u(x-ct) = cG(x-ct)$$

on the multipulson collision dynamics are reflected in the Hamiltonian (13.19) that governs this dynamics. For example, see Figure 7, in which the pulsons are *triangular*.

Exercise. Verify the hydrodynamic form of the b-equation in eqn (13.3). ★

Exercise. Verify that the b-equation (13.1) with $c_0 = 0$ and $\gamma = 0$ admits peakon-train solutions of the form (12.5) for any value of b. ★

Exercise. Verify that the b-equation (13.1) with $c_0 = 0$ and $\gamma = 0$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u\|^2_{H^1} = (b-2)\int u_x^3\,\mathrm{d}x\,,$$

for any value of b and for solutions that vanish sufficiently rapidly at spatial infinity that no endpoint contributions arise upon integration by parts. ★

Exercise. Prove a steepening lemma for the b-equation (13.1) with $c_0 = 0$ and $\gamma = 0$ that avoids the assumption that $u^2(\overline{x}(t), t)$ remains finite. That is, establish a necessary and sufficient condition depending only on the initial data for blow-up to occur in finite time. How does this condition depend on the value of b? Does this steepening lemma hold for every value of $b > 1$? ★

Exercise. Are the equations of peakon dynamics for the b-equation (13.1) with $c_0 = 0$ and $\gamma = 0$ canonically Hamiltonian for every value of b? HINT: try $b = 3$. ★

13.3. **Peakons.**

13.3.1. *Pulson–Pulson interactions.* The solution of EPDiff in 1D

$$\partial_t m + u m_x + 2 u_x m = 0\,, \tag{13.20}$$

with $u = G * m$ for the momentum $m = Q_{op} u$ is given for the interaction of only two pulsons by the sum of delta functions in eqn (13.14) with $N = 2$,

$$m(x,t) = \sum_{i=1}^{2} p_i(t)\, \delta(x - q_i(t))\,. \tag{13.21}$$

The parameters satisfy the finite dimensional geodesic canonical Hamiltonian equations (13.12), in which the Hamiltonian for $N = 2$ is given by

$$H_{N=2}(q_1, q_2, p_1, p_2) = \frac{1}{2}(p_1^2 + p_2^2) + p_1 p_2\, G(q_1 - q_2)\,. \tag{13.22}$$

13.3.2. *Conservation laws and reduction to quadrature.* Provided the Green's function G is symmetric under spatial reflections, $G(-x) = G(x)$, the two-pulson Hamiltonian system conserves the total momentum

$$P = p_1 + p_2\,. \tag{13.23}$$

Conservation of P ensures integrability, by Liouville's theorem, and reduces the 2-pulson system to quadratures. To see this, we introduce sum and difference variables as

$$P = p_1 + p_2\,, \quad Q = q_1 + q_2\,, \quad p = p_1 - p_2\,, \quad q = q_1 - q_2\,. \tag{13.24}$$

In these variables, the Hamiltonian (13.22) becomes

$$H(q, p, P) = \frac{1}{4}(P^2 - p^2)\big(1 - G(q)\big)\,. \tag{13.25}$$

Likewise, the 2-pulson equations of motion transform to sum and difference variables as

$$\begin{aligned}
\frac{\mathrm{d}P}{\mathrm{d}t} &= -2\frac{\partial H}{\partial Q} = 0\,, & \frac{\mathrm{d}Q}{\mathrm{d}t} &= 2\frac{\partial H}{\partial P} = P(1 + G(q))\,,\\
\frac{\mathrm{d}p}{\mathrm{d}t} &= -2\frac{\partial H}{\partial q} = \frac{1}{2}(p^2 - P^2)G'(q)\,, & \frac{\mathrm{d}q}{\mathrm{d}t} &= 2\frac{\partial H}{\partial p} = -\,p(1 - G(q))\,.
\end{aligned}$$

Eliminating p^2 between the formula for H and the equation of motion for q yields

$$\begin{aligned}
\left(\frac{\mathrm{d}q}{\mathrm{d}t}\right)^2 &= P^2\big(1 - G(q)\big)^2 - 4H\big(1 - G(q)\big)\\
&=: Z(G(q); P, H) \ge 0\,,
\end{aligned} \tag{13.26}$$

which rearranges into the following quadrature,

$$\mathrm{d}t = \frac{\mathrm{d}G(q)}{G'(q)\sqrt{Z(G(q); P, H)}}\,. \tag{13.27}$$

For the peakon case, we have $G(q) = e^q$ so that $G'(q) = G(q)$ and the quadrature (13.27) simplifies to an elementary integral. Having obtained $q(t)$ from the quadrature, the momentum difference $p(t)$ is found from eqn (13.25) via the algebraic expression

$$p^2 = P^2 - \frac{4H}{1 - G(q)}\,, \tag{13.28}$$

in terms of q and the constants of motion P and H. Finally, the sum $Q(t)$ is found by a further quadrature.

Upon writing the quantities H and P as

$$H = c_1 c_2\,, \quad P = c_1 + c_2\,, \quad \frac{1}{2}c_1^2 + \frac{1}{2}c_2^2 = \frac{1}{2}P^2 - H\,, \tag{13.29}$$

in terms of the asymptotic speeds of the pulsons, c_1 and c_2, we find the relative momentum relation,

$$p^2 = (c_1 + c_2)^2 - \frac{4c_1c_2}{1 - G(q)}\,. \tag{13.30}$$

This equation has several implications for the qualitative properties of the 2-pulson collisions.

Definition

13.8. *Overtaking, or rear-end, pulson collisions satisfy $c_1c_2 > 0$, while head-on pulson collisions satisfy $c_1c_2 < 0$.*

The pulson order $q_1 < q_2$ is preserved in an overtaking, or rear-end, collision. This follows, as

Proposition

13.9 (Preservation of pulson order)**.** *For overtaking, or rear-end, collisions, the 2-pulson dynamics preserves the sign condition*

$$q = q_1 - q_2 < 0\,.$$

Proof. Suppose the peaks were to overlap in an overtaking collision with $c_1c_2 > 0$, thereby producing $q = 0$ during a collision. The condition $G(0) = 1$ implies the second term in eqn (13.30) would diverge if this overlap were to occur. However, such a divergence would contradict $p^2 \geq 0$. □

Consequently, seen as a collision between two 'particles' with initial speeds c_1 and c_2 that are initially well separated, the separation $q(t)$ reaches a non-zero distance of closest approach q_{min} in an overtaking, or rear-end, collision that may be expressed in terms of the pulse shape, as follows.

Corollary

13.10 (Minimum separation distance)**.**
The minimum separation distance reachable in two-pulson collisions with $c_1c_2 > 0$ is given by,

$$1 - G(q_{min}) = \frac{4c_1c_2}{(c_1 + c_2)^2}\,. \tag{13.31}$$

Proof. Set $p^2 = 0$ in eqn (13.30). □

Proposition

13.11 (Head-on collisions admit $q \to 0$)**.**
The 2-pulson dynamics allows the overlap $q \to 0$ in head-on collisions.

Proof. Because $p^2 \geq 0$, the overlap $q \to 0$ implying $g \to 1$ is only possible in eqn (13.30) for $c_1c_2 < 0$. That is, for the head-on collisions. □

Remark

13.12 (Divergence of head-on momentum).
Equation (13.30) implies that $p^2 \to \infty$ diverges when $q \to 0$ in head-on collisions. As we shall discuss, this signals the development of a vertical slope in the velocity profile of the solution at the moment of collision.

13.4. Pulson–anti-pulson interactions.

13.4.1. *Head-on pulson–anti-pulson collision.* In a ***completely anti-symmetric*** head-on collision of a pulson and anti-pulson, one has $p_1 = -p_2 = p/2$ and $q_1 = -q_2 = q/2$ (so that $P = 0$ and $Q = 0$). In this case, the quadrature formula (13.27) reduces to

$$\pm\,(t - t_0) = \frac{1}{\sqrt{-4H}} \int_{q(t_0)}^{q(t)} \frac{dq\,'}{\big(1 - G(q\,')\big)^{1/2}}\,, \tag{13.32}$$

and the second constant of motion in eqn (13.25) satisfies

$$-\,4H = p^2\big(1 - G(q)\big) \geq 0\,. \tag{13.33}$$

After the collision, the pulson and anti-pulson separate and travel apart in opposite directions; so that asymptotically in time $g(q) \to 0$, $p \to 2c$, and $H \to -c^2$, where c (or $-c$) is the asymptotic speed (and amplitude) of the pulson (or anti-pulson). Setting $H = -c^2$ in eqn (13.33) gives a relation for the pulson–anti-pulson (p, q) phase trajectories for any kernel,

$$p = \pm\,\frac{2c}{\big(1 - G(q)\big)^{1/2}}\,. \tag{13.34}$$

Notice that p diverges (and switches branches of the square root) when $q \to 0^+$, because $G(0) = 1$. The convention of switching branches of the square root allows one to keep $q > 0$ throughout, so the particles retain their order. That is, the particles 'bounce' elastically at the moment when $q \to 0^+$ in the perfectly anti-symmetric head-on collision in Figure 6.

Remark

13.13 (Preservation of particle identity in collisions).

- *The relative separation distance $q(t)$ in pulson–anti-pulson collisions is determined by following a phase point along a level surface of the Hamiltonian H in the phase space with coordinates (q, p).*
- *Because H is quadratic, the relative momentum p has two branches on such a level surface, as indicated by the $\pm$ sign in eqn (13.34). At the pulson–anti-pulson collision point, both $q \to 0^+$ and either $1/p \to 0^+$ or $p \to 0^+$, so following a phase point through a collision requires that one must choose a convention for which branch of the level surface is taken after the collision.*
- *Taking the convention that p changes sign (corresponding to a **bounce**), but q does not change sign (so the **particles keep their identity**) is convenient, because it allows the phase points to be followed more easily through multiple collisions. This choice is also consistent with the pulson–pulson and anti-pulson–anti-pulson collisions. In these other **rear-end collisions**, as implied by eqn (13.30), the separation distance always remains positive and again the particles retain their identity.*

Theorem

13.14 (Pulson–anti-pulson exact solution).
The exact analytical solution for the pulson–anti-pulson collision for any symmetric G may be written as a function of position x and the separation between the pulses q for any pulse shape or kernel $G(x)$ as

$$u(x, q) = \frac{c}{\big(1 - G(q)\big)^{1/2}}\Big[G(x + q/2) - G(x - q/2)\Big], \tag{13.35}$$

where c is the pulson speed at sufficiently large separation and the dynamics of the separation $q(t)$ is given by the quadrature (13.32) with $\sqrt{-4H} = 2c$.

Proof. The solution for the velocity $u(x,t)$ in the head-on pulson–anti-pulson collision may be expressed in this notation as

$$u(x,t) = \frac{p}{2}G(x+q/2) - \frac{p}{2}G(x-q/2)\,. \tag{13.36}$$

In using eqn (13.34) to eliminate p this solution becomes eqn (13.35). □

Exercise. According to eqn (13.32), how much time is required for the head-on pulson–anti-pulson collision, when $G(q) = \mathrm{e}^{-q^2/2}$ is a Gaussian? ★

Exercise. For the case that $G(x) = \mathrm{e}^{-|x|}$, which is Green's function for the Helmholtz operator in 1D with $\alpha = 1$, show that solution (13.36) for the peakon–anti-peakon collision yields

$$q = -\log \operatorname{sech}^2(ct)\,, \quad p = \frac{\pm 2c}{\tanh(ct)}\,, \tag{13.37}$$

so the peakon–anti-peakon collision occurs at time $t = 0$ and eqn (13.36) results in

$$\begin{aligned} m(x,t) &= u - \alpha^2 u_{xx} \\ &= \frac{2c}{\tanh(ct)}\Big[\delta\Big(x - \frac{1}{2}q(t)\Big) - \delta\Big(x + \frac{1}{2}q(t)\Big)\Big]\,. \end{aligned} \tag{13.38}$$

Discuss the behaviour of this solution. What happens to the slope and amplitude of the peakon velocity just at the moment of impact? ★

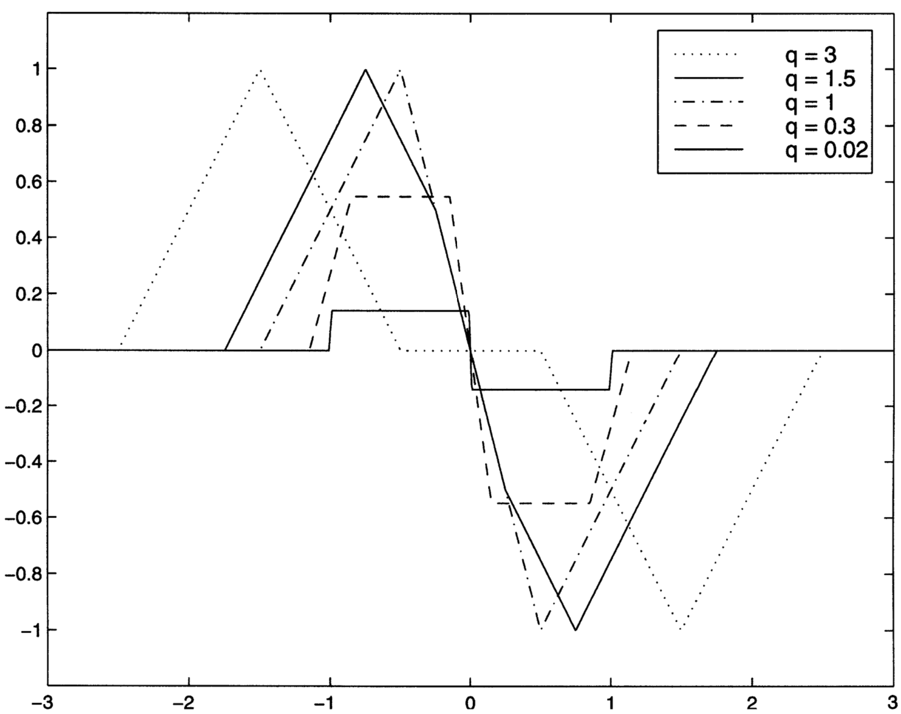


FIGURE 8. Velocity profile (13.35) for the head-on collision of the triangular peakon–anti-peakon pair as a function of separation between the peaks [FH01].

14. Integrability of EPDiff in 1D

In the previous section, we discussed the CH equation for unidirectional shallow-water waves derived in [CH93], as a special case of the b-equation (13.1) with $b = 2$,

$$m_t + um_x + 2mu_x = -c_0 u_x + \gamma u_{xxx}\,, \qquad m = u - \alpha^2 u_{xx}\,. \tag{14.1}$$

This partial differential equation (PDE) describes shallow-water dynamics at quadratic order in the asymptotic expansion for unidirectional shallow-water waves on a free surface under gravity. The previous chapter discussed its elastic particle-collision solution properties in the dispersionless case for which the linear terms on the right side of eqn (14.1) are absent. These elastic-collision solution properties hold for any Green's function $G(x)$ in the convolution relation $u = G * m$ between velocity u and momentum m. For the CH equation $G(x) = \mathrm{e}^{-|x|/\alpha}$ is the Green's function for the 1D Helmholtz operator on the real line with homogeneous boundary conditions.

> This section explains the noncanonical Hamiltonian properties of the CH equation (14.1) in one spatial dimension. In fact, the CH equation has two compatible Hamiltonian structures, so it is ***bi-Hamiltonian***. In this situation, Magri's lemmas for bi-Hamiltonian PDE in 1D imply systematically that CH arises as a different compatibility condition for an ***isospectral eigenvalue problem*** and a linear evolution equation for the corresponding eigenfunctions in the case when $G(x) = \mathrm{e}^{-|x|/\alpha}$. The properties of being bi-Hamiltonian and possessing an associated isospectral problem are ingredients for proving the one-dimensional CH equation (14.1) is ***completely integrable*** as a Hamiltonian system and is solvable by the ***inverse scattering transform (IST) method***.

14.1. The CH equation is bi-Hamiltonian.

The CH equation is ***bi-Hamiltonian***. This means that eqn (14.1) may be written in two compatible Hamiltonian forms, namely as

$$m_t = -B_2\,\frac{\delta H_1}{\delta m} = -B_1\,\frac{\delta H_2}{\delta m}\,, \tag{14.2}$$

where B_1 and B_2 are Poisson operators. For the CH equation, the pairs of Hamiltonians and Poisson operators are given by

$$\begin{aligned}
H_1 &= \frac{1}{2}\int (u^2 + \alpha^2 u_x^2)\,\mathrm{d}x\,,\\
B_2 &= \partial_x m + m\partial_x + c_0\partial_x + \gamma\,\partial_x^3\,, \qquad (14.3)\\
H_2 &= \frac{1}{2}\int u^3 + \alpha^2 u\,u_x^2 + c_0 u^2 - \gamma\,u_x^2\,\mathrm{d}x\,,\\
B_1 &= \partial_x - \alpha^2\partial_x^3\,. \qquad (14.4)
\end{aligned}$$

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,

$$\{f, h\}(m) = -\int \frac{\delta f}{\delta m}(c_1 B_1 + c_2 B_2)\frac{\delta h}{\delta m}\,\mathrm{d}x\,, \tag{14.5}$$

as a bilinear skew-symmetric operation that satisfies the Jacobi identity. (In general, the sum of the Poisson brackets would fail to satisfy 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 [Olv00]. This fact was already known in the literature for KdV, see [Fuc96].

14.2. Magri's lemmas.

The property of ***compatibility*** of the two Hamiltonian operators for a bi-Hamiltonian equation enables the construction under certain conditions of an infinite hierarchy of Poisson-commuting Hamiltonians. The property of compatibility was used by Magri [Mag78] in proving the following important pair of lemmas (see also [Olv00] for a clear discussion of Magri's lemmas):

Lemma

14.1 (Magri 1978)**.** *If B_1 and B_2 are compatible Hamiltonian operators, with B_1 non-degenerate, and if*

$$B_2\frac{\delta H_1}{\delta m} = B_1\frac{\delta H_2}{\delta m} \quad \textit{and} \quad B_2\frac{\delta H_2}{\delta m} = B_1\mathcal{K}\,, \tag{14.6}$$

for Hamiltonians H_1, H_2, and some function $\mathcal{K}$, then there exists a third Hamiltonian functional H such that $\mathcal{K} = \delta H/\delta m$.

To prove the existence of an infinite hierarchy of Hamiltonians, H_n, $n = 1, 2, \ldots$, related to the two compatible Hamiltonian operators B_1, B_2, we need to check that the following two conditions hold:

: (i) There exists an infinite sequence of functions $\mathcal{K}_1$, $\mathcal{K}_2$, ... satisfying

$$B_2\mathcal{K}_n = B_1\mathcal{K}_{n+1}\,; \tag{14.7}$$

: (ii) There exist two functionals H_1 and H_2 such that

$$\mathcal{K}_1 = \frac{\delta H_1}{\delta m}\,, \qquad \mathcal{K}_2 = \frac{\delta H_2}{\delta m}\,. \tag{14.8}$$

It then follows from Lemma 14.1 that there exist functionals H_n such that

$$\mathcal{K}_n = \frac{\delta H_n}{\delta m}\,, \quad \text{for all} \quad n \geq 1\,. \tag{14.9}$$

Lemma

14.2 (Magri 1978)**.** *Let $\{\,\cdot\,,\,\cdot\,\}_1$ and $\{\,\cdot\,,\,\cdot\,\}_2$ denote the Poisson brackets defined, respectively, by B_1 and B_2, which are assumed to be compatible Hamiltonian operators. Let H_1, H_2, ... be an infinite sequence of Hamiltonian functionals constructed from eqns (14.7) and (14.9). Then, these Hamiltonian functionals mutually commute under both Poisson brackets:*

$$\{\,H_m,\,H_n\,\}_1 = \{\,H_m,\,H_n\,\}_2 = 0\,, \quad \textit{for all} \quad m, n \geq 1\,. \tag{14.10}$$

Definition

14.3. *A set of functionally independent Hamiltonians that Poisson-commute among themselves is said to be **in involution**.*

Remark

14.4. *The condition for a canonical Hamiltonian system with N degrees of freedom to be **completely integrable** is that it possess N constants of motion in involution. The bi-Hamiltonian property is important because it produces the corresponding condition for an infinite-dimensional system. The infinite-dimensional case introduces additional questions, such as the completeness of the infinite set of independent constants of motion in involution. However, such questions are beyond our present scope.*

14.3. **Applying Magri's lemmas.** The bi-Hamiltonian property of eqn (14.1) allows one to construct an infinite number of Poisson-commuting conservation laws for it by applying Magri's lemmas. According to [Mag78], these conservation laws may be constructed for non-degenerate B_1 by defining the transpose operator $R^T = B_1^{-1}B_2$ that leads from the variational derivative of one conservation law to the next,

$$\frac{\delta H_n}{\delta m} = R^T\frac{\delta H_{n-1}}{\delta m}\,, \quad n = -1, 0, 1, 2, \ldots. \tag{14.11}$$

The operator $R^T = B_1^{-1}B_2$ recursively takes the variational derivative of H_{-1} to that of H_0, to that of H_1, then to that of H_2, etc. 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*** $R = B_2B_1^{-1}$ leads to a hierarchy of commuting flows, defined by $\mathcal{K}_{n+1} = R\mathcal{K}_n$, for $n = 0, 1, 2, \dots$,

$$\begin{aligned} m_t^{(n+1)} &= \mathcal{K}_{n+1}[m] = -B_1\,\frac{\delta H_n}{\delta m} \\ &= -\,B_2\frac{\delta H_{n-1}}{\delta m} = B_2B_1^{-1}\mathcal{K}_n[m]\,. \end{aligned} \tag{14.12}$$

The first three flows in the 'positive hierarchy' when $c_0, \gamma \to 0$ are

$$m_t^{(1)} = 0\,, \quad m_t^{(2)} = -\,m_x\,, \quad m_t^{(3)} = -\,(m\partial + \partial m)u\,, \tag{14.13}$$

the third being EPDiff. The next flow is too complicated to be usefully written here. However, by Magri's 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 in 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 second Hamiltonian in the negative hierarchy of EPDiff is given by

$$m_t = B_1\,\frac{\delta H_{-1}}{\delta m} = B_2\,\frac{\delta H_{-2}}{\delta m}\,, \tag{14.14}$$

which gives

$$H_{-2} = \frac{1}{2}\int_{-\infty}^{\infty}\left[\frac{\alpha^2}{4}\,\frac{m_x^2}{m^{5/2}} - \frac{2}{\sqrt{m}}\right]. \tag{14.15}$$

The flow defined by eqn (14.14) is

$$m_t = -(\partial - \alpha^2\partial^3)\left(\frac{1}{2\sqrt{m}}\right). \tag{14.16}$$

For $m = u - \alpha^2 u_{xx}$, this flow is similar to the Dym equation,

$$u_{xxt} = \partial^3\left(\frac{1}{2\sqrt{u_{xx}}}\right), \tag{14.17}$$

which is also a completely integrable soliton equation [AS06].

14.4. **The CH equation is isospectral.** The isospectral eigenvalue problem associated with eqn (14.1) may be found by using the recursion relation of the bi-Hamiltonian structure, following a standard technique due to Gelfand and Dorfman [GD79]. Let us introduce a spectral parameter λ and multiply by λ^n the nth step of the recursion relation (14.12), then taking the sum yields

$$B_1\sum_{n=0}^{\infty}\lambda^n\frac{\delta H_n}{\delta m} = \lambda B_2\sum_{n=0}^{\infty}\lambda^{(n-1)}\frac{\delta H_{n-1}}{\delta m}\,, \tag{14.18}$$

or, by introducing the squared-eigenfunction ψ^2

$$\psi^2(x,t;\lambda) := \sum_{n=0}^{\infty}\lambda^n\frac{\delta H_n}{\delta m}\,, \tag{14.19}$$

one finds, formally,

$$B_1\psi^2(x,t;\lambda) = \lambda B_2\psi^2(x,t;\lambda)\,. \tag{14.20}$$

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.

Proposition

14.5. *If ψ satisfies*

$$\lambda\left(\frac{1}{4} - \alpha^2\partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi\,, \tag{14.21}$$

then ψ^2 is a solution of eqn (14.20).

Proof. This is is straightforward computation. □

Now, assuming that λ will be independent of time, we seek, in analogy with the KdV equation, an evolution equation for ψ of the form,

$$\psi_t = a\psi_x + b\psi \,, \tag{14.22}$$

where a and b are functions of u and its derivatives. These functions are determined from the requirement that the ***compatibility condition*** $\psi_{xxt} = \psi_{txx}$ between eqns (14.21) and (14.22) implies eqn (14.1). Cross-differentiation shows

$$b = -\frac{1}{2}a_x \,, \quad \text{and} \quad a = -(\lambda + u)\,. \tag{14.23}$$

Consequently,

$$\psi_t = -(\lambda + u)\psi_x + \frac{1}{2}u_x\psi \,, \tag{14.24}$$

is the desired evolution equation for the eigenfunction ψ.

Summary of the isospectral property of eqn (14.1).
The Gelfand–Dorfman theory [GD79] determines the isospectral problem for integrable equations via the squared-eigenfunction approach. Its bi-Hamiltonian property implies that the nonlinear shallow-water wave equation (14.1) arises as a compatibility condition for two linear equations. These are the ***isospectral eigenvalue problem***,

$$\lambda\left(\frac{1}{4} - \alpha^2\partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi \,, \tag{14.25}$$

and the ***evolution equation*** for the eigenfunction ψ,

$$\psi_t = -(u + \lambda)\,\psi_x + \frac{1}{2}u_x\,\psi \,. \tag{14.26}$$

Compatibility of these linear equations ($\psi_{xxt} = \psi_{txx}$) together with isospectrality ($\mathrm{d}\lambda/\mathrm{d}t = 0$) imply the CH equation,

$$m_t + um_x + 2mu_x = -c_0u_x + \gamma u_{xxx} \,, \qquad m = u - \alpha^2 u_{xx} \,. \tag{14.27}$$

Remark

14.6 (Implications of Isospectrality).

- *The isospectral eigenvalue problem (14.25) for the nonlinear CH water-wave equation (14.27) restricts to the isospectral problem for KdV (namely, the Schrödinger equation) when $\alpha^2 \to 0$.*
- *The evolution equation (14.26) for the isospectral eigenfunctions in the cases of KdV and CH are identical.*
- *The isospectral eigenvalue problem and the evolution equation for its eigenfunctions are two linear equations whose compatibility implies a nonlinear equation for the unknowns in the KdV and CH equations.*
- *This formulation for the KdV equation led to the famous method of the* ***inverse scattering transform (IST)*** *for the solution of its initial-value problem, reviewed, e.g., in* [AS06].
- *The CH equation also admits the IST solution approach, but for a different isospectral eigenvalue problem that limits to the Schrödinger equation when $\alpha^2 \to 0$. The isospectral eigenvalue problem (14.25) for CH arises in the study of the fundamental oscillations of a non-uniform string under tension.*

EPDiff(H^1) is the dispersionless case of CH. In the dispersionless case $c_0 = 0 = \gamma$, the shallow-water equation (14.1) becomes the 1D geodesic equation EPDiff(H^1)

$$m_t + um_x + 2mu_x = 0\,, \qquad m = u - \alpha^2 u_{xx}\,. \tag{14.28}$$

The solitary travelling-wave solution of 1D EPDiff (14.28) in this dispersionless case is the ***peakon***,

$$u(x,t) = c\,G(x-ct) = \frac{c}{2\alpha}\,\mathrm{e}^{-|x-ct|/\alpha}\,.$$

The EPDiff equation (12.1) may also be written as a conservation law for momentum,

$$\partial_t m = -\partial_x\Big(um + \frac{1}{2}u^2 - \frac{\alpha^2}{2}u_x^2\Big)\,. \tag{14.29}$$

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 (14.25) in the dispersionless case $c_0 = 0 = \Gamma$ has a ***purely discrete spectrum*** on the real line and the N-soliton solutions for this equation may be expressed as a ***peakon wave train***,

$$u(x,t) = \sum_{i=1}^{N} p_i(t)\mathrm{e}^{-|x-q_i(t)|/\alpha}\,. \tag{14.30}$$

As before, $p_i(t)$ and $q_i(t)$ satisfy the finite-dimensional geodesic motion equations obtained as Hamilton's canonical equations

$$\dot q_i = \frac{\partial H_N}{\partial p_i} \quad \text{and} \quad \dot p_i = -\frac{\partial H_N}{\partial q_i}\,, \tag{14.31}$$

where the Hamiltonian is given by,

$$H_N = \frac{1}{2}\sum_{i,j=1}^{N} p_i\,p_j\,\mathrm{e}^{-|q_i-q_j|/\alpha}\,. \tag{14.32}$$

Thus, we have proved the following.

Theorem

14.7. *CH peakons are an integrable subcase of EPDiff pulsons in one dimension for the choice of the H^1 norm.*

Remark

14.8. *The discrete process of peakon creation via the steepening lemma 12.2 is consistent with the discreteness of the isospectrum for the eigenvalue problem (14.25) 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.*

Constants of motion for integrable N-peakon dynamics. One may verify the integrability of the N-peakon dynamics by substituting the N-peakon solution (14.30) (which produces the sum of delta functions (12.6) for the momentum m) into the isospectral problem (14.25). This substitution reduces (14.25) to an $N \times N$ matrix eigenvalue problem.

In fact, the canonical equations (14.31) for the peakon Hamiltonian (14.32) may be written directly in Lax matrix form,

$$\frac{\mathrm{d}L}{\mathrm{d}t} = [L, A] \quad \Longleftrightarrow \quad L(t) = U(t)L(0)U^\dagger(t)\,, \tag{14.33}$$

with $A = \dot U U^\dagger(t)$ and $UU^\dagger = Id$. Explicitly, L and A are $N \times N$ matrices with entries

$$L_{jk} = \sqrt{p_j p_k}\,\phi(q_i - q_j)\,, \quad A_{jk} = -2\sqrt{p_j p_k}\,\phi'(q_i - q_j)\,. \tag{14.34}$$

Here, $\phi'(x)$ denotes derivative with respect to the argument of the function ϕ, given by $\phi(x) = \mathrm{e}^{-|x|/2\alpha} = 2\alpha G(x/2)$. The Lax matrix L in eqn (14.33) evolves by time-dependent unitary transformations, which leave its spectrum invariant. Isospectrality then implies that the traces $\mathrm{tr}\, L^n$, $n = 1, 2, \ldots, N$ of the powers of the matrix L (or, equivalently, its N eigenvalues) yield N constants of the motion. These turn out to be functionally independent, non-trivial and in involution under the canonical Poisson bracket. Hence, the canonically Hamiltonian N-peakon dynamics (14.31) is completely integrable in the finite-dimensional (Liouville) sense.

Exercise. Verify that the compatibility condition (equality of cross derivatives $\psi_{xxt} = \psi_{txx}$) obtained from the eigenvalue equation (14.25) and the evolution equation (14.26) do indeed yield the CH shallow-water wave equation (14.1) when the eigenvalue λ is constant. ★

Exercise. Show that the peakon Hamiltonian H_N in (14.32) may be expressed as a function of the invariants of the matrix L, as

$$H_N = -\mathrm{tr}\, L^2 + 2(\mathrm{tr}\, L)^2 \,. \tag{14.35}$$

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 that no difficulties arise, either due to the non-analyticity of $\phi(x)$, or the sign in the square roots in the Lax matrices L and A. ★

Hunter–Saxton equation. Retrace the progress of this chapter for the EPDiff equation

$$m_t + u m_x + 2 m u_x = 0 \,, \quad \text{with} \quad m = -u_{xx} \,. \tag{14.36}$$

This integrable Hamiltonian partial differential equation arises in the theory of liquid crystals. Its peakon solutions are the compactly supported triangles in Figure 7 and Figure 8. It may also be regarded as the $\alpha \to \infty$ limit of the CH equation. For more results and discussion of this equation, see [HZ94]. ★

References

[AS06] Ablowitz, M. J. and Segur, H. [2006] *Solitons and the Inverse Scattering Transform.* Cambridge University Press, New York.

[AbMa1978] Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, 2nd ed. Reading, MA: Addison-Wesley.

[BlIs2006] Bloch, A. M. and Iserles, A. [2006] On an isospectral Lie-Poisson system and its Lie algebra. *Found. Comput. Math.* **6**, 121–144.

[Bo1985] Bogoyavlensky, O. I. [1985] Integrable Euler equations on Lie algebras arising in problems of mathematical physics. *Math. USSR Izvestiya* **25**, 207–257.

[CH93] Camassa,R. and Holm, D. D. [1993] An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71:1661–64.

[LDP2000] *Memoirs of Lorenzo Da Ponte*, [2000] Translated by E. Abbott; edited and annotated by A. Livingston; introduction by C. Rosen. The New York Review of Books, New York.

[Da1994] Darling, R. W. R. [1994] *Differential Forms and Connections.* Cambridge: Cambridge University Press.

[DHH03] Degasperis, A., Holm, D. D. and Hone,A. N. W. [2003] Integrable and nonintegrable equations with peakons. In M. J. Ablowitz, M. Boiti, F. Pempinelli, and B. Prinari, editors, *Nonlinear Physics: Theory and Experiment (Gallipoli 2002) Vol II*, pages 37–43. World Scientific, Singapore.

[doCa1976] do Carmo, M. P. [1976] *Differential Geometry of Curves and Surfaces.* Englewood Cliffs, NJ: Prentice-Hall.

[DGH04] Dullin, H., Gottwald, G. and Holm, D. D. [2004] On asymptotically equivalent shallow water wave equations. *Physica D*, 190:1–14.

[EM70] Ebin, D. G. and Marsden, J. E. [1970] Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math.*, 92:102–163.

[FaTa1987] Faddeev, L. D. and Takhtajan, L. A. [1987] *Hamiltonian Methods in the Theory of Solitons.* New York: Springer.

[Fl1963] Flanders, H. [1963] *Differential Forms with Applications to the Physical Sciences.* New York: Academic Press.

[FH01] Fringer, O. B. and Holm, D. D. [2001] Integrable vs nonintegrable geodesic soliton behavior. *Physica D*, 150:237–263.

[Fuc96] Fuchssteiner, B. [1996] Some tricks from the symmetry-toolbox for nonlinear equations: Generalizations of the Camassa-Holm equation. *Physica D*, 95:229–243.

[GBRa2008] Gay-Balmaz, F. and Ratiu, T.S. [2008] The geometric structure of complex fluids. *Adv. in Appl. Math.* **42** (2), 176–275.

[GaVi2010] Gay-Balmaz F. and Vizman, C. [2011] Dual pairs in fluid dynamics. *Ann. Glob. Anal. Geom.* **41**, 1–24.

[GD79] Gelfand, I. M. and Dorfman, I. Ya. R. [1979] Hamiltonian operators and algebraic structures associated with them. *Funct. Anal. Appl.*, 13:248–254.

[GiHoKu1982] Gibbons, J., Holm, D. D. and Kupershmidt, B. A. [1982] Gauge-invariant Poisson brackets for chromo-hydrodynamics. *Phys. Lett. A* **90**, 281–283.

[GrKrMa1988] Grossman, R., Krishnaprasad, P. S. and Marsden, J. E. [1988] The dynamics of two coupled rigid bodies. In *Dynamical Systems Approaches to Nonlinear Problems in Systems and Circuits*, edited by F. M. Abdel Salam and M. Levi, Engineering Foundation, United States Army Research Office, and U.S. National Science Foundation. Philadelphia, PA: SIAM, pp. 373–378.

[Ho2002] Holm, D. D. [2002] Euler–Poincaré dynamics of perfect complex fluids. In *Geometry, Mechanics, and Dynamics: In Honor of the 60th Birthday of Jerrold E. Marsden*, edited by P. Newton, P. Holmes and A. Weinstein. New York: Springer, pp. 113–167.

[Ho2005] Holm, D. D. [2005] The Euler–Poincaré variational framework for modeling fluid dynamics. In *Geometric Mechanics and Symmetry: The Peyresq Lectures*, edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306. Cambridge: Cambridge University Press.

[Ho2011GM] Holm, D. D. [2011]
Geometric Mechanics I: Dynamics and Symmetry,
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.

[HoIvPe2012] Holm, D. D., Ivanov, R. I. and Percival, J. R. [2012] *G*-Strands, *J. Nonlin. Sci.*, to appear.

[HoKu1988] Holm, D. D. and Kupershmidt, B. A. [1988] The analogy between spin glasses and Yang–Mills fluids. *J. Math Phys.* **29**, 21–30.

[HoMa2004] Holm, D. D. and Marsden, J. E. [2004] Momentum maps and measure valued solutions (peakons, filaments, and sheets) for the EPDiff equation. In *The Breadth of Symplectic and Poisson Geometry*, edited by J. Marsden, and T. Ratiu. Boston: Birkhäuser. arXiv.nlin.CD/0312048.

[HoMaRa1998] Holm, D. D., Marsden, J. E. and Ratiu, T. S. [1998] The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* **137**, 1–81.

[HoSmSt2009] Holm, D. D., Schmah, T. and Stoica, C. [2009]
Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions,
Oxford University Press.

[HS03] Holm, D. D. and Staley, M. F. [2003] Wave structures and nonlinear balances in a family of 1+1 evolutionary pdes. *SIAM J. Appl. Dyn. Syst.*, 2:323–380.

[HoVi2010] Holm, D. D. and Vizman, C. [2010] *The n:m resonance dual pair*. Preprint at http://arxiv.org/abs/1102.4377.

[Ho1931] Hopf, H. [1931] Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche. *Math. Ann.* **104**, 637–665.

[HZ94] Hunter, J. K. and Zheng, Y. [1994] On a completely integrable nonlinear hyperbolic variational equation. *Physica D*, 79:361–386.

[Iv2006] Ivanov, R. [2006] Hamiltonian formulation and integrability of a complex symmetric nonlinear system. *Phys. Lett. A* **350**, 232–235.

[JKLOR2011] Jia, S., Kirsten, J., Lai, E., Ortmann, K. and Raithatha, A. [2011] An Application (and Translation) of Henri Poincaré's Paper, "*Sur une forme nouvelle des équations de la mécanique*" [Po1901].

[Ku1978] Kummer, M. [1978] On resonant classical Hamiltonians with two equal frequencies. *Commun. Math. Phys.* **58**, 85–112.

[Mag78] Magri, F. [1978] A simple model of the integrable Hamiltonian equation. *J. Math. Phys.*, 19:1156–1162.

[Man1976] Manakov, S. V. [1976] Note on the integration of Euler's equations of the dynamics of and n-dimensional rigid body. *Funct. Anal. Appl.* **10**, 328–329.

[MaRa1994] Marsden, J. E. and Ratiu, T. S. [1994]
Introduction to Mechanics and Symmetry.
Texts in Applied Mathematics, Vol. 75. New York: Springer-Verlag.

[MaRaWe1984a] Marsden, J. E., Ratiu, T. S. and Weinstein, A. [1984a] Semidirect products and reduction in mechanics. *Trans. Amer. Math. Soc.* **281**(1), 147–177.

[MaRaWe1984b] Marsden, J. E., Ratiu, T. S. and Weinstein, A. [1984b] Reduction and Hamiltonian structures on duals of semidirect product Lie algebras. *Contemporary Math.* **28**, 55–100.

[MN02] Mikhailov, A. V. and Novikov, V. S. [2002] Perturbative symmetry approach. *J. Phys. A*, 35:4775–4790.

[Ne1997] Needham, T. [1997] *Visual Complex Analysis.* Oxford: Clarendon Press.

[OcoRo1998] O'Connor, J. J. and Robertson, E. F. [1998] *Sir William Rowan Hamilton.* Available at http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Hamilton.html.

[Olv00] Olver, P. J. [2000] *Applications of Lie Groups to Differential Equations.* Springer, New York.

[OrRa2004] Ortega, J.-P. and Ratiu, T. S. [2004] *Momentum Maps and Hamiltonian Reduction.* Progress in Mathematics, Vol. 222. Boston: Birkhäuser.

[Po1892] Poincaré, H. [1892] *Théorie Mathématique de la Lumière*, Paris: Georges Carré, p. 275.

[Po1901] Poincaré, H. [1901] Sur une forme nouvelle des équations de la méchanique,
C.R. Acad. Sci. **132** 369-371.

[Ra1982] Ratiu, T. [1982] Euler–Poisson equations on Lie algebras and the N-dimensional heavy rigid body. *Am. J. Math.* **104**, 409–448.

[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.

[RaVM1982] Ratiu, T. and van Moerbeke, P. [1982] The Lagrange rigid body motion. *Annales de l´institut Fourier* **32**(1), 211–234.

[Sk1961] Skyrme, T. H. R. [1961] A nonlinear field theory. *Proc. R. Soc. A* **260**, 127–138.

[Tay96] Taylor, M. E. [1996] *Partial Differential Equations I.* Springer-Verlag, New York.

[We1983a] Weinstein, A. [1983a] Sophus Lie and symplectic geometry. *Expo. Math.* **1**, 95–96.

[We1983b] Weinstein, A. [1983b] The local structure of Poisson manifolds. *J. Differ. Geom.* **18**, 523–557.

[We2002] Weinstein, A. [2002] Geometry of momentum. ArXiv:math/SG0208108 v1.

Index