M3-4-5 A34 Handout: Dynamics, Symmetry and Integrability

Professor Darryl D Holm   Spring Term 2017
Imperial College London   d.holm@ic.ac.uk
`http://wwwf.imperial.ac.uk/~dholm/classnotes/`

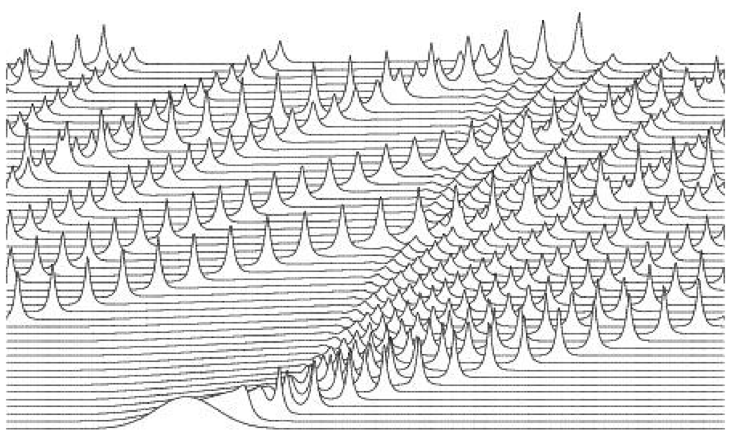
Course meets $\boxed{\text{12-2pm Tues @ Hux 342 \& 1-2pm Thurs @ Hux 140}}$

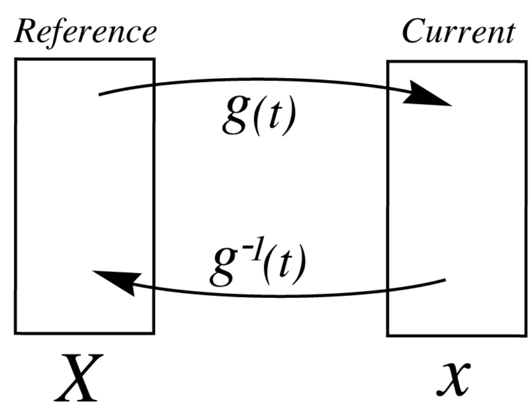
10% of mark from 3 assessed homeworks of about 5 questions at 3-week intervals, Exam taken mainly from these.

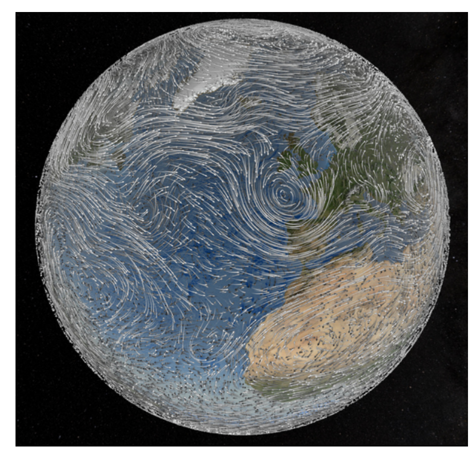
Primary References:

[HoScSt2009] *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*, by DD Holm,T Schmah and C Stoica, Oxford University Press, (2009). ISBN 978-0-19-921290-3

[Ho2011GM2] *Geometric Mechanics II: Rotating, Translating & Rolling* (aka GM2) by DD Holm, World Scientific: Imperial College Press, 2*nd* edition (2011). ISBN 978-1-84816-777-3







Abstract

Classical mechanics, one of the oldest branches of science, has undergone a long evolution, developing hand in hand with many areas of mathematics, including calculus, differential geometry, and the theory of Lie groups and Lie algebras. The modern formulations of Lagrangian and Hamiltonian mechanics, in the coordinate-free language of differential geometry, are elegant and general. They provide a unifying framework for many seemingly disparate physical systems, such as n-particle systems, rigid bodies, fluids and other continua, and electromagnetic and quantum systems.

This course on Geometric Mechanics and Symmetry is a friendly and fast-paced introduction to the geometric approach to classical mechanics, suitable for PhD students or advanced undergraduates. It fills a gap between traditional classical mechanics texts and advanced modern mathematical treatments of the subject. After a summary of the setting of mechanics using variational calculus of transformations on smooth manifolds and basic Lie group theory illustrated in matrix multiplication, the rest of the course considers how symmetry reduction of Hamilton's principle allows one to derive and analyze the Euler-Poincaré equations for dynamics generated by the actions of Lie groups on manifolds. The main topics are shallow water waves, ideal incompressible fluid dynamics and geophysical fluid dynamics (GFD).

Three worked examples (rigid body rotation, shallow water solitons and ideal fluid dynamics) illustrating the course material are given in full detail in the course notes. These will be assigned as outside reading and then discussed in both taught lectures and Q&A sessions in class.

Contents

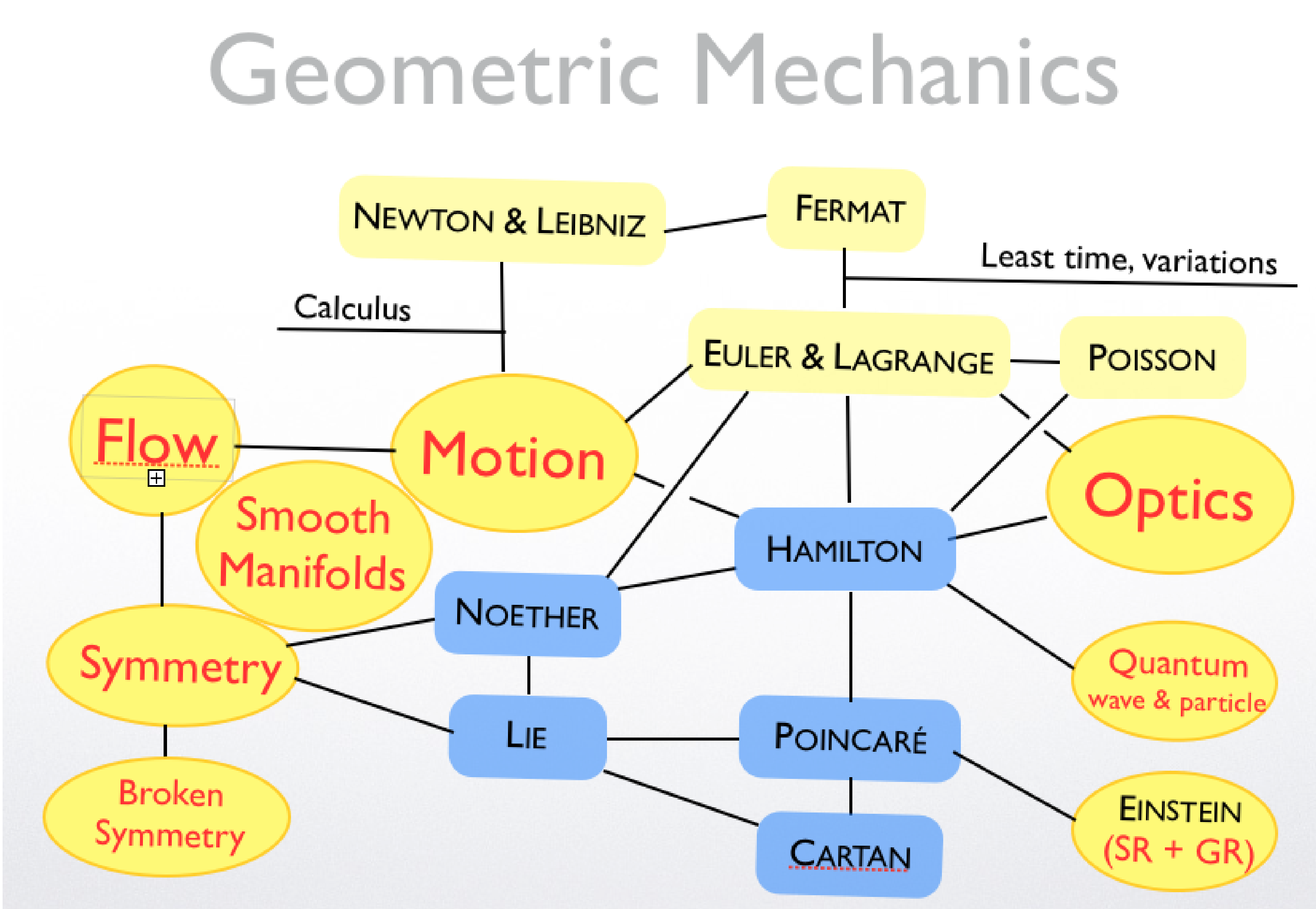


Figure 1: Geometric Mechanics has involved many great mathematicians!

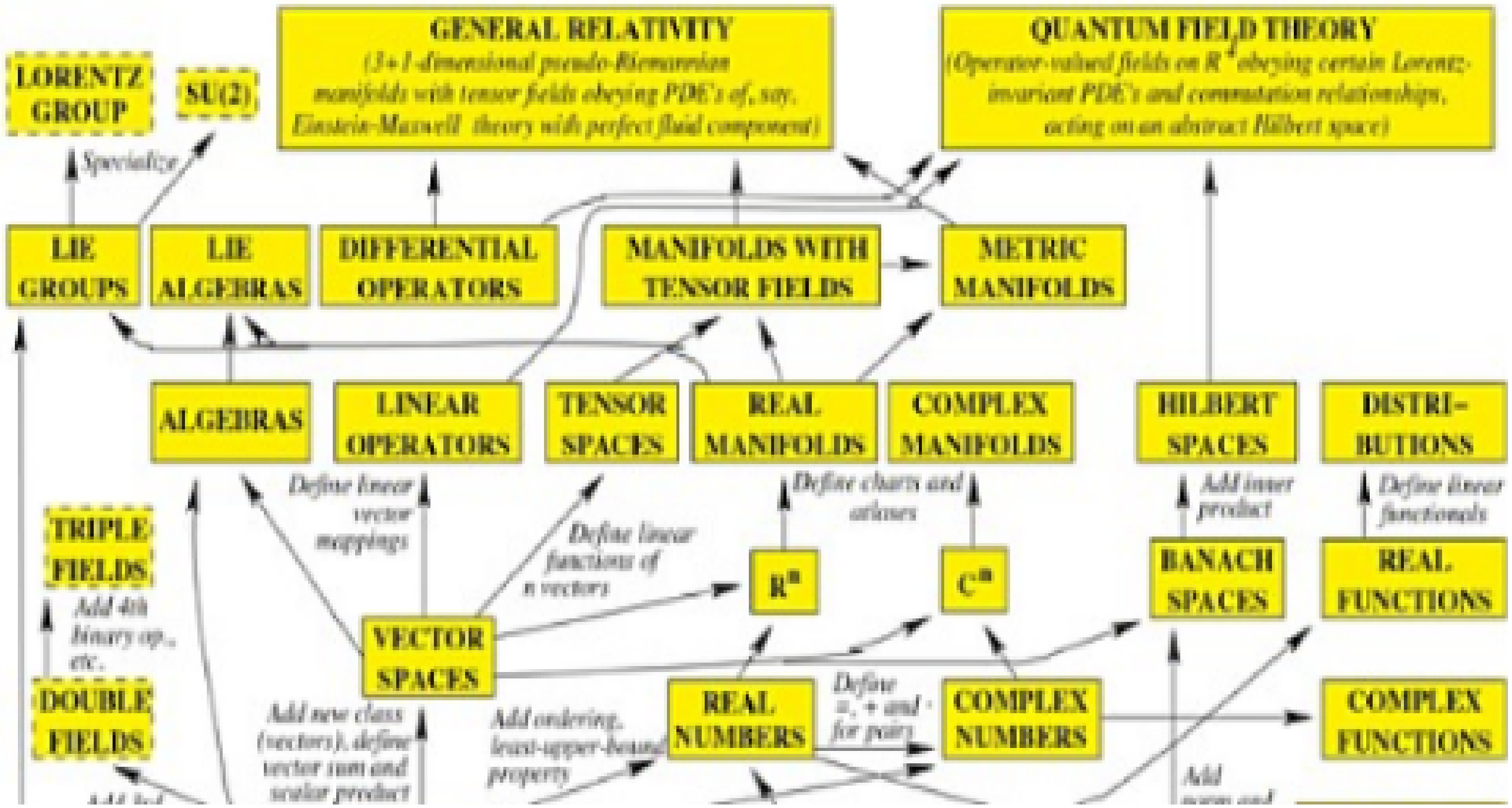


Figure 2: Geometric Mechanics involves many related fields of mathematics which are also fundamental in physics! Here are some definitions: An algebra is a vector space which is also a ring (a set with two binary operations: addition and multiplication); in finite dimensions, a Lie group is a group which is also a manifold (a space which is locally $\mathbb{R}^n$ or $\mathbb{C}^n$); and a Lie algebra is the vector space given by the tangent space at the identity element of a Lie group. A Lie algebra inherits a nonassociative product called a Lie bracket; denoted as the commutator $[x, y] = xy - yx$ for a finite-dimensional Lie algebra's vector space operations. In infinite dimensions, Lie-algebra vector spaces become vector fields; so differential operators and function spaces are involved. See M. Tegmark, *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*, Figure 12.1.

1 Introduction

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

Background reading: Chapter 2 of [Ho2011GM1].

Time

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

Space

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

As a special case, consider the motion of a particle at position $\mathbf{q}(t) \in \mathbb{R}^3$ that is constrained to move on a sphere. This motion may be expressed as time-dependent rotations $O(t) \in SO(3)$ such that

$$\mathbf{q}(t) = O(t)\mathbf{q}(0)\,, \quad \dot{\mathbf{q}}(t) = \dot{O}(t)\mathbf{q}(0) = \dot{O}O^{-1}(t)\mathbf{q}(t) = \widehat{\omega}(t)\mathbf{q}(t) =: \boldsymbol{\omega}(t) \times \mathbf{q}(t)$$

with 3×3 antisymmetric matrix

$$\widehat{\omega}(t) = \dot{O}O^{-1}(t) = -\,\widehat{\omega}(t)^T \quad \text{since} \quad O^{-1} = O^T \quad \text{so that} \quad 0 = \frac{d}{dt}(OO^T) = \dot{O}O^T + (\dot{O}O^T)^T = \widehat{\omega} + \widehat{\omega}^T$$

Motion

Motion is a map $\phi_t : T \to Q$, where subscript t denotes dependence on time t. For example, when $T = \mathbb{R}$, the motion is a curve $q_t = \phi_t \circ q_0$ obtained by composition of functions.

The motion is called a *flow* if $\phi_{t+s} = \phi_t \circ \phi_s$, for $s, t \in \mathbb{R}$, and $\phi_0 = \mathrm{Id}$, so that $\phi_t^{-1} = \phi_{-t}$. Note that the composition of functions is associative, $(\phi_t \circ \phi_s) \circ \phi_r = \phi_t \circ (\phi_s \circ \phi_r) = \phi_t \circ \phi_s \circ \phi_r = \phi_{t+s+r}$, but it is not commutative, in general. Thus, we should anticipate flows that arise as Lie group actions on manifolds.

We have already seen the example of $\mathbf{q}_t = O(t)\mathbf{q}_0$ for the action of $O(t) \in SO(3)$ on the manifold $Q = \mathbb{R}^3$.

Velocity

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

Motion equation

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

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

where $f(q)$ is a prescribed *vector field* over Q. For example, if the curve $q_t = \phi_t \circ q_0$ is a *flow* (that is, $\phi_t \circ \phi_s = \phi_{t+s}$), then

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

so that

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

which defines the *pullback* of f by ϕ_t.

Optimal motion equation – Hamilton's principle

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

$$\delta S[q_t] = 0 \quad \text{for} \quad S[q_t] = \int_{t_0}^{t_1} L(q_t, \dot{q}_t)\, dt\,,$$

in which variational derivatives are given by

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

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

For example,

$$\text{if} \quad S = \int_{t_0}^{t_1} L(q, \dot{q})\, dt\,, \quad \text{then} \quad \delta S = \int_{t_0}^{t_1} \left\langle \frac{\partial L}{\partial \dot{q}_t}\,,\, \delta \dot{q}_t) \right\rangle + \left\langle \frac{\partial L}{\partial q_t}\,,\, \delta q_t) \right\rangle dt = 0$$

leads to the *Euler-Lagrange equations*

$$-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_t} + \frac{\partial L}{\partial q_t} = 0\,, \quad \text{when} \quad \left\langle \frac{\partial L}{\partial \dot{q}_t}\,,\,\delta q \right\rangle\Big|_{t_0}^{t_1} = 0\,.$$

The endpoint term yields **Noether's theorem**, when $\delta q = \pounds_\xi q$ is an infinitesimal Lie symmetry of the Lagrangian, so that $\pounds_\xi L(q,\dot{q}) = 0$.

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

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

In this case, the *phase-space action principle*

$$0 = \delta \int_{t_0}^{t_1} \langle\, p\,,\,\dot{q}\,\rangle - H(q,p)\,dt\,,$$

gives *Hamilton's canonical equations*

$$\dot{q} = H_p \quad \text{and} \quad \dot{p} = -\,H_q\,, \quad \text{with} \quad \langle\, p\,,\,\delta q\,\rangle|_{t_0}^{t_1} = 0\,,$$

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

Exercise. Derive Hamilton's canonical equations from the the phase-space action principle. ★

Symmetry

Lie group symmetries of the Lagrangian will be particularly important, both in reducing the number of independent degrees of freedom in Hamilton's principle and in finding conservation laws by Noether's theorem. According to Noether's theorem, each Lie group symmetry of the Lagrangian has a corresponding conserved quantity under evolution by the Euler-Lagrange equations.

Dynamics!

Dynamics is the science of deriving, analysing, solving and interpreting the solutions of motion equations. The main ideas of our course will often be illuminated by considering dynamics in the example that the configuration manifold Q is a Lie group itself G and the Lagrangian $TG \to \mathbb{R}$ transforms simply (e.g., is invariant) under the action of G. When the Lagrangian $TG \to \mathbb{R}$ is *invariant* under G, the dynamics may be reformulated for a symmetry-reduced Lagrangian defined on $TG/G \simeq \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of the Lie group G. With an emphasis on applications in mechanics, we will discuss a variety of interesting properties and results that are inherited from this formulation of dynamics on Lie groups.

What shall we study?

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

> Lie: Groups of transformations that depend smoothly on parameters
> Poincaré: Mechanics on Lie groups, e.g., $SO(3)$, $SU(2)$, $Sp(2)$, $SE(3) \simeq SO(3)Ⓢ\mathbb{R}^3$
> Noether: Implications of symmetry in variational principles

These accomplishments lead to a new view of dynamics. In particular, Poincaré's view of it lead to mechanics on Lie groups.

2 Elements of Transformation theory for motion on smooth manifolds

Vocabulary

smooth manifold	equilibrium	tangent lift
tangent space	linearisation	commutator
motion equation	infinitesimal transformation	differential, d
vector field	pull-back	differential k-form
diffeomorphism	push-forward	wedge product, $\wedge$
flow	Jacobian matrix	Lie derivative, $\pounds_Q$
fixed point	directional derivative	product rule

- Let M be a *smooth manifold*, $\dim M = n$. That is, M is a smooth space that is locally isomorphic to $\mathbb{R}^n$.

  (There is much more than this to say about manifolds, but it must be dealt with in another course.)

- The *tangent space* TM contains velocity $v_q = \dot{q}(t) \in T_qM$, tangent to curve $q(t) \in M$ at point q. The coordinates are $(q, v_q) \in TM$. Note, $\dim TM = 2n$ and subscript q reminds us that v_q is an element of the tangent space at the point q and that on manifolds we must keep track of base points.

  The union of tangent spaces $TM := \cup_{q \in M} T_qM$ is also called the *tangent bundle* of the manifold M.

  The curve $\dot{q}(t) \in TM$ is called the *tangent lift* of the curve $q(t) \in M$.

- A *motion* is defined as a smooth curve $q(t) \in M$ parameterised by $t \in \mathbb{R}$ that solves the *motion equation*, which is a system of differential equations

$$\dot{q}(t) = \frac{dq}{dt} = f(q) \in TM\,, \tag{2.1}$$

  or in components

$$\dot{q}^i(t) = \frac{dq^i}{dt} = f^i(q) \quad i = 1, 2, \ldots, n\,, \tag{2.2}$$

- The map $f : q \in M \to f(q) \in T_qM$ is a *vector field*.

  According to standard theorems about differential equations that are not proven in this course, the solution, or integral curve, $q(t)$ exists, provided f is sufficiently smooth, which will always be assumed to hold.

  Vector fields can also be defined as *differential operators* that act on functions, as

$$\frac{d}{dt}G(q) = \dot{q}^i(t)\frac{\partial G}{\partial q^i} = f^i(q)\frac{\partial G}{\partial q^i} \quad i = 1, 2, \ldots, n, \quad \text{(sum on repeated indices)} \tag{2.3}$$

for any smooth function $G(q) : M \to \mathbb{R}$.

- To indicate the dependence of the solution of its initial condition $q(0) = q_0$, we write the motion as a smooth transformation

$$q(t) = \phi_t(q_0)\,.$$

Because the vector field f is independent of time t, for any fixed value of t we may regard ϕ_t as mapping from M into itself that satisfies the composition law

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

and

$$\phi_0 = \mathrm{Id}\,.$$

Setting $s = -t$ shows that ϕ_t has a smooth inverse. A smooth mapping that has a smooth inverse is called a *diffeomorphism*. Geometric mechanics deals with diffeomorphisms.

- The smooth mapping $\phi_t : \mathbb{R} \times M \to M$ that determines the solution $\phi_t \circ q_0 = q(t) \in M$ of the motion equation (2.1) with initial condition $q(0) = q_0$ is called the *flow* of the vector field Q.

A point $q^\star \in M$ at which $f(q^\star) = 0$ is called a *fixed point* of the flow ϕ_t, or an *equilibrium*.

Vice versa, the vector field f is called the *infinitesimal transformation* of the mapping ϕ_t, since

$$\frac{d}{dt}\bigg|_{t=0} (\phi_t \circ q_0) = f(q)\,.$$

That is, $f(q)$ is the *linearisation* of the flow map ϕ_t at the point $q \in M$.

More generally, the *directional derivative* of the function h along the vector field f is given by the action of a differential operator, as

$$\frac{d}{dt}\bigg|_{t=0} h \circ \phi_t = \left[\frac{\partial h}{\partial \phi_t}\frac{d}{dt}(\phi_t \circ q_0)\right]_{t=0} = \frac{\partial h}{\partial q^i}\dot q^i = \frac{\partial h}{\partial q^i} f^i(q) =: Qh\,.$$

- Under a smooth change of variables $q = c(r)$ the vector field Q in the expression Qh transforms as

$$Q = f^i(q)\frac{\partial}{\partial q^i} \quad \mapsto \quad R = g^j(r)\frac{\partial}{\partial r^j} \quad \text{with} \quad g^j(r)\frac{\partial c^i}{\partial r^j} = f^i(q(r)) \quad \text{or} \quad g = c_r^{-1} f \circ c\,, \tag{2.4}$$

where c_r is the *Jacobian matrix* of the transformation. That is,

$$(Qh) \circ c = R(h \circ c)\,.$$

We express the transformation between the vector fields as $R = c^*Q$ and write this relation as

$$(Qh) \circ c =: c^*Q(h \circ c)\,. \tag{2.5}$$

The expression c^*Q is called the *pull-back* of the vector field Q by the map c. Two vector fields are equivalent under a map c, if one is the pull-back of the other, and fixed points are mapped into fixed points.

The inverse of the pull-back is called the *push-forward*. It is the pull-back by the inverse map.

- The *commutator*

$$QR - RQ =: \big[Q,\, R\big]$$

of two vector fields Q and R defines another vector field. Indeed, if

$$Q = f^i(q)\frac{\partial}{\partial q^i} \quad \text{and} \quad R = g^j(q)\frac{\partial}{\partial q^j}$$

then

$$\big[Q,\, R\big] = \left(f^i(q)\frac{\partial g^j(q)}{\partial q^i} - g^i(q)\frac{\partial f^j(q)}{\partial q^i}\right)\frac{\partial}{\partial q^j}$$

because the second-order derivative terms cancel. By the pull-back relation (2.5) one finds

$$c^*\big[Q,\, R\big] = \big[c^*Q,\, c^*R\big]$$

under a change of variables defined by a smooth map, c. This means the definition of the vector field commutator is independent of the choice of coordinates.

Exercise. Show that the tangent to the relation $c_t^*\big[Q,\, R\big] = \big[c_t^*Q,\, c_t^*R\big]$ evaluated at $t = 0$ is the *Jacobi identity* for the vector fields to form an algebra. ★

- The *differential* of a smooth function $f : M \to M$ is defined as

$$df = \frac{\partial f}{\partial q^i}dq^i\,.$$

- Under a smooth change of variables $s = \phi \circ q = \phi(q)$ the differential of the composition of functions $d(f \circ \phi)$ transforms according to the chain rule as

$$df = \frac{\partial f}{\partial q^i} dq^i \,, \quad d(f \circ \phi) = \frac{\partial f}{\partial \phi^j(q)} \frac{\partial \phi^j}{\partial q^i} dq^i = \frac{\partial f}{\partial s^j} ds^j \quad \Longrightarrow \quad d(f \circ \phi) = (df) \circ \phi \tag{2.6}$$

  That is, the differential d commutes with the pull-back ϕ^* of a smooth transformation ϕ,

$$d(\phi^* f) = \phi^* df \,. \tag{2.7}$$

  In a moment, this pull-back formula will give us the rule for transforming differential forms of any order.

- Differential k-forms on an n-dimensional manifold are defined in terms of the differential d and the antisymmetric *wedge product* $(\wedge)$ satisfying

$$dq^i \wedge dq^j = -\, dq^j \wedge dq^i \,, \quad \text{for} \quad i, j = 1, 2, \ldots, n \tag{2.8}$$

  By using wedge product, any k-form $\alpha \in \Lambda^k$ on M may be written locally at a point $q \in M$ in the differential basis dq^j as

$$\alpha_m = \alpha_{i_1 \ldots i_k}(m) dq^{i_1} \wedge \cdots \wedge dq^{i_k} \in \Lambda^k \,, \quad i_1 < i_2 < \cdots < i_k \,, \tag{2.9}$$

  where the sum over repeated indices is ordered, so that it must be taken over all i_j satisfying $i_1 < i_2 < \cdots < i_k$. Roughly speaking differential forms Λ^k are objects that can be integrated. As we shall see, vector fields also act on differential forms in interesting ways.

- Pull-backs of other differential forms may be built up from their basis elements, the dq^{i_k}. By equation (2.7), we have

Theorem

2.1 (Pull-back of a wedge product)**.** *The pull-back of a wedge product of two differential forms is the wedge product of their pull-backs:*

$$\phi_t^*(\alpha \wedge \beta) = \phi_t^* \alpha \wedge \phi_t^* \beta \,. \tag{2.10}$$

Definition

2.2 (Lie derivative of a differential k-form)**.** *The* Lie derivative *of a differential k-form Λ^k by a vector field Q is defined by linearising its flow ϕ_t around the identity $t = 0$,*

$$\pounds_Q \Lambda^k = \frac{d}{dt}\bigg|_{t=0} \phi_t^* \Lambda^k \quad \textit{maps} \quad \pounds_Q \Lambda^k \mapsto \Lambda^k \,.$$

Hence, by equation (2.10), the Lie derivative satisfies the product rule for the wedge product.

Corollary

2.3 (Product rule for the Lie derivative of a wedge product)**.**

$$\pounds_Q(\alpha\wedge\beta) = \pounds_Q\alpha\wedge\beta + \alpha\wedge\pounds_Q\beta\,. \tag{2.11}$$

2.1 Variational principles

Vocabulary

kinetic energy	Hamilton's principle	momentum
Riemannian metric	variational derivative	fibre derivative
Lagrangian	Legendre transformation	pairing

- Define *kinetic energy*, $KE : TM \to \mathbb{R}$, via a *Riemannian metric* $g_q(\,\cdot\,,\,\cdot\,) : TM \times TM \to \mathbb{R}$.
- Choose *Lagrangian* $L : TM \to \mathbb{R}$. (For example, one could choose L to be KE.)
- *Hamilton's principle* is $\delta S = 0$ with $S = \int_a^b L(q, \dot q)dt$, where for a family of curves parameterised smoothly by (t, ϵ) the linearisation

$$\delta S = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_a^b L(q(t,\epsilon), \dot q(t,\epsilon))dt$$

defines the *variational derivative* δS of S near the identity $\epsilon = 0$. The variations in q are assumed to vanish at endpoints in time, so that $q(a, \epsilon) = q(a)$ and $q(b, \epsilon) = q(b)$.

Exercise. Show that Hamilton's principle $\delta S = 0$ with $S = \int_a^b L(q, \dot q)dt$ implies Euler-Lagrange (EL) equations:

$$\dot p(q, \dot q) = \frac{d}{dt}\frac{\partial L(q, \dot q)}{\partial \dot q} = \frac{\partial L(q, \dot q)}{\partial q}\,. \tag{2.12}$$

★

- *Legendre transformation* $LT : (q, \dot q) \in TM \to (q, p) \in T^*M$ defines *momentum* p as the *fibre derivative* of L, namely

$$p := \frac{\partial L(q, \dot q)}{\partial \dot q} \in T^*M\,.$$

The LT is invertible for $\dot q = f(q, p)$, provided *Hessian* $\partial^2 L(q, \dot q)/\partial \dot q \partial \dot q$ has nonzero determinant. Note, $\dim T^*M = 2n$.

In terms of LT, the *Hamiltonian* $H : T^*M \to \mathbb{R}$ is defined by

$$H(q, p) = \langle p, \dot q\rangle - L(q, \dot q) \tag{2.13}$$

in which the expression $\langle p, \dot{q}\rangle$ in this calculation identifies a *pairing* $\langle\,\cdot\,,\,\cdot\,\rangle : T^*M \times TM \to \mathbb{R}$.

Taking the differential of this definition yields

$$dH = \langle H_p, dp\rangle + \langle H_q, dq\rangle = \langle dp, \dot{q}\rangle + \langle p - L_{\dot{q}}, d\dot{q}\rangle - \langle L_q, dq\rangle \tag{2.14}$$

from which Hamilton's principle $\delta S = 0$ for $S = \int_{t_0}^{t_1} \langle p, \dot{q}\rangle - H(q,p)\, dt$ produces Hamilton's canonical equations,

$$H_p = \dot{q} \quad \text{and} \quad H_q = -\,L_q = -\,\dot{p}\,. \tag{2.15}$$

Exercise. When $L = KE = \frac{1}{2}g_q(\dot{q},\dot{q}) =: \frac{1}{2}\|\dot{q}\|^2$ for the metric g_q, the solution $q(t)$ of the EL equations that passes from point $q(a)$ to $q(b)$ is a *geodesic* with respect to that metric. Compute the corresponding Christoffel symbols. ★

2.2 Hamilton's canonical equations

Theorem

2.4 (Hamiltonian equations)**.** *When the Lagrangian is* ***non-degenerate*** *(hyperregular), the Euler–Lagrange equations*

$$[\,L\,]_q := \frac{d}{dt}\frac{\partial L(q,\dot{q})}{\partial \dot{q}} - \frac{\partial L(q,\dot{q})}{\partial q} = 0\,,$$

in (2.12) are equivalent to ***Hamilton's canonical equations***

$$\dot{q} = \frac{\partial H}{\partial p}\,, \qquad \dot{p} = -\,\frac{\partial H}{\partial q}\,, \tag{2.16}$$

where $\partial H/\partial q$ and $\partial H/\partial p$ are the gradients of $H(p,q) = \langle\, p\,,\, \dot{q}\,\rangle - L(q,\dot{q})$ with respect to q and p, respectively.

Proof. The derivatives of the Hamiltonian follow from the differential of its defining Equation (2.13) as

$$\begin{aligned} dH &= \left\langle \frac{\partial H}{\partial p}, dp \right\rangle + \left\langle \frac{\partial H}{\partial q}, dq \right\rangle \\ &= \langle\, \dot{q}\,,\, dp\,\rangle - \left\langle \frac{\partial L}{\partial q}, dq \right\rangle + \left\langle p - \frac{\partial L}{\partial \dot{q}}, d\dot{q} \right\rangle. \end{aligned}$$

Consequently,

$$\frac{\partial H}{\partial p} = \dot{q} = \frac{dq}{dt}\,, \quad \frac{\partial H}{\partial q} = -\,\frac{\partial L}{\partial q} \quad \text{and} \quad \frac{\partial H}{\partial \dot{q}} = p - \frac{\partial L}{\partial \dot{q}} = 0\,.$$

The Euler–Lagrange equations $[\,L\,]_{q^a} = 0$ then imply

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

This proves the equivalence of the Euler–Lagrange equations and Hamilton's canonical equations for nondegenerate or hyperregular Lagrangians. □

Remark

2.5. *The Euler–Lagrange equations are second-order and they determine curves in configuration space $q \in M$. In contrast, Hamilton's equations are first-order and they determine curves in phase space $(q, p) \in T^*M$, a space whose dimension is twice the dimension of the configuration space M.*

Definition

2.6 (Number of degrees of freedom)**.** *The dimension of the configuration space is called the **number of degrees of freedom**.*

Remark

2.7. *Each degree of freedom has its own coordinate and momentum in phase space.*

Remark

2.8 (Momentum vs position in phase space)**.** *The momenta $p = (p_1, \dots, p_n)$ are coordinates in the cotangent bundle at $q = (q^1, \dots, q^n)$ corresponding to the basis $dq^1, \dots, dq^n$ for T_q^*M. This basis for one-forms in T_q^*M is dual to the vector basis $\partial/\partial q^1, \dots, \partial/\partial q^n$ for the tangent bundle T_qM at $q = (q^1, \dots, q^n)$.*

2.3 Phase-space action principle

Hamilton's principle on the tangent space of a manifold M may be augmented by imposing the relation $v = \frac{dq}{dt}$ as an additional constraint in terms of generalised coordinates $(q, v) \in T_qM$. In this case, the **constrained action** is given by

$$0 = \delta S := \delta \int_a^b L(q(t), v) \, + \Big\langle \, p \, , \, \frac{dq}{dt} - v \, \Big\rangle \, dt \, . \tag{2.17}$$

where p is a **Lagrange multiplier** for the constraint. The variations of this action result in

$$\delta S = \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q} - \frac{dp}{dt} \right) \delta q + \left(\frac{\partial L}{\partial v} - p \right) \delta v + \Big(\frac{dq}{dt} - v \Big) \delta p \, dt + \Big[p \, \delta q \Big]_{t_a}^{t_b} . \tag{2.18}$$

The contributions at the endpoints t_a and t_b in time vanish, because the variations δq are assumed to vanish then. Thus, stationarity of this action under these variations imposes the relations

$$\begin{aligned} \delta q : \quad & \frac{\partial L}{\partial q} = \frac{dp}{dt} \, , \\ \delta v : \quad & \frac{\partial L}{\partial v} = p \, , \\ \delta p : \quad & v = \frac{dq}{dt} \, . \end{aligned}$$

- Combining the first and second of these relations recovers the Euler–Lagrange equations (2.12) in the form

$$\frac{d}{dt} \frac{\partial L(q, v)}{\partial v} \bigg|_{v = \frac{dq}{dt}} = \frac{\partial L(q, v)}{\partial q} \bigg|_{v = \frac{dq}{dt}} . \tag{2.19}$$

  This calculation shows how it is that one may vary q and $\frac{dq}{dt}$ independently.

- The third relation constrains the variable v to be the time derivative of the trajectory $q(t)$ at any time t. The relation $v = \frac{dq}{dt}$ is called the **tangent lift** of the curve $q(t)$.

Substituting the Legendre-transform relation (2.13) into the constrained action (2.17) yields the **phase-space action**

$$S = \int_{t_a}^{t_b} \Big(p \frac{dq}{dt} - H(q, p) \Big) \, dt \, . \tag{2.20}$$

Upon varying the phase-space action in (2.20), Hamilton's principle yields

$$0 = \delta S = \int_{t_a}^{t_b} \left(\frac{dq}{dt} - \frac{\partial H}{\partial p} \right) \delta p - \left(\frac{dp}{dt} + \frac{\partial H}{\partial q} \right) \delta q \, dt + \Big[p \, \delta q \Big]_{t_a}^{t_b} . \tag{2.21}$$

Because the variations δq vanish at the endpoints t_a and t_b in time, the last term vanishes. Thus, stationary variations of the phase-space action in (2.20) recover Hamilton's canonical equations (2.16).

Remark

2.9 (Poisson bracket and Hamiltonian vector field).

Hamiltonian evolution along a curve $(q(t), p(t)) \in T^*M$ satisfying Equations (2.16) induces the evolution of a given function $F(q,p) : T^*M \to \mathbb{R}$ on the phase space T^*M of a manifold M, as

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial q}\frac{dq}{dt} + \frac{\partial H}{\partial q}\frac{dp}{dt} \\ &= \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} =: \{F \,,\, H\} \qquad (2.22) \\ &= \left(\frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p} \right) F =: X_H F \,. \qquad (2.23) \end{aligned}$$

*The second and third lines of this calculation introduce notation for two natural operations that will be investigated further in the next few sections. These are the **Poisson bracket** $\{\,\cdot\,,\,\cdot\,\}$ and the **Hamiltonian vector field** $X_H = \{\,\cdot\,,\,H\,\}$.*

2.4 Poisson brackets

Definition

2.10 (Canonical Poisson bracket). *Hamilton's canonical equations are associated with the **canonical Poisson bracket** for functions on phase space, defined by*

$$\frac{dp}{dt} = \{p \,,\, H\} \,, \quad \frac{dq}{dt} = \{q \,,\, H\} \,. \tag{2.24}$$

Hence, the evolution of a smooth function on phase space is expressed as

$$\frac{dF(q,p)}{dt} = \{F \,,\, H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} \,. \tag{2.25}$$

This expression defines the canonical Poisson bracket as a map $\{F,H\} : C^\infty \times C^\infty \to C^\infty$ *for smooth, real-valued functions* F, G *on phase space.*

Remark

2.11. *For one degree of freedom, the canonical Poisson bracket is the same as the determinant for a change of variables*

$$(q,p) \to (F(q,p), H(q,p)),$$

namely,

$$dF \wedge dH = \det \frac{\partial(F,H)}{\partial(q,p)} dq \wedge dp = \{F,H\} dq \wedge dp \,. \tag{2.26}$$

Here the wedge product $\wedge$ *denotes the antisymmetry of the determinant of the Jacobian matrix under exchange of rows or columns, so that*

$$dF \wedge dH = -\, dH \wedge dF.$$

Proposition

2.12 (Canonical Poisson bracket)**.** *The definition of the canonical Poisson bracket in (2.25) implies the following properties. By direct computation, the bracket*

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

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

  for the product of any two phase-space functions F *and* G*, and*
- *satisfies the Jacobi identity*

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

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

2.5 Canonical transformations

Definition

2.13 (Transformation)**.** *A **transformation** is a one-to-one mapping of a set onto itself.*

Example

2.14. *For example, under a change of variables*

$$(q,p) \to (Q(q,p), P(q,p))$$

in phase space T^*M, the Poisson bracket in (2.26) transforms via the Jacobian determinant, as

$$\begin{aligned} dF \wedge dH &= \{F, H\} dq \wedge dp \\ &= \{F, H\} \det \frac{\partial(q,p)}{\partial(Q,P)} dQ \wedge dP \,. \end{aligned} \tag{2.27}$$

Definition

2.15 (Canonical transformations)**.** *When the Jacobian determinant is equal to unity, that is, when*

$$\det \frac{\partial(q,p)}{\partial(Q,P)} = 1 \,, \quad \textit{so that} \quad dq \wedge dp = dQ \wedge dP \,, \tag{2.28}$$

*then the Poisson brackets $\{F, H\}$ have the same values in either set of phase-space coordinates. Such transformations of phase space T^*M are said to be **canonical transformations**, since in that case Hamilton's canonical equations keep their forms, as*

$$\frac{dP}{dt} = \{P \,,\, H\} \,, \quad \frac{dQ}{dt} = \{Q \,,\, H\} \,. \tag{2.29}$$

Remark

2.16. *If the Jacobian determinant above were equal to any nonzero constant, then Hamilton's canonical equations would still keep their forms, after absorbing that constant into the units of time. Hence, transformations for which*

$$\det \frac{\partial(q,p)}{\partial(Q,P)} = \textit{constant} \tag{2.30}$$

may still be said to be canonical.

Definition

2.17 (Lie transformation groups).

- *A set of transformations is called a* ***group****, provided:*
  - *it includes the identity transformation and the inverse of each transformation;*
  - *it contains the result of the consecutive application of any two transformations; and*
  - *composition of that result with a third transformation is associative.*
- *A group is a* ***Lie group****, provided its transformations depend smoothly on a parameter.*

Proposition

2.18. *The canonical transformations form a group.*

Proof. Composition of change of variables $(q, p) \to (Q(q, p), P(q, p))$ in phase space T^*M with constant Jacobian determinant satisfies the defining properties of a group. □

Remark

2.19. *The smooth parameter dependence needed to show that the canonical transformations actually form a Lie group will arise from their definition in terms of the Poisson bracket.*

2.6 Flows of Hamiltonian vector fields

The Leibniz property (product rule) in Proposition 2.12 suggests the canonical Poisson bracket is a type of derivative. This derivation property of the Poisson bracket allows its use in the definition of a Hamiltonian vector field.

Definition

2.20 (Hamiltonian vector field). *The Poisson bracket expression*

$$X_H = \{\cdot\,,\,H\} = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p} \tag{2.31}$$

defines a ***Hamiltonian vector field*** X_H*, for any smooth phase-space function* $H : T^*M \to \mathbb{R}$*.*

Proposition

2.21. *Solutions of Hamilton's canonical equations $q(t)$ and $p(t)$ are the* ***characteristic paths*** *of the first-order linear partial differential operator X_H. That is, X_H corresponds to the time derivative along these characteristic paths.*

Proof. Verify directly by applying the product rule for vector fields and Hamilton's equations in the form $\dot{p} = X_H p$ and $\dot{q} = X_H q$. □

Definition

2.22 (Hamiltonian flow)**.** *The union of the characteristic paths of the Hamiltonian vector field X_H in phase space T^*M is called the* ***flow*** *of the Hamiltonian vector field X_H. That is, the flow of X_H is the collection of maps $\phi_t : T^*M \to T^*M$ satisfying*

$$\frac{d\phi_t}{dt} = X_H\big(\phi_t(q,p)\big) = \{\phi_t,\, H\}\,, \tag{2.32}$$

for each $(q,p) \in T^*M$ for real t and $\phi_0(q,p) = (q,p)$.

Theorem

2.23. *Canonical transformations result from the smooth flows of Hamiltonian vector fields. That is, Poisson brackets generate canonical transformations.*

Proof. According to Definition 2.15, a transformation

$$(q(0), p(0)) \to (q(\epsilon), p(\epsilon))$$

which depends smoothly on a parameter ϵ is canonical, provided it preserves area in phase space (up to a constant factor that defines the units of area). That is, provided it satisfies the condition in Equation (2.28), discussed further in Chapter **??**, namely,

$$dq(\epsilon) \wedge dp(\epsilon) = dq(0) \wedge dp(0)\,. \tag{2.33}$$

Let this transformation be the flow of a Hamiltonian vector field X_F. That is, let it result from integrating the characteristic equations of

$$\frac{d}{d\epsilon} = X_F = \{\cdot\,,\, F\} = \frac{\partial F}{\partial p}\frac{\partial}{\partial q} - \frac{\partial F}{\partial q}\frac{\partial}{\partial p} =: F_{,p}\partial_q - F_{,q}\partial_p\,,$$

for a smooth function F on phase space. Then applying the Hamiltonian vector field to the area in phase space and exchanging the differential and derivative with respect to ϵ yields

$$\begin{aligned}\frac{d}{d\epsilon}\Big(dq(\epsilon)\wedge dp(\epsilon)\Big) &= d(X_F q)\wedge dp + dq\wedge d(X_F p)\\ &= d(F_{,p})\wedge dp + dq\wedge d(F_{,q})\\ &= (F_{,pq}dq + F_{,pp}dp)\wedge dp + dq\wedge(F_{,qq}dq + F_{,qp}dp)\\ &= (F_{,pq} - F_{,qp})dq\wedge dp\\ &= 0\,,\end{aligned}\tag{2.34}$$

by equality of the cross derivatives of F and asymmetry of the wedge product. Therefore, condition (2.33) holds and the transformation is canonical. □

Corollary

2.24. *The canonical transformations of phase space form a Lie group.*

Proof. The flows of the Hamiltonian vector fields are canonical transformations that depend smoothly on their flow parameters. □

Exercise.

Bohlin [Bo1911] discovered the following invertible conformal transformation between the Kepler problem with canonically conjugate pairs $(x, p_x)\,, (y, p_y)$ and the planar isotropic simple harmonic oscillator $(q_1, p_1)\,, (q_2, p_2)$:

$$p_1 + ip_2 = (q_1 - iq_2)(p_x + ip_y)\,, \qquad x + iy = (q_1 + iq_2)^2\,.$$

(a) Show that under this transformation $H_{osc}(q_j, p_j) \Longleftrightarrow H_{Kepler}(x, p_x; y, p_y)$ with real positive constants ω^2 and μ, as

$$H_{osc}(q_1, p_1; q_2, p_2) = \frac{1}{2}\left(p_1^2 + \omega^2 q_1^2\right) + \frac{1}{2}\left(p_2^2 + \omega^2 q_2^2\right) - 2\mu \quad\Longleftrightarrow\quad p_x^2 + p_y^2 - \frac{\mu}{\sqrt{x^2+y^2}} + \omega^2 = H_{Kepler}(x, p_x; y, p_y)\,.$$

(b) Show that the Bohlin [Bo1911] transformation is canonical; since $dx\wedge dp_x + dy\wedge dp_y = 2(dq_1\wedge dp_1 + dq_2\wedge dp_2)$.

This exercise is relevant to the integrable systems part of our course, because it shows that the two most well-known super-integrable systems in classical mechanics, with their closed orbits, extra conservation laws and various other bonus properties, such as scaling laws, are actually equivalent to each other. That is, Hooke's Law for linear restoring force is canonically equivalent to Newton's Law of Gravitation, and vice versa! Perhaps, if the two science giants knew that their two separate discoveries were one, that knowledge would set them both spinning in their graves! ★

Canonical transformation to oscillator variables in 2D

In oscillator variables with two complex components defined by,

$$\mathbf{q} + i\mathbf{p} = \mathbf{a} \in \mathbb{C}^2\,,$$

the Hamiltonian H_{osc} for two uncoupled linear oscillators transforms into the Hamiltonian $H_{1-1\,res}$ for the $1-1$ resonance

$$H_{osc} = \frac{1}{2}\Big(|\mathbf{p}|^2 + |\mathbf{q}|^2\Big) = \frac{1}{2}|\mathbf{a}|^2 = \frac{1}{2}\mathbf{a}\cdot\mathbf{a}^* = H_{1-1\,res}\,.$$

The linear transformation to oscillator variables is canonical: its symplectic two-form is

$$dq\wedge dp = \frac{1}{(-2i)}(dq + idp)\wedge(dq - idp) = \frac{1}{(-2i)}da\wedge da^*$$

where we ignore subscripts for brevity.

Likewise, the Poisson bracket transforms by the chain rule as

$$\{a, a^*\} = \{q + ip\,,\, q - ip\} = -2i\,\{q\,,\,p\} = -2i\,\mathrm{Id}$$

Thus, in oscillator variables Hamilton's canonical equations become

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

The corresponding Hamiltonian vector field is

$$X_H = \{\,\cdot\,, H\} = -2i\,\frac{\partial H}{\partial a^*}\frac{\partial}{\partial a} + 2i\,\frac{\partial H}{\partial a}\frac{\partial}{\partial a^*}$$

satisfying

$$X_H \,\lrcorner\, \frac{1}{(-2i)}da\wedge da^* = dH$$

where $\lrcorner$ (contract) is one of the standard symbols for the substitution of a vector field into a differential form.

Oscillator variables for two degrees of freedom As we have seen, the Hamiltonian for the 2D isotropic harmonic oscillator may be canonically transformed into

$$H = \frac{1}{2}(\,|\mathbf{p}|^2 + |\mathbf{q}|^2) = \frac{1}{2}(\,|a_1|^2 + |a_2|^2) = \frac{1}{2}|\mathbf{a}|^2\,,$$

with complex two-component vector $\mathbf{a} = (a_1\,,\,a_2) \in \mathbb{C}^2$. In terms of this complex vector, the corresponding canonical equations are

$$\dot{\mathbf{a}} = \{\mathbf{a}, H\} = -2i\,\mathbf{a} \quad \text{and} \quad \dot{\mathbf{a}}^* = \{\mathbf{a}^*, H\} = 2i\,\mathbf{a}^*$$

whose solutions are immediately found as simple S^1 phase shifts, linear in time:

$$\mathbf{a}(t) = e^{-i\,t}\mathbf{a}(0) \quad \text{and} \quad \mathbf{a}^*(t) = e^{i\,t}\mathbf{a}^*(0)$$

Being invariant under an arbitrary S^1 phase shift $\mathbf{a} \to e^{i\phi}\mathbf{a}$, the three quadratic quantities,

$$Y_1 + iY_2 = 2a_1 a_2^* \quad \text{and} \quad Y_3 = |a_1|^2 - |a_2|^2\,,$$

are all conserved by the simple S^1-shift motion of the 2D isotropic harmonic oscillator.

The three quadratic S^1-invariants are not independent. They form a vector $\mathbf{Y}$ with components $(Y_1, Y_2, Y_3) \in \mathbb{R}^3$ whose magnitude is related to the Hamiltonian that leaves them invariant,

$$|\mathbf{Y}|^2 := Y_1^2 + Y_2^2 + Y_3^2 = (|a_1|^2 + |a_2|^2)^2 = 4H^2\,.$$

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

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

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

$$\{F\,,\,G\}(\mathbf{Y}) = -\,4\mathbf{Y}\cdot\frac{\partial F}{\partial \mathbf{Y}} \times \frac{\partial G}{\partial \mathbf{Y}}\,.$$

The Hamiltonian for the 2D isotropic harmonic oscillator may be expressed in these variables as $H = |\mathbf{Y}|^2/2$. This Hamiltonian has derivative $\partial H/\partial \mathbf{Y} = \mathbf{Y}$; so it is a **Casimir** for this Poisson bracket. That is, $H = |\mathbf{Y}|^2/2$ Poisson-commutes (that is, H satisfies $\{F\,,\,H\} = 0$) with any smooth function $F(\mathbf{Y})$. In particular, it Poisson-commutes with each of the components (Y_1, Y_2, Y_3). Hence, as expected, each component of $\mathbf{Y} \in \mathbb{R}^3$ is conserved under the dynamics generated by this Hamiltonian, and of course this dynamics takes place on one of the level sets of H, which are spheres in $\mathbb{R}^3$.

Remark The Poisson brackets for the S^1 invariants $4\mathbf{Y} \in \mathbb{R}^3$ in the $1-1$ resonance problem coincide with the Poisson brackets for the three components of the body angular momentum $\mathbf{\Pi} \in \mathbb{R}^3$ for a freely rotating rigid body. This coincidence arises because the dual Lie algebras $su(2)^*$ and $so(3)^*$ arising from reduction by symmetry are both isomorphic to $\mathbb{R}^3$. To understand this remark, notice that two different sets of reductions and dual pairs are involved, one to $su(2)^*$ for the $1-1$ resonance and the other to $so(3)^*$ for the rigid body rotating in $\mathbb{R}^3$. The $1-1$ resonance dual pair of momentum maps $H \in \mathbb{R} \longleftarrow (\mathbb{C}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2, \omega) \longrightarrow \mathbf{Y} \in su(2)^* \simeq \mathbb{R}^3$ arises from the action of the Lie group $U(2) \simeq U(1) \times SU(2)$ on oscillator phase space $\mathbb{C}^2$ [HoVi2012]. In contrast, the dual pair for rigid body motion $\mathbf{\Pi} \in so(3)^* \simeq \mathbb{R}^3 \longleftarrow (T^*SO(3), \omega) \longrightarrow \boldsymbol{\pi} \in so(3)^* \simeq \mathbb{R}^3$ arises from the actions of the Lie group $SO(3)$ from the left and the right on the phase space $T^*SO(3)$ and the recognition that $T^*SO(3)/SO(3) = so(3)^*$. For more discussion of these two dual pairs, see [MaRa1999].

2.7 Noether's theorem on the Hamiltonian side

Exercise. (Prologue to Noether's theorem on the Hamiltonian side) Suppose the phase-space action (2.20) is invariant under the infinitesimal transformation $q \to q + \delta q$, with $\delta q = \xi_M(q) \in TM$ for $q \in M$ under the transformations of a Lie group G acting on a manifold M. That is, suppose S in (2.20) satisfies $\delta S = 0$ for $\delta q = \xi_M(q) \in TM$.

Does the invariance of the phase-space action principle in (2.20) also imply a conservation law, as in does in Hamilton's principle? ★

Answer. One would expect from Noether's theorem on the Lagrangian side that invariance of the phase-space action (2.20) under a canonical transformation would also imply conservation of the end-point quantity,

$$J^\xi(q,p) = < p,\, \xi_M(q) > \in \mathbb{R}\,, \tag{2.35}$$

arising from integration by parts evaluated at the endpoints. This notation introduces a pairing $< \cdot\,, \cdot >: T^*M \times TM \to \mathbb{R}$. The conservation of the phase-space function $J^\xi(q,p)$ is expressed in Hamiltonian form as

$$\frac{dJ^\xi}{dt} = \{J^\xi,\, H\} = 0\,. \tag{2.36}$$

That is, $X_H J^\xi = 0$, or, equivalently,

$$\begin{aligned} 0 = X_{J^\xi} H &= \frac{\partial J^\xi}{\partial p}\frac{\partial H}{\partial q} - \frac{\partial J^\xi}{\partial q}\frac{\partial H}{\partial p} \\ &= \xi_M(q)\partial_q H - p\xi_M'(q)\partial_p H \\ &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} H\big(q(\epsilon),\, p(\epsilon)\big) = \delta H\,. \end{aligned} \tag{2.37}$$

This means that H is invariant under $(\delta q,\,\delta p)=(\xi_M(q),\,-p\xi_M'(q))$. That is, H is invariant under the cotangent lift to T^*M of the infinitesimal **point transformation** $q\to q+\xi_M(q)$ of the Lie group G acting by canonical transformations on the manifold M.

Conversely, if the Hamiltonian $H(q,p)$ is invariant under the canonical transformation generated by X_{J^ξ}, then the Noether endpoint quantity J^ξ in (2.35) will be a constant of the canonical motion under H.

It remains to prove the claim that conservation of the Noether phase space quantity in (2.35) actually does follow from invariance of the phase-space action (2.20). ▲

Theorem

2.25 (Noether's theorem on the Hamiltonian side).

When Hamilton's canonical equations are satisfied, a sufficient condition for $J^\xi(p,q) :=< p,\xi_M(q) >$ to be conserved is that the corresponding Hamiltonian vector field X_{J^ξ} generates a canonical transformation which is a symmetry of the Hamiltonian H.

Proof. Define the Noether quantity $J^\xi(p,q) :=< p,\delta q >=< p,\xi_M(q) >$ for a function $\xi_M(q)$ which represents the infinitesimal action of a Hamiltonian vector field on coordinate $q\in M$. The corresponding Hamiltonian vector field is

$$X_{J^\xi}=\xi_M(q)\frac{\partial}{\partial q}-p\frac{\partial\xi_M(q)}{\partial q}\frac{\partial}{\partial p}=\Big\{\,\cdot\,,\,J^\xi(p,q)\Big\}\quad\text{so that}\quad\delta H=\Big\{\,H\,,\,J^\xi(p,q)\Big\}$$

The Hamiltonian vector field X_{J^ξ} generates a canonical transformation, so it preserves the symplectic form $\omega=dq\wedge dp$. In fact, X_{J^ξ} also preserves pdq, as may be seen from an elementary calculation using $\delta p=\{p,J^\xi\}=X_{J^\xi}p$ and $\delta q=\{q,J^\xi\}=X_{J^\xi}q$. Namely,

$$\delta(pdq)=(\delta p)dq+pd(\delta q)=-p\frac{\partial\xi_M(q)}{\partial q}dq+pd\big(\xi_M(q)\big)=0\,.$$

(More sophisticated demonstrations the $\delta(pdq)=0$ are also available.) Now, let's rewrite equation (2.21) for the variation of the phase-space action (2.20) in the following two ways

$$\begin{aligned}
0=\delta S&=\int_{t_a}^{t_b}\left(\left(\frac{dq}{dt}-\frac{\partial H}{\partial p}\right)\delta p-\left(\frac{dp}{dt}+\frac{\partial H}{\partial q}\right)\delta q\right)dt+\Big[p\,\xi_M(q)\Big]_{t_a}^{t_b}\\
&=\delta\int_{t_a}^{t_b}\Big(pdq-H(q,p)\,dt\Big)=\int_{t_a}^{t_b}\Big(\delta(pdq)-\delta H(q,p)\,dt\Big)\\
&=-\int_{t_a}^{t_b}\{H,J^\xi\}\,dt\,.
\end{aligned}\tag{2.38}$$

Comparing the first and third lines of equation (2.38) now proves the statement above of Noether's theorem on the Hamiltonian side. $\square$

Remark

2.26. *As expected, the theorem proves that the Noether quantity J^ξ associated with the symmetry of the phase-space action principle Poisson-commutes with the Hamiltonian, since $X_{J^\xi}H = \{H, J^\xi\} = 0$. Of course, the condition for the conservation of $J^\xi(p,q)$ under the Hamiltonian dynamics was clear from the start, because on the solutions of Hamilton's canonical equations we have $\frac{dJ^\xi}{dt} = \{J^\xi, H\}$, which vanishes, if $X_{J^\xi}H = \{H, J^\xi\} = 0$. However, even if this is clear by hindsight, one should not forget that the issue here is still Noether's result: Each symmetry of the phase-space action principle under a canonical transformation of the configuration manifold, when cotangent lifted to a Hamiltonian vector field on phase space is a symmetry of the Hamiltonian in that action principle and, thus, is associated with a conserved quantity under that Hamiltonian flow.*

Definition

2.27 (Cotangent lift momentum map)**.** *On introducing a pairing $\langle\,\cdot\,,\,\cdot\,\rangle : \mathfrak{g}^*\times\mathfrak{g}\to\mathbb{R}$, one may define a map $J : T^*M \to \mathfrak{g}^*$ in terms of this pairing and the Noether endpoint quantity in (2.35) as*

$$J^\xi = \big\langle p,\, \xi_M(q)\big\rangle_{T^*M\times TM} =: \big\langle J(q,p),\, \xi\big\rangle_{\mathfrak{g}^*\times\mathfrak{g}}\,, \tag{2.39}$$

for any fixed element of the Lie algebra $\xi\in\mathfrak{g}$. The map $J(q,p)$ is called the ***cotangent lift momentum map*** *associated with the infinitesimal transformation $\delta q = \xi_M(q)\in TM$ and its cotangent lift $\delta p = -p\xi_M'(q)\in TM^*$.*

Exercise. (Cotangent lift momentum maps are Poisson) Show that cotangent lift momentum maps are Poisson. That is, show that, for smooth functions F and H,

$$\big\{F\circ J\,,\,H\circ J\big\} = \big\{F\,,\,H\big\}\circ J\,. \tag{2.40}$$

This relation defines a **Lie–Poisson bracket** on $\mathfrak{g}^*$ that inherits the properties in Proposition 2.12 of the canonical Poisson bracket. ★

2.8 Properties of Hamiltonian vector fields

By associating Poisson brackets with Hamiltonian vector fields on phase space, one may quickly determine their shared properties.

Definition

2.28 (Hamiltonian vector field commutator). *The **commutator** of the Hamiltonian vector fields X_F and X_H is defined as*

$$[X_F\,,\,X_H] = X_F X_H - X_H X_F\,, \tag{2.41}$$

which is again a Hamiltonian vector field.

Exercise. Verify directly that the commutator of two Hamiltonian vector fields yields yet another one. ★

Lemma

2.29. *Hamiltonian vector fields satisfy the Jacobi identity,*

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

Proof. Write $[X_G\,,\,X_H] = G(H) - H(G)$ symbolically, so that

$$[X_F\,,\,[X_G\,,\,X_H]\,] = F(G(H)) - F(H(G)) - G(H(F)) + H(G(F))\,.$$

Summation over cyclic permutations then yields the result. □

Lemma

2.30. *The Jacobi identity holds for the canonical Poisson bracket $\{\cdot\,,\,\cdot\}$,*

$$\{F,\,\{G,\,H\}\} + \{G,\,\{H,\,F\}\} + \{H,\,\{F,\,G\}\} = 0\,. \tag{2.42}$$

Proof. Formula (2.42) may be proved by direct computation, as in Proposition 2.12. This identity may also be verified formally by the same calculation as in the proof of the previous lemma, by writing $\{G,\,H\} = G(H) - H(G)$ symbolically.

□

Remark

2.31 (Lie algebra of Hamiltonian vector fields)**.** *The Jacobi identity defines the Lie algebra property of Hamiltonian vector fields, which form a Lie subalgebra of all vector fields on phase space.*

Theorem

2.32 (Poisson bracket and commutator)**.** *The canonical Poisson bracket $\{F, H\}$ is put into one-to-one correspondence with the commutator of the corresponding Hamiltonian vector fields X_F and X_H by the equality (Lie algebra anti-homomorphism)*

$$X_{\{F\,,\,H\}} = -\,[X_F\,,\,X_H]\,. \tag{2.43}$$

Proof. One recalls that

$$\begin{aligned}
[X_G, X_H] &= X_G X_H - X_H X_G \\
\text{For Hamiltonian vector fields} &:= \{G,\,\cdot\,\}\{H,\,\cdot\,\} - \{H,\,\cdot\,\}\{G,\,\cdot\,\} \\
\text{Upon rearranging} &= \{G,\,\{H,\,\cdot\,\}\} - \{H,\,\{G,\,\cdot\,\}\} \\
\text{By Jacobi} &= \{\{G,\,H\},\,\cdot\,\} = -\,X_{\{G,H\}}\,.
\end{aligned}$$

The first line is the definition of the commutator of vector fields. The second line is the definition of Hamiltonian vector fields in terms of the Poisson bracket. The third line is a substitution. The fourth line uses the Jacobi identity (2.42) and skew-symmetry. □

3 Geometric Mechanics stems from the work of H. Poincaré [Po1901]

3.1 Poincaré's work in 1901 was based on earlier work of S. Lie in 1870's

Vocabulary

group	conjugation map	structure constants
Lie group, G	Lie algebra bracket,	reduced Lagrangian
identity element, e	$[\,\cdot\,,\,\cdot\,]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$	dual Lie algebra, $\mathfrak{g}^*$
Lie algebra, $\mathfrak{g}$	Jacobi identity	dual basis, $e^k\in\mathfrak{g}^*$
tangent vectors	basis vectors, $e_k\in\mathfrak{g}$	pairing, $\mathfrak{g}^*\times\mathfrak{g}\to\mathbb{R}$

- A *group* G is a set of elements $g\in G$ with an associative binary product $G\times G\to G$ that has a unique inverse g^{-1} such that $gg^{-1}=e$ with identity element e. The identity element e satisfies $eg=g=ge$ for all $g\in G$, with group product gh denoted by concatenation.

- A *Lie group* G is a group that depends smoothly on a set of parameters in $\mathbb{R}^{\dim(G)}$.

  A Lie group is also a *manifold* (a locally $\mathbb{R}^n$ space where the rules of calculus apply), so it is an arena for geometric mechanics.

- One can choose the manifold M for mechanics to be either the Lie group G or another manifold Q on which G acts $G\times Q\to Q$.

- The *Lie algebra* $\mathfrak{g}$ of the Lie group G is defined as the space of *tangent vectors* $\mathfrak{g}\cong T_eG$ at the identity e of the group.

  The Lie algebra has a *bracket* operation $[\,\cdot\,,\,\cdot\,]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$, which it inherits from linearisation at the identity e of the *conjugation map* $h\cdot g=hgh^{-1}$ for $g,h\in G$. For this, one begins with the conjugation map $h(t)\cdot g(s)=h(t)g(s)h(t)^{-1}$ for curves $g(s),h(t)\in G$, with $g(0)=e=h(0)$. One linearises at the identity, first in s to get the operation $\mathrm{Ad}:G\times\mathfrak{g}\to\mathfrak{g}$ and then in t to get the operation $\mathrm{ad}:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$, which yields the Lie bracket. The bracket operation is antisymmetric $[a,b]=-[b,a]$ and satisfies the *Jacobi identity* for $a,b,c\in\mathfrak{g}$,

  $$\text{Jacobi identity:}\quad [a,[b,c]]+[b,[c,a]]+[c,[a,b]]=0\,.$$

  The bracket operation among the *basis vectors* $e_k\in\mathfrak{g}$ with $k=1,2,\dots,\dim(\mathfrak{g})$ with $\xi=\sum_{k=1}^{\dim(\mathfrak{g})}\xi^k e_k$ for $\xi\in\mathfrak{g}$ defines the Lie algebra by its *structure constants* ${c_{ij}}^k$ in (summing over repeated indices)

  $$[e_i\,,\,e_j]={c_{ij}}^k e_k\,.$$

  The requirements of skew-symmetry and the Jacobi condition put constraints on the structure constants. These constraints are

  $$\text{skew-symmetry}\quad c_{ji}^k=-\,c_{ij}^k\,,\qquad \text{Jacobi identity}\quad c_{ij}^k c_{lk}^m+c_{li}^k c_{jk}^m+c_{jl}^k c_{ik}^m=0\,. \tag{3.1}$$

  Conversely, any set of constants c_{ij}^k that satisfy relations (3.1) defines a Lie algebra $\mathfrak{g}$.

Exercise. Prove that the Jacobi identity requires the relations (3.1) .
Hint: the Jacobi identity involves summing three terms of the form

$$[\mathbf{e}_l\,,\,[\mathbf{e}_i\,,\,\mathbf{e}_j\,]\,] = c^k_{ij}[\,\mathbf{e}_l\,,\,\mathbf{e}_k] = c^k_{ij}c^m_{lk}\mathbf{e}_m\,.$$

★

Adjoint and coadjoint actions. Before introducing the concepts AD, Ad, ad, Ad^* and ad^* for the various actions of a Lie group on itself, on its Lie algebra (its tangent space at the identity), the action of the Lie algebra on itself, and their dual actions, we give a short "Executive summary" of Poincaré's famous paper [Po1901], which laid down the foundations of geometric mechanics. Not all of the notation will be defined in advance, but the fundamental spirit and importance of the work should be clear.

3.2 H. Poincaré (1901) Executive summary

Sur une forme nouvelle des équations de la méchanique. *C.R. Acad. Sci.* **132,** 369-371.

In 1901 Poincaré noticed that when a Lie group G acts transitively on a manifold Q, taken as the *configuration space* of a mechanical system, then an opportunity arises, "to cast the equations of mechanics into a new form which could be interesting to know" [Po1901]. The familiar Euler-Lagrange equations of mechanics are

$$\frac{d}{dt}\frac{\partial L}{\partial v}\bigg|_{v=\dot q} = \frac{\partial L}{\partial q}\,, \tag{3.2}$$

for a given Lagrangian $L(q,v)$, defined on the tangent bundle T_qQ of the manifold Q along a curve passing through position $q(t)$ at time t. That is, the Lagrangian is a map $L : T_qM \to \mathbb{R}$, of the *velocity tangent space* into the real numbers, in which the tangent vector, v, is constrained by the Lagrange multiplier p to be $v = \dot q$, the *tangent lift* of the solution curve $q(t)$.

The Euler-Lagrange equations arise from Hamilton's principle

$$0 = \delta S := \delta \int_a^b L(q(t),v) + <\ p\,,\,\dot q(t) - v\ >\ dt =: \delta \int_a^b \overline{L}(q(t),\dot q(t))\,dt\,. \tag{3.3}$$

The "new form" of the equations of mechanics emerges when the Euler-Lagrange equations are lifted to a set of dynamical equations for a curve $g(t) \in G$, parameterized by time, t, by setting $q(t) = g^{-1}(t)q_0$, with $g(0) = e$, the identity element of the group. The action integral S in Hamilton's principle transforms under G in [Po1901] to produced a *reduced Lagrangian*, $\ell(u, g^{-1}q_0)$,

$$\int_a^b \overline{L}(q,\dot q)dt =: \int_a^b \tilde L(g,\dot g; q_0)dt = \int_a^b \tilde L(e, g^{-1}\dot g, g^{-1}q_0)dt =: \int_a^b \ell(u, g^{-1}q_0) + \big\langle m\,,\, g^{-1}\dot g - u\big\rangle_{\mathfrak g}\, dt\,, \tag{3.4}$$

where $\langle\cdot,\cdot\rangle_{\mathfrak g} : \mathfrak g^* \times \mathfrak g \to \mathbb R$ denotes the nondegenerate pairing between the Lie algebra $\mathfrak g$ and its dual $\mathfrak g^*$. The Lagrange multiplier m introduced in (3.4) imposes the *reconstruction relation* $g^{-1}\dot g = u$ for the curve $g(t) \in G$ which, in turn, generates the motion $q(t) = g^{-1}(t)q_0$ along a solution curve in Q.

Upon taking variations in Hamilton's principle, Poincaré cast the Euler-Lagrange equations for vanishing endpoint conditions $\delta g(a) = 0 = \delta g(b)$ into his "new form", which yields, in modern notation [HoScSt2009, MaRa1999], upon taking variations in g,

$$\int_a^b \underbrace{\left\langle \frac{d}{dt}\frac{\partial L}{\partial \dot q} - \frac{\partial L}{\partial q}\,,\, \delta q \right\rangle_Q}_{\text{Euler-Lagrange eqns}} dt = 0 = \int_a^b \underbrace{\left\langle \frac{d}{dt}\frac{\partial \ell}{\partial u} - \mathrm{ad}^*_u \frac{\partial \ell}{\partial u} - \frac{\partial \ell}{\partial q}\diamond q\,,\, v \right\rangle_{\mathfrak g}}_{\text{Euler-Poincaré eqns}} dt\ , \tag{3.5}$$

where $<\ \cdot,\cdot\ >_Q$: $T^*Q \times TQ \to \mathbb R$ denotes the natural duality pairing, taken fiberwise at each q, between the tangent bundle TQ and its dual T^*Q, the *phase space* of the mechanical system; $\mathrm{ad}^*_{(\cdot)}(\cdot) : \mathfrak g^* \times \mathfrak g \ni (\mu, u) \mapsto \mathrm{ad}^*_u\,\mu \in \mathfrak g^*$ is the dual of the Lie algebra commutator $\mathrm{ad}_{(\cdot)}(\cdot) : \mathfrak g \times \mathfrak g \ni (u, v) \mapsto \mathrm{ad}_u\, v := [u, v] \in \mathfrak g$; and the diamond operation is a map $\diamond : Q^* \times Q \to \mathfrak g^*$.[1]

In Poincaré's illustrative example, $G = SO(3)$ was the group of rotations in three dimensions; the manifold $Q = \mathbb R^3$ was three dimensional Euclidean space; $\mathfrak g = \mathfrak{so}^*(3) \simeq \mathbb R^3$ and $\mathfrak g^* = \mathfrak{so}^*(3) \simeq \mathbb R^3$ were isomorphic to $\mathbb R^3$; the pairings $\langle\cdot,\cdot\rangle_Q$ and $\langle\cdot,\cdot\rangle_{\mathfrak g}$ were both the Euclidean scalar product; the operations ad, ad* and $\diamond$ were all (plus, or minus) the vector cross product in $\mathbb R^3$; and Poincaré's new form of the equations of mechanics in that case reduced to Euler's equations for a heavy top. Thus began geometric mechanics. This fundamental source of geometric mechanics was carefully reviewed recently from a modern perspective in [Marle2013]. For textbook discussions of geometric mechanics, see, e.g., [HoScSt2009, MaRa1999].

[1] In general, the configuration space Q can be a manifold, not just a vector space. In his calculations, though, Poincaré treated his "states" in Q as coordinate vectors. By simply following his calculation, we find the diamond operation as a map $\diamond : Q^* \times Q \to \mathfrak g^*$, in which one regards $q \in Q$ as a (state) vector. In this case, the diamond map is given by $\langle p \diamond q, u\rangle_{\mathfrak g} :=< p, -\pounds_u q >_Q$ where $\pounds_u$ denotes Lie derivative as action of the Lie algebra element $u \in \mathfrak g$ on the vector space at state $q \in Q$. We write $\mu \in \mathfrak g^*$ as $\mu = \sum_{j=1}^{\dim(\mathfrak g)} \mu_j e^j$, with pairing $\langle e^j\,,\, e_k\rangle_{\mathfrak g} = \delta^j_k$.

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

The notation AD, Ad, and ad follows the standard notation for the corresponding actions of a Lie group on itself, on its Lie algebra (its tangent space at the identity), the action of the Lie algebra on itself, and their dual actions.

3.3.1 ADjoint, Adjoint and adjoint operations for matrix Lie groups

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

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

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

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

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

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

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

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

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{g(t)} \, \eta &= \frac{d}{dt}\Big|_{t=0} \Big(g(t)\eta g^{-1}(t) \Big) \\ &= \dot{g}(0)\eta g^{-1}(0) - g(0)\eta g^{-1}(0)\dot{g}(0)g^{-1}(0) \\ &= \xi \, \eta - \eta \, \xi = [\xi, \eta] = \mathrm{ad}_\xi \, \eta , \end{aligned} \tag{3.6}$$

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

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

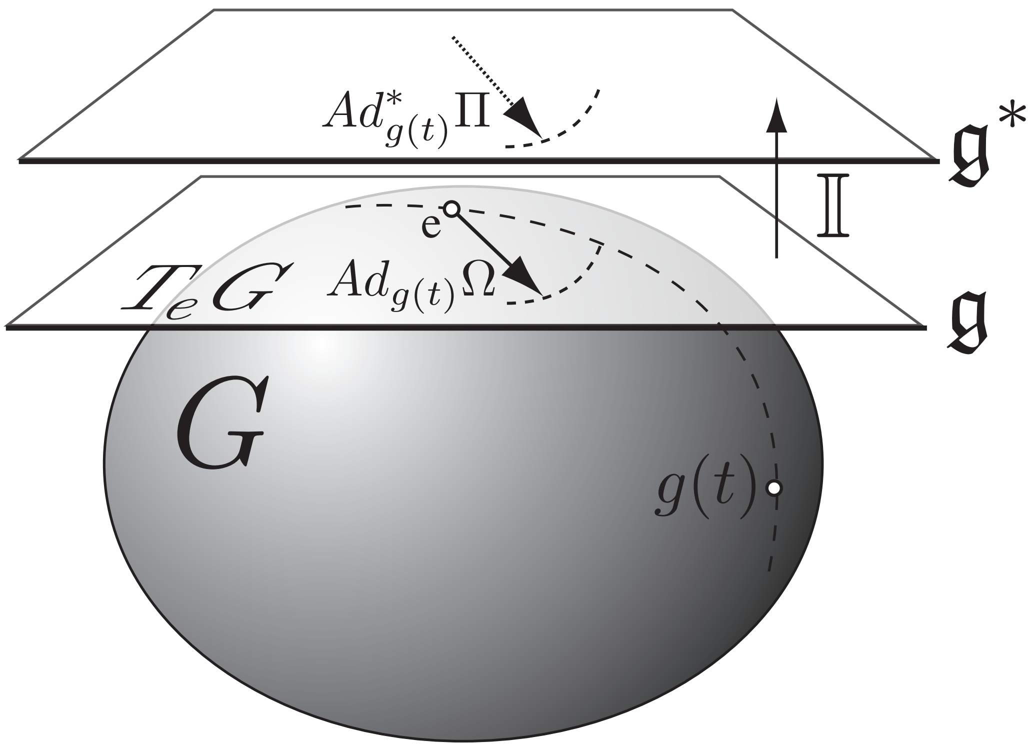
  is the matrix commutator for a matrix Lie algebra.

Remark

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

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

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



Exercise. Verify that (note the minus sign)

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

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

Proposition

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

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

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

Proof. By Equation (3.6), for a curve $\eta(t) \in \mathfrak{g}$,

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=t_0} \mathrm{Ad}_{g(t)}\,\eta(t) &= \frac{d}{dt}\Big|_{t=t_0}\Big(g(t)\eta(t)g^{-1}(t)\Big) \\
&= g(t_0)\Big(\dot{\eta}(t_0) + g^{-1}(t_0)\dot{g}(t_0)\eta(t_0) \\
&\quad - \eta(t_0)g^{-1}(t_0)\dot{g}(t_0)\Big)g^{-1}(t_0) \\
&= \left[\mathrm{Ad}_{g(t)}\left(\frac{d\eta}{dt} + \mathrm{ad}_{\xi}\eta\right)\right]_{t=t_0}.
\end{aligned} \tag{3.7}$$

□

Exercise. (Inverse Adjoint motion relation) Verify that

$$\frac{d}{dt}\mathrm{Ad}_{g(t)^{-1}}\eta = -\mathrm{ad}_{\xi}\mathrm{Ad}_{g(t)^{-1}}\eta\,, \tag{3.8}$$

for any fixed $\eta \in \mathfrak{g}$. Note the placement of $\mathrm{Ad}_{g(t)^{-1}}$ and compare with Exercise on page 38. ★

3.3.2 Compute the coAdjoint and coadjoint operations by taking duals

The pairing

$$\big\langle \,\cdot\,,\,\cdot\,\big\rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R} \tag{3.9}$$

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

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

$$\mathrm{Ad}^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* : \quad \langle\, \mathrm{Ad}^*_g\, \mu \,,\, \eta \,\rangle := \langle\, \mu \,,\, \mathrm{Ad}_g\, \eta \,\rangle \,, \tag{3.10}$$

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

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

$$\mathrm{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g}^* : \quad \langle\, \mathrm{ad}^*_\xi\, \mu \,,\, \eta \,\rangle := \langle\, \mu \,,\, \mathrm{ad}_\xi\, \eta \,\rangle \,, \tag{3.11}$$

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

Definition

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

$$\Phi^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* \quad \textit{given by} \quad (g, \mu) \mapsto \mathrm{Ad}^*_{g^{-1}}\mu \tag{3.12}$$

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

Remark

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

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

However, taking the inverse g^{-1} in Definition 3.3 of the coAdjoint action Φ^* restores the order and thereby allows it to represent the group composition law when acting on the dual Lie algebra, for then

$$\mathrm{Ad}^*_{g^{-1}} \mathrm{Ad}^*_{h^{-1}} = \mathrm{Ad}^*_{h^{-1}g^{-1}} = \mathrm{Ad}^*_{(gh)^{-1}} \,. \tag{3.13}$$

(See [MaRa1999] for further discussion of this point.)

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

Proposition

3.5 (Coadjoint motion relation). *Let $g(t)$ be a path in a matrix Lie group G and let $\mu(t)$ be a path in $\mathfrak{g}^*$, the dual (under the Frobenius pairing) of the matrix Lie algebra of G. The corresponding Ad^* operation satisfies*

$$\frac{d}{dt} Ad^*_{g(t)^{-1}} \mu(t) = Ad^*_{g(t)^{-1}} \left[\frac{d\mu}{dt} - ad^*_{\xi(t)} \mu(t) \right] , \tag{3.14}$$

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

Proof. The Exercise on page 39 introduces the inverse Adjoint motion relation (3.8) for any fixed $\eta \in \mathfrak{g}$, repeated as

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

Relation (3.8) may be proven by the following computation,

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

in which for the last step one recalls

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

Relation (3.8) plays a key role in demonstrating relation (3.14) in the theorem, as follows. Using the pairing $\langle \,\cdot\, , \,\cdot\, \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$

between the Lie algebra and its dual, one computes

$$
\begin{aligned}
\left\langle \frac{d}{dt}\mathrm{Ad}^*_{g(t)^{-1}}\mu(t),\eta\right\rangle &= \frac{d}{dt}\left\langle \mathrm{Ad}^*_{g(t)^{-1}}\mu(t),\eta\right\rangle \\
\text{by (3.10)} \quad &= \frac{d}{dt}\left\langle \mu(t),\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle \\
&= \left\langle \frac{d\mu}{dt},\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle + \left\langle \mu(t),\frac{d}{dt}\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle \\
\text{by (3.8)} \quad &= \left\langle \frac{d\mu}{dt},\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle + \left\langle \mu(t),-\mathrm{ad}_{\xi(t)}\left(\mathrm{Ad}_{g(t)^{-1}}\eta\right)\right\rangle \\
\text{by (3.11)} \quad &= \left\langle \frac{d\mu}{dt},\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle - \left\langle \mathrm{ad}^*_{\xi(t)}\mu(t),\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle \\
\text{by (3.10)} \quad &= \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\frac{d\mu}{dt},\eta\right\rangle - \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\mathrm{ad}^*_{\xi(t)}\mu(t),\eta\right\rangle \\
&= \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\left[\frac{d\mu}{dt}-\mathrm{ad}^*_{\xi(t)}\mu(t)\right],\eta\right\rangle .
\end{aligned}
$$

This concludes the proof.

□

Corollary

3.6. *The **coadjoint orbit relation***

$$\mu(t) = Ad^*_{g(t)}\mu(0) \tag{3.15}$$

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

$$\frac{d\mu}{dt} - ad^*_{\xi(t)}\mu(t) = 0\,, \quad \textit{with} \quad \xi(t) = g(t)^{-1}\dot{g}(t)\,. \tag{3.16}$$

Proof. Substituting Equation (3.16) into Equation (3.14) yields

$$\mathrm{Ad}^*_{g(t)^{-1}}\mu(t) = \mu(0)\,. \tag{3.17}$$

Operating on this equation with $\mathrm{Ad}^*_{g(t)}$ and recalling the composition rule for Ad^* from Remark 3.4 yields the result (3.15). □

Remark

3.7. *As it turns out, the equations in Poincaré (1901) for which we have been preparing describe coadjoint motion! Moreover, by equation (3.17) in the proof, coadjoint motion implies that $Ad^*_{g(t)^{-1}}\mu(t)$ is a conserved quantity.*

3.4 Preparation for understanding H. Poincaré's contribution [Po1901].

To understand [Po1901], let's introduce two more definitions.

1. Define a $\boxed{\textbf{\textit{reduced Lagrangian}}}$ $l:\mathfrak{g}\to\mathbb{R}$ and an associated variational principle $\delta S=0$ with $S=\int_a^b l(\xi)dt$ where $\xi=\xi^k e_k\in\mathfrak{g}$ has components ξ^k in the set of basis vectors e_k.

2. Define elements of the $\boxed{\textbf{\textit{dual Lie algebra}}}$ $\mathfrak{g}^*$ by using the fibre derivative of the Lagrangian $l:\mathfrak{g}\to\mathbb{R}$ to acquire a pairing as

$$\mu := \frac{\partial l(\xi)}{\partial \xi}\in\mathfrak{g}^*\,,\quad \text{written in components as}\quad \mu_i := \frac{\partial l(\xi)}{\partial \xi^i}\,,\quad \text{with a basis}\quad \mu=\mu_j e^j\,,\quad \text{and pairing}\quad \langle\, e^j,\, e_i\,\rangle=\delta^j_i\,.$$

In particular, the relation $dl=\langle\mu,d\xi\rangle$ defines a natural pairing $\langle\,\cdot\,,\,\cdot\,\rangle:\mathfrak{g}^*\times\mathfrak{g}\to\mathbb{R}$.

The natural $\boxed{\textbf{\textit{dual basis}}}$ for $\mathfrak{g}^*$ satisfies $\langle e^j,e_k\rangle=\delta^j_k$ in this pairing and an element $\mu\in\mathfrak{g}^*$ has components in this dual basis given by $\mu=\mu_k e^k$, again with with $k=1,2,\dots,\dim(\mathfrak{g})$.

- $\boxed{\textbf{\textit{Exercise:}}}$

(a) Show that Hamilton's principle $\delta S=0$ with $S=\int_a^b l(\xi)\,dt$ implies the Euler-Poincaré (EP) equations:

$$\frac{d}{dt}\mu_i = -\,c_{ij}{}^k\xi^j\mu_k\,,\quad \text{with}\quad \mu_k=\frac{\partial l(\xi)}{\partial \xi^k}\,,$$

for variations given by

$$\delta\xi=\dot\eta+[\xi,\eta]\quad\text{with}\quad \xi,\eta\in\mathfrak{g}\,. \tag{3.18}$$

For this, explain how this type of variations arises from variations of the group elements.

Note: $[e_j,\,e_k]=c_{jk}{}^ie_i$, so

$$[\xi,\eta]=[\xi^je_j,\,\eta^ke_k]=\xi^j[e_j,\,e_k]\eta^k=\xi^j\eta^kc_{jk}{}^ie_i=[\xi,\eta]^ie_i\,.$$

- $\boxed{\textbf{\textit{Answer:}}}$ Variations given by $\delta\xi = \dot\eta + [\xi,\eta]$ with $\xi,\eta\in\mathfrak{g}$ arise from variations of the group elements, as follows, by a direct computation,

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

On taking the difference, the terms with cross derivatives cancel and one finds the variational formula (3.18),

$$
\xi\,' - \dot\eta = [\,\xi\,,\,\eta\,] \quad\text{with}\quad [\,\xi\,,\,\eta\,] := \xi\,\eta - \eta\,\xi = \mathrm{ad}_\xi\,\eta\,. \tag{3.19}
$$

(See Remark 6.5 for more details.)

Upon using formula (6.4), the left-invariant variations in of the action in Hamilton's principle yield

$$
\begin{aligned}
\delta S = \delta\int_a^b l(\xi)dt &= \int_a^b \left\langle \frac{\partial l}{\partial\xi},\,\delta\xi\right\rangle dt = \int_a^b \left\langle \frac{\partial l}{\partial\xi},\,\dot\eta + [\xi,\eta]\right\rangle dt\\
&= \int_a^b \left\langle \frac{\partial l}{\partial\xi^n}e^n,\,\dot\eta^i e_i + \xi^j\eta^k c_{jk}{}^i e_i\right\rangle dt \quad\text{since}\quad \left\langle e^n,\,e_i\right\rangle = \delta^n_i\\
&= \int_a^b \underbrace{\left(-\frac{d}{dt}\frac{\partial l}{\partial\xi^i} + \frac{\partial l}{\partial\xi^k}\xi^j c_{ji}{}^k\right)}_{\text{Euler-Poincaré equation}} \eta^i\,dt + \left[\frac{\partial l}{\partial\xi^i}\,\eta^i\right]_a^b
\end{aligned}
$$

where, in the last step, we integrated by parts and relabelled indices. Hence, when η^i vanishes at the endpoints in time, but is otherwise arbitrary, we recover the EP equations as

$$
\frac{d}{dt}\frac{\partial l}{\partial\xi^i} + \frac{\partial l}{\partial\xi^k}\xi^j c_{ij}{}^k = 0\,,
$$

where we have used the antisymmetry of the structure constant $c_{ij}{}^k = -\,c_{ji}{}^k$.

These are the equations introduced by Poincaré in [Po1901], which we now write as $\frac{d}{dt}\frac{\partial l}{\partial\xi} - \mathrm{ad}^*_\xi\frac{\partial l}{\partial\xi} = 0$.

Here the notation ad^* is defined by $\left\langle -\,\mathrm{ad}^*_\xi\frac{\partial l}{\partial\xi},\,\eta\right\rangle := \frac{\partial l}{\partial\xi^k}\xi^j c_{ij}{}^k\eta^i = \frac{\partial l}{\partial\xi^k}[e_i\eta^i,\,e_j\xi^j]^k =: \left\langle\frac{\partial l}{\partial\xi},\,-\,\mathrm{ad}_\xi\eta\right\rangle$.

- $\boxed{\textbf{\textit{Exercise:}}}$ Write Noether's theorem for the Euler-Poincaré theory.
- $\boxed{\textbf{\textit{Answer:}}}$ To each continuous symmetry group G of the Lagrangian $l(\xi)$, the quantity $(\frac{\partial l}{\partial \xi^i}\,\eta^i)$ is conserved by the Euler-Poincaré motion equation, where $\eta^i e_i \in \mathfrak{g}$ is the infinitesimal transformation of the action of the group $G \times \mathfrak{g} \to \mathfrak{g}$.
  Proof: Look at the end point terms in the variation of the action, assuming $\delta S = 0$ because of a symmetry of the Lagrangian $l(\xi)$.
- $\boxed{\textbf{\textit{Exercise:}}}$ The Lie algebra $\mathfrak{so}(3)$ of the Lie group $SO(3)$ of rotations in three dimensions has structure constants ${c_{ij}}^k = {\epsilon_{ij}}^k$, where ${\epsilon_{ij}}^k$ with $i, j, k \in \{1, 2, 3\}$ is totally antisymmetric under pairwise permutations of its indices, with ${\epsilon_{12}}^3 = 1$, ${\epsilon_{21}}^3 = -1$, etc.

  Identify the Lie bracket $[a, b]$ of two elements $a = a^i e_i, b = b^j e_j \in \mathfrak{so}(3)$ with the cross product $\mathbf{a} \times \mathbf{b}$ of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ according to [2]

$$[a, b] = [a^i e_i, b^j e_j] = a^i b^j {\epsilon_{ij}}^k e_k = (\mathbf{a} \times \mathbf{b})^k e_k\,.$$

  (a) Show that in this case the EP equation

$$\dot{\mu}_i = -{\epsilon_{ij}}^k \xi^j \mu_k$$

  is equivalent to the vector equation for $\boldsymbol{\xi}, \boldsymbol{\mu} \in \mathbb{R}^3$

$$\dot{\boldsymbol{\mu}} = -\,\boldsymbol{\xi} \times \boldsymbol{\mu}\,.$$

  (b) Show that when the Lagrangian is given by the quadratic

$$l(\boldsymbol{\xi}) = \frac{1}{2}\|\boldsymbol{\xi}\|_K^2 = \frac{1}{2}\boldsymbol{\xi} \cdot K\boldsymbol{\xi} = \frac{1}{2}\xi^i K_{ij} \xi^j$$

  for a symmetric constant Riemannian metric $K^T = K$, then Euler's equations for a rotating rigid body are recovered.

> That is, Euler's equations for rigid body motion are contained in Poincaré's equations for motion on Lie groups!
>
> And Poincaré's equations generalise Euler's equations for rigid body motion from $\mathbb{R}^3$ to motion on Lie groups!

  (c) Identify the functional dependence of $\boldsymbol{\mu}$ on $\boldsymbol{\xi}$ and give the physical meanings of the symbols $\boldsymbol{\xi}, \boldsymbol{\mu}$ and K in Euler's rigid body equations.

---

[2] (a') Show that this formula implies the Jacobi identity for the cross product of vectors in $\mathbb{R}^3$. This is no surprise because, that familiar cross product relation for vectors may be proven directly by using the antisymmetric tensor ${\epsilon_{ij}}^k$.

3.5 Euler-Poincaré variational principle for the rigid body

The Euler rigid-body equations on $T\mathbb{R}^3$ are

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega}\,, \tag{3.20}$$

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

Definition

3.8 (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

3.9 (Euler's rigid-body equations)**.** *Euler's rigid-body equations* (3.20) *arise from Hamilton's principle,*

$$\delta S(\mathbf{\Omega}) = \delta \int_a^b l(\mathbf{\Omega})\, dt = 0, \tag{3.21}$$

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

$$l(\mathbf{\Omega}) = \frac{1}{2}\langle \mathbb{I}\mathbf{\Omega}, \mathbf{\Omega}\rangle = \frac{1}{2}\,\mathbb{I}\mathbf{\Omega}\cdot\mathbf{\Omega} = \frac{1}{2}\,(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)\,, \tag{3.22}$$

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

$$\delta\mathbf{\Omega} = \dot{\mathbf{\Xi}} + \mathbf{\Omega} \times \mathbf{\Xi}\,, \tag{3.23}$$

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

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

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

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

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

Since $\mathbf{\Xi}$ is otherwise arbitrary, (3.21) is equivalent to

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

which recovers Euler's equations (3.20) in vector form. □

Proposition

3.10 (Derivation of the restricted variation)**.** *The restricted variation in (3.23) arises via the following steps:*

(i) Vary the definition of body angular velocity, $\widehat{\Omega} = O^{-1}\dot{O}$.

(ii) Take the time derivative of the variation, $\widehat{\Xi} = O^{-1}O'$.

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

(iv) Apply the hat map.

Proof. One computes directly that

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

On taking the difference, the cross derivatives cancel and one finds a variational formula equivalent to (3.23),

$$\widehat{\Omega}\,' - \widehat{\Xi}^{\,\cdot} = \Big[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,\Big] \quad \text{with} \quad [\,\widehat{\Omega}\,,\,\widehat{\Xi}\,] := \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega}\,. \tag{3.24}$$

Under the bracket relation

$$[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,] = (\boldsymbol{\Omega}\times\boldsymbol{\Xi})\widehat{\ } \tag{3.25}$$

for the hat map, this equation recovers the vector relation (3.23) in the form

$$\boldsymbol{\Omega}\,' - \dot{\boldsymbol{\Xi}} = \boldsymbol{\Omega}\times\boldsymbol{\Xi}\,. \tag{3.26}$$

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

$$\mathbb{I}\dot{\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega}\,, \tag{3.27}$$

follow from the variational principle (3.21) with variations of the form (3.23) derived from the definition of body angular velocity $\widehat{\Omega}$. □

Exercise. What conservation law does Noether's theorem imply for the rigid-body equations (3.20). Hint, is the Lagrangian in (3.22) invariant under rotations around $\boldsymbol{\Xi}$? ★

3.6 Clebsch variational principle for the rigid body

Proposition

3.11 (Clebsch variational principle)**.**
The Euler rigid-body equations on $T\mathbb{R}^3$ *given in equation* (3.20) *as*

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

are equivalent to the ***constrained variational principle****,*

$$\delta S(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P}) = \delta \int_a^b l(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P}) \, dt = 0, \tag{3.28}$$

for a ***constrained action integral***

$$\begin{aligned} S(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}) &= \int_a^b l(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}) \, dt \\ &= \int_a^b \frac{1}{2}\mathbf{\Omega} \cdot \mathbb{I}\mathbf{\Omega} + \mathbf{P} \cdot \Big(\dot{\mathbf{Q}} + \mathbf{\Omega} \times \mathbf{Q} \Big) \, dt \,. \end{aligned} \tag{3.29}$$

Remark

3.12 (Reconstruction as constraint)**.**

- *The first term in the Lagrangian (3.29),*

$$l(\mathbf{\Omega}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2) = \frac{1}{2}\mathbf{\Omega}^T\mathbb{I}\mathbf{\Omega} \,, \tag{3.30}$$

  is the (rotational) kinetic energy of the rigid body.

- *The second term in the Lagrangian (3.29) introduces the Lagrange multiplier* $\mathbf{P}$ *which imposes the constraint*

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

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

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

which satisfies

$$\begin{aligned}\dot{\mathbf{Q}}(t) &= -(O^{-1}\dot{O})O^{-1}(t)\mathbf{Q}(0) \\ &= -\hat{\Omega}(t)\mathbf{Q}(t) = -\boldsymbol{\Omega}(t)\times\mathbf{Q}(t)\,. \end{aligned} \tag{3.31}$$

Proof. The variations of S are given by

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

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

$$\mathbb{I}\boldsymbol{\Omega} = \mathbf{P}\times\mathbf{Q}\,, \quad \dot{\mathbf{Q}} = -\,\boldsymbol{\Omega}\times\mathbf{Q}\,, \quad \dot{\mathbf{P}} = -\,\boldsymbol{\Omega}\times\mathbf{P}\,. \tag{3.32}$$

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

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

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

Remark

3.13. *The Clebsch approach is 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

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

show that the Poisson brackets for $\mathbf{\Pi}=\mathbf{P}\times\mathbf{Q}\in\mathbb{R}^3$ are

$$\{\Pi_a, \Pi_i\} = \{\epsilon_{abc}P_bQ_c\,,\, \epsilon_{ijk}P_jQ_k\} = -\,\epsilon_{ail}\Pi_l\,.$$

Derive the corresponding Lie–Poisson bracket $\{f, h\}(\mathbf{\Pi})$ for functions of the $\mathbf{\Pi}$'s. ★

Answer. The $\mathbb{R}^3$ components of the angular momentum $\mathbf{\Pi} = \mathbb{I}\mathbf{\Omega} = \mathbf{P}\times\mathbf{Q}$ in (3.32) are

$$\Pi_a = \epsilon_{abc}P_bQ_c\,,$$

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

$$\{\Pi_a, \Pi_i\} = \{\epsilon_{abc}P_bQ_c\,,\, \epsilon_{ijk}P_jQ_k\} = -\,\epsilon_{ail}\Pi_l\,.$$

Consequently, the derivative property of the canonical Poisson bracket yields

$$\{f, h\}(\mathbf{\Pi}) = \frac{\partial f}{\partial \Pi_a}\{\Pi_a, \Pi_i\}\frac{\partial h}{\partial \Pi_b} = -\,\epsilon_{abc}\Pi_c\frac{\partial f}{\partial \Pi_a}\frac{\partial h}{\partial \Pi_b} = -\mathbf{\Pi}\cdot\frac{\partial f}{\partial \mathbf{\Pi}}\times\frac{\partial h}{\partial \mathbf{\Pi}}\,, \tag{3.33}$$

which is the Lie–Poisson bracket on functions of the $\mathbf{\Pi}$'s. This Poisson bracket satisfies the Jacobi identity as a result of the Jacobi identity for the vector cross product on $\mathbb{R}^3$. ▲

Remark

3.14. *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 (3.32) 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****.*

Exercise.

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

$$\frac{dM}{dt} = [\, M \,,\, \Omega \,]\,, \tag{3.34}$$

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

$$S(\Omega, Q, P) = \int_a^b l(\Omega) + \mathrm{tr}\Big(P^T \left(\dot{Q} - Q\Omega\right)\Big)\, dt\,, \tag{3.35}$$

for which $M = \delta l/\delta\Omega = P^T Q - Q^T P$ and $Q, P \in SO(n)$ satisfy the following symmetric equations reminiscent of those in (3.32),

$$\dot{Q} = Q\Omega \quad \text{and} \quad \dot{P} = P\Omega\,, \tag{3.36}$$

as a result of the constraints.

(b) How does equation (3.34) for the $SO(n)$ rigid body dynamics change, if the Lagrangian $l(\Omega)$ in (3.35) is changed to accommodate dependence on Q, i.e., if we have $l(\Omega, Q)$?

(c) Derive the Lie-Poisson bracket for the Hamiltonian formulation of the N-dimensional heavy top. ★

Answer.

(a)

$$\begin{aligned}
0 = \delta S(\Omega, Q, P) &= \delta \int_a^b l(\Omega) + \Big\langle P \,,\, \dot{Q} - Q\Omega \Big\rangle \, dt \\
&= \int_a^b \Big\langle \frac{\partial l}{\partial \Omega} - Q^T P \,,\, \delta\Omega \Big\rangle + \Big\langle \delta P \,,\, \dot{Q} - Q\Omega \Big\rangle \\
&\qquad - \Big\langle \dot{P} - P\Omega \,,\, \delta Q \Big\rangle \, dt + \Big\langle P \,,\, \delta Q \Big\rangle \Big|_a^b \,.
\end{aligned}$$

Thus, we have the variational equations,

$$\begin{aligned} \delta\Omega: &\quad \frac{\partial l}{\partial \Omega} = Q^T P \\ \delta P: &\quad \dot{Q} = Q\Omega \\ \delta Q: &\quad \dot{P} = P\Omega \end{aligned}$$

To derive the Euler equation, we compute

$$(Q^T P)^{\cdot} = \dot{Q}^T P + Q^T \dot{P} = \Omega^T Q^T P + Q^T P \Omega = [Q^T P\,,\,\Omega]$$

since $\Omega^T = -\Omega$. Likewise, $(P^T Q)^{\cdot} = [P^T Q\,,\,\Omega]$.
Consequently, upon antisymmetrising because $\delta\Omega^T = -\delta\Omega$, we find that $M = \frac{1}{2}(\delta l/\delta\Omega - \delta l/\delta\Omega^T) = \frac{1}{2}(Q^T P - P^T Q)$ satisfies the Euler equation, $\dot{M} = [M, \Omega]$.

(b) By slightly modifying the previous calculation to include $\partial l/\partial Q$, we find

$$\frac{dM}{dt} = [\,M\,,\,\Omega\,] + \frac{1}{2}\left(Q^T \frac{\partial l}{\partial Q} - \frac{\partial l}{\partial Q}^T Q \right), \qquad \frac{dQ}{dt} = Q\Omega \tag{3.37}$$

where we have antisymmetrised the term $Q^T \frac{\partial l}{\partial Q}$ so the equation transforms properly under taking transpose.

(c) By Legendre transforming to the Hamiltonian

$$h(M,Q) = \langle M, \Omega \rangle - l(\Omega, Q)$$

and by taking the time derivative and rearranging using $M^T = -M$ and the Frobenius pairing $\langle A, B\rangle = \mathrm{tr}(A^T B)$ as

$$\begin{aligned} \frac{d}{dt} f(M,Q) &= \left\langle \frac{\partial f}{\partial M}\,,\,\frac{dM}{dt} \right\rangle + \left\langle \frac{\partial f}{\partial Q}\,,\,\frac{dQ}{dt} \right\rangle = \left\langle \frac{\partial f}{\partial M}\,,\,\left[M\,,\,\frac{\partial h}{\partial M} \right] - Q^T \frac{\partial h}{\partial Q} \right\rangle + \left\langle \frac{\partial f}{\partial Q}\,,\,Q \frac{\partial h}{\partial M} \right\rangle \\ &= -\left\langle M\,,\,\left[\frac{\partial f}{\partial M}\,,\,\frac{\partial h}{\partial M} \right] \right\rangle - \left\langle Q\,,\,\frac{\partial f}{\partial Q}\frac{\partial h}{\partial M} - \frac{\partial h}{\partial Q}\frac{\partial f}{\partial M} \right\rangle =: \Big\{ f\,,\,h \Big\} \end{aligned}$$

we have built the Lie-Poisson bracket for the Hamiltonian formulation! For $S)(3)$, this Lie-Poisson bracket becomes, via the hat map,

$$\Big\{ f\,,\,h \Big\} = -\boldsymbol{\Pi} \cdot \frac{\partial f}{\partial \boldsymbol{\Pi}} \times \frac{\partial h}{\partial \boldsymbol{\Pi}} - \boldsymbol{\Gamma} \cdot \left(\frac{\partial f}{\partial \boldsymbol{\Gamma}} \times \frac{\partial h}{\partial \boldsymbol{\Pi}} - \frac{\partial h}{\partial \boldsymbol{\Gamma}} \times \frac{\partial f}{\partial \boldsymbol{\Pi}} \right)$$

▲

4 Poincaré's formulation of mechanics on Lie groups

If time runs out for treating [Po1901] in detail, we'll go to Section 5

4.1 Reading assignment: H. Poincaré (1901) Paraphrased! cf. [Marle2013]

Sur une forme nouvelle des équations de la méchanique. *C.R. Acad. Sci.* 132, 369-371.

Poincaré began by explaining that the opportunity to work on the rotational motion of hollow solid bodies filled with liquid, had led him to cast the equations of mechanics into a new form which he said, "could be interesting to know".

Suppose a dynamical system has n degrees of freedom, and let $q = \{q^1, ..., q^n\}$ lying in a smooth manifold M be the variables describing the state of the system. Let T and V be the kinetic and potential energy of the system.

Consider a Lie group G, which acts transitively on the smooth manifold M. (That is, suppose the action of the Lie group G covers the entire manifold, M.) Let $X_\alpha(f)$ for a smooth function $f : M \to \mathbb{R}$ be an *infinitesimal transformation* of this group such that[3]

$$X_\alpha(f) = \sum_{i=1}^{n} X^i_\alpha(q) \frac{\partial f}{\partial q^i} = X^1_\alpha \frac{\partial f}{\partial q^1} + X^2_\alpha \frac{\partial f}{\partial q^2} + \cdots + X^n_\alpha \frac{\partial f}{\partial q^n} \,. \tag{4.1}$$

Exercise. Lie showed that the characteristic equations of Lie algebra vector fields determine the finite transformations of their Lie groups, For good discussions of this point, see Peter Olver's book on Group Theory and Differential Equations.

$$\frac{dq^i}{d\epsilon} = \sum_{\alpha=1}^{r} \eta^\alpha X_\alpha(q^i) = \sum_{\alpha=1}^{r} \eta^\alpha X^i_\alpha(q) \quad \Longrightarrow \quad d\epsilon = \frac{dq^i}{X^i_\alpha(q)} \quad \text{(for each } \alpha\text{, no sum on } i\text{)}$$

Compute the finite transformations and commutator table for $(n, r = 1, 3)$, $X_1 = \partial_q, X_2 = -q\partial_q, X_3 = -q^2\partial_q$. Find 2×2 matrix representations of the subalgebras of this 3-dimensional algebra. Find vector fields producing the classical matrix Lie groups: upper triangular, $SL(2, \mathbb{R})$, $SE(3)$, the Galilean group, and the group of real projective transformations. ★

Vector fields on the real line.

We integrate the characteristic equations of the following vector fields, as

[3] As a special case of the finite dimensional systems considered here, the set of infinitesimal transformations $\{X_\alpha\}$ may be restricted to be linear in q, *represented* by a set of r constant $n \times n$ matrices acting linearly on the set of n states $\{q\}$. Then, for example, $X^i_\alpha(q) = \sum_{j=1}^n [X_\alpha]^i_j q^j$.

1. $v_1 = X_1 \partial_q = \partial_q, \quad \frac{dq}{d\epsilon_1} = 1 \quad \Longrightarrow \quad q(\epsilon_1) = q(0) + \epsilon_1$
2. $v_2 = X_2 \partial_q = -q\partial_q \quad \frac{dq}{d\epsilon_2} = -q \quad \Longrightarrow \quad q(\epsilon_2) = e^{-\epsilon_2} q(0)$
3. $v_3 = X_3 \partial_q = -\, q^2 \partial_q, \quad \frac{dq}{d\epsilon_3} = -\, q^2 \quad \Longrightarrow \quad q(\epsilon_3) = \frac{q(0)}{1+\epsilon_3 q(0)}$

Theorem

4.1. *The finite transformations generated by vector fields v_1, v_2 and v_3 with infinitesimal transformations X_1, X_2 and X_3 may be identified with the projective group of the real line and the group $SL(2,\mathbb{R})$ of unimodular ($\det = 1$) 2×2 real matrices,*

$$\text{by identifying the composition} \quad g.q = \frac{aq+b}{cq+d} \quad \text{with the } SL(2,\mathbb{R}) \text{ matrices} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \begin{bmatrix} 1 & \epsilon_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} e^{-\epsilon_2/2} & 0 \\ 0 & e^{\epsilon_2/2} \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \epsilon_3 & 1 \end{bmatrix}. \tag{4.2}$$

Exercise. Show that these 2×2 matrices form a three-parameter Lie group. ★

Proof. The projective group transformations of the real line may be identified with the group $SL(2,\mathbb{R})$ of unimodular ($\det = 1$) 2×2 real matrices, as follows

$$g_2.(g_1.q) = \frac{(a_1a_2 + b_1c_2)q + (a_1b_2 + b_1d_2)}{(c_1a_2 + d_1c_2)q + (c_1b_2 + d_1d_2)} \quad \text{and} \quad \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} = \begin{bmatrix} (a_1a_2 + b_1c_2) & (a_1b_2 + b_1d_2) \\ (c_1a_2 + d_1c_2) & (c_1b_2 + d_1d_2) \end{bmatrix}$$

This means the Lie group of nonlinear projective transformations has a linear matrix representation in terms of $SL(2,\mathbb{R})$. □

Commutators. Commutators of the vector fields $v_\alpha = X_\alpha(q)\partial_q$ with $X_1 = 1$, $X_2 = -q$ and $X_3 = -q^2$ are given by

$$[v_1, v_2] = -v_1\,, \quad [v_1, v_3] = 2v_2\,, \quad [v_2, v_3] = -v_3\,.$$

These may be assembled into a commutator table, as

$$X_\alpha(q)\partial_q X_\beta(q) - X_\beta(q)\partial_q X_\alpha(q) =: \big[X_\alpha(q)\partial_q\,,\, X_\beta(q)\partial_q\big] =: [v_\alpha,\, v_\beta] = \tag{4.3}$$

$[\,\cdot\,,\,\cdot\,]$	v_1	v_2	v_3
v_1	0	$-v_1$	$2v_2$
v_2	v_1	0	$-v_3$
v_3	$-2v_2$	v_3	0

or, in index notation,

$$[v_\alpha\,,\,v_\beta]^i = v^j_\alpha \frac{\partial v^i_\beta}{\partial q^j} - v^j_\beta \frac{\partial v^i_\alpha}{\partial q^j} = c^\gamma_{\alpha\beta} v^i_\gamma\,, \tag{4.4}$$

or, upon suppressing Latin indices

$$\big[v_\alpha\,,\,v_\beta\big] = v_\alpha \frac{\partial v_\beta}{\partial q} - v_\beta \frac{\partial v_\alpha}{\partial q} = c^\gamma_{\alpha\beta} v_\gamma\,, \tag{4.5}$$

with

$$c^1_{12} = c^3_{23} = -1 = -c^1_{21} = -c^3_{32}, \qquad c^2_{13} = 2 = -c^2_{31}\,, \tag{4.6}$$

while the other $c^\gamma_{\alpha\beta}$'s are zero.

Anti-homomorphism. We note that *minus* the same commutator table (4.3) arises from the following three linearly independent 2×2 traceless matrices comprising a basis for $\mathfrak{sl}(2,\mathbb{R})$, obtained by taking the derivatives at the identity of the $SL(2,\mathbb{R})$ matrices in (4.2),

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{for which} \quad [A_\alpha,\,A_\beta] =$$

$[\,\cdot\,,\,\cdot\,]$	A_1	A_2	A_3
A_1	0	A_1	$-2A_2$
A_2	$-A_1$	0	A_3
A_3	$2A_2$	$-A_3$	0

This overall relative minus sign means the matrix commutation relations will match the vector field commutation relations, provided we define the *Jacobi-Lie bracket* of vector fields to be

$$\big[v_\alpha\,,\,v_\beta\big]_{JL} = \frac{\partial v_\alpha}{\partial q} v_\beta - \frac{\partial v_\beta}{\partial q} v_\alpha$$

▲

Back to Poincaré [Po1901]. Since these transformations form a group, G, the infinitesimal transformations of this group (i.e., elements of T_eG) must satisfy,

$$\Big[X_\alpha(X_\beta) - X_\beta(X_\alpha)\Big]^k = \sum_{i=1}^n \left[X^i_\alpha(q)\frac{\partial X^k_\beta}{\partial q^i} - X^i_\beta(q)\frac{\partial X^k_\alpha}{\partial q^i}\right] = \sum_{\gamma=1}^r c_{\alpha\beta}{}^\gamma(q) X^k_\gamma\,, \quad r = \dim(\mathfrak{g}) \quad \text{with} \quad \mathfrak{g} \simeq T_eG\,. \tag{4.7}$$

This relation follows, because the commutator of any two smooth vector fields representing the infinitesimal transformations produces another smooth vector field. Note that $c_{\alpha\beta}{}^\gamma(q)$ is a function of q, whenever $X^\mu_\alpha(q)$ is nonlinear in q. For the linear case, however, we have $X^i_\alpha(q) = \sum_{j=1}^n [X_\alpha]^i_j q^j$, with *constant* matrices $[X_\alpha]^i_j$, so that equation (4.7) becomes

$$\Big[X_\alpha(X_\beta) - X_\beta(X_\alpha)\Big]^k = \sum_{i,j=1}^n \Big([X_\alpha]^i_j q^j [X_\beta]^k_i - [X_\beta]^i_j q^j [X_\alpha]^k_i\Big) = \sum_{\gamma=1}^r c_{\alpha\beta}{}^\gamma [X_\gamma]^k_j q^j \,. \tag{4.8}$$

In this case, the $c_{\alpha\beta}{}^\gamma$ are matrix Lie algebra structure *constants*, given by the matrix commutation relation

$$\Big[X_\alpha\,,\,X_\beta\Big]^k_j := \Big[X_\alpha X_\beta - X_\beta X_\alpha\Big]^k_j = \sum_{i=1}^n \Big([X_\alpha]^i_j [X_\beta]^k_i - [X_\beta]^i_j [X_\alpha]^k_i\Big) = \sum_{\gamma=1}^r c_{\alpha\beta}{}^\gamma [X_\gamma]^k_j \,. \tag{4.9}$$

Since the group is transitive, we can write for the velocity of the evolution of the dynamical system

$$\dot q^i(t) := \frac{dq^i}{dt} = \sum_{\alpha=1}^r \eta^\alpha(t) X^i_\alpha(q) = \eta^1(t)X^i_1 + \eta^2(t)X^i_2 + \cdots + \eta^r(t)X^i_r \,, \tag{4.10}$$

in such a way that we can go from the state $(q^1,\dots,q^n)$ of the system to a state $(q^1+\dot q^1 \mathrm{d}t,\dots,q^n+\dot q^n\mathrm{d}t)$ by using the infinitesimal transformation of the group,

$$\frac{df}{dt} = \sum_{\alpha=1}^r \eta^\alpha X_\alpha(f)\,. \tag{4.11}$$

Equation (4.10) allows us to rewrite the functional dependence of the kinetic energy as

$$T(q,\dot q) \to T(q,\eta)\,.$$

That is, instead of being expressed as a function of the q and $\dot q$, the kinetic energy can be written as a function of the η and q.

If we perturb the η and q by virtual displacements $\delta\eta$ and δq, respectively, the following virtual displacements in T and V will result,

$$\delta T = \sum_{\alpha=1}^r \frac{\partial T}{\partial \eta^\alpha}\delta\eta^\alpha + \sum_{i=1}^n \frac{\partial T}{\partial q^i}\delta q^i \quad\text{and}\quad \delta V = \sum_{i=1}^n \frac{\partial V}{\partial q^i}\delta q^i\,. \tag{4.12}$$

Since the group is transitive, one will also be able to write the virtual displacement

$$\delta q^i = \sum_{\alpha=1}^r \xi^\alpha X^i_\alpha(q) = \xi^1 X^i_1 + \xi^2 X^i_2 + \cdots + \xi^r X^i_r \,, \tag{4.13}$$

in such a way that we can go from the state q^i of the system to the state $q^i + \delta q^i$ by using the infinitesimal transformation of the group, $\delta q^i = \sum_{\alpha=1}^r \xi^\alpha X^i_\alpha(q)$.

One may then write[4]

$$\delta T - \delta V = \sum_{\alpha=1}^{r} \frac{\partial T}{\partial \eta^\alpha}\delta\eta^\alpha + \sum_{i=1}^{n}\left(\frac{\partial T}{\partial q^i} - \frac{\partial V}{\partial q^i}\right)\delta q^i =: \sum_{\alpha=1}^{r} \frac{\partial T}{\partial \eta^\alpha}\delta\eta^\alpha + \sum_{\alpha=1}^{r}\Xi_\alpha \xi^\alpha\,, \tag{4.14}$$

with $\Xi_\alpha(q,\eta)$ defined by

$$\Xi_\alpha(q,\eta) := \sum_{i=1}^{n}\left(\frac{\partial T}{\partial q^i} - \frac{\partial V}{\partial q^i}\right)X^i_\alpha(q)\,.$$

Next, let the action integral in Hamilton's principle be $S = \int (T - V)\,\mathrm{d}t$, so we will have

$$0 = \delta S = \int \left(\sum_{\gamma=1}^{r}\frac{\partial T}{\partial \eta^\gamma}\delta\eta^\gamma + \sum_{\gamma=1}^{r}\Xi_\gamma\xi^\gamma\right)\mathrm{d}t\,, \tag{4.15}$$

and we can easily find (by equality of cross derivatives for $\frac{d}{dt}\delta q^i = \delta \frac{d}{dt} q^i$) that

$$\delta\eta^\gamma = \frac{d\xi^\gamma}{dt} + \sum_{\alpha,\beta=1}^{r} c_{\alpha\beta}{}^\gamma(q)\,\eta^\beta\xi^\alpha =: \frac{d\xi^\gamma}{dt} + \big[\eta\,,\,\xi\big]^\gamma =: \frac{d\xi^\gamma}{dt} + \big(\mathrm{ad}_\eta\,\xi\big)^\gamma\,. \tag{4.16}$$

The principle of stationary action then gives

$$\frac{d}{dt}\frac{\partial T}{\partial \eta^\gamma} = \sum c_{\gamma\beta}{}^\alpha(q)\,\frac{\partial T}{\partial \eta^\alpha}\eta^\beta + \Xi_\gamma\,. \tag{4.17}$$

Proof.

$$\begin{aligned}
0 = \delta S &= \int_a^b \sum_{\gamma=1}^{r}\left(\frac{\partial T}{\partial \eta^\gamma}\delta\eta^\gamma + \Xi_\gamma\xi^\gamma\right)\mathrm{d}t \\
&= \int_a^b \sum_{\gamma=1}^{r}\left(\frac{\partial T}{\partial \eta^\gamma}\frac{d\xi^\gamma}{dt} + \sum_{\alpha,\beta=1}^{r}\frac{\partial T}{\partial \eta^\alpha}c^\alpha_{\gamma\beta}(q)\,\eta^\beta\xi^\gamma + \Xi_\gamma\xi^\gamma\right)\mathrm{d}t \\
&= \int_a^b \sum_{\gamma=1}^{r}\left(-\frac{d}{dt}\frac{\partial T}{\partial \eta^\gamma} + \sum_{\alpha,\beta=1}^{r}\frac{\partial T}{\partial \eta^\alpha}c^\alpha_{\gamma\beta}(q)\,\eta^\beta + \Xi_\gamma\right)\xi^\gamma\,\mathrm{d}t + \left[\sum_{\gamma=1}^{r}\frac{\partial T}{\partial \eta^\gamma}\xi^\gamma\right]_a^b\,.
\end{aligned} \tag{4.18}$$

[4]Here Poincaré's formula reveals for the linear case that $\Xi_\alpha = \sum_{i,j=1}^n \frac{\partial L}{\partial q^i}(\,[X_\alpha]^i_j q^j\,)$ with $L = T - V$, or equivalently $\Xi = \frac{\partial L}{\partial q}\diamond q$ in modern notation.

We will rewrite this derivation in terms of modern ad and ad* notation, then convert to Hamel equations.

Setting

$$\sum_{\gamma=1}^{r}\frac{\partial T}{\partial\eta^\gamma}(\mathrm{ad}_\eta\,\xi)^\gamma=\sum_{\gamma=1}^{r}\left(\mathrm{ad}^*_\eta\frac{\partial T}{\partial\eta}\right)_\gamma\xi^\gamma \tag{4.19}$$

yields the more compact form of equation (4.17)

$$\frac{d}{dt}\frac{\partial T}{\partial\eta}=\mathrm{ad}^*_\eta\frac{\partial T}{\partial\eta}+\Xi\,. \tag{4.20}$$

If we now write $\ell(q,\eta)=T(q,\eta)-V(q)$, then these equations may be written even more compactly as

$$\frac{d}{dt}\frac{\partial\ell}{\partial\eta}=\mathrm{ad}^*_\eta\frac{\partial\ell}{\partial\eta}+\big[\ell\big]\,, \tag{4.21}$$

with $[\ell]$ denoting the standard directional derivative of $\ell(\eta,q)$ along $X^i_\alpha(q)$. That is,

$$\big[\ell\big]_\alpha:=\sum_{i=1}^{n}\frac{\partial\ell}{\partial q^i}X^i_\alpha(q)\,. \tag{4.22}$$

The rearrangement of equation (4.20) into the form (4.21) was first noticed by Hamel in [Ha1904]. For a modern review of Hamel's form of the dynamical equations of mechanics, see [BlMaZe2009].

□

Poincaré mentioned that these equations encompass some particular cases:

(1.) The Euler–Lagrange equations arise when the infinitesimal transformations are simply translations, all commuting amongst each other, which each shift one of the variables q by an infinitesimally small constant.

In this case, $r=n$, since the number of independent translations is the dimension of space. In this case, $X^i_\alpha=\delta^i_\alpha$, and

$$\dot q^i=\sum_{\alpha=1}^{r}\eta^\alpha X^i_\alpha=\sum_{\alpha=1}^{r}\eta^\alpha\delta^i_\alpha=\eta^i(t)\,.$$

Then

$$\frac{df}{dt}=\sum_{\alpha=1}^{r}\eta^\alpha X_\alpha(f)=\sum_{\alpha=1}^{r}\eta^\alpha X^i_\alpha\frac{\partial f}{\partial q^i}=\sum_{\alpha=1}^{r}\eta^\alpha\delta^i_\alpha\frac{\partial f}{\partial q^i}=\sum_{i=1}^{n}\dot q^i\frac{\partial f}{\partial q^i}\,.$$

In addition, $c_{\alpha\beta}{}^{\gamma} = 0$, so

$$\frac{d}{dt}\frac{\partial \ell}{\partial \eta} = \mathrm{ad}^*_\eta \frac{\partial \ell}{\partial \eta} + \big[\ell\big] \quad \text{with} \quad [\ell]_\alpha = \sum_{i=1}^n \frac{\partial \ell}{\partial q^i}\delta^i_\alpha \quad \Longrightarrow \quad \frac{d}{dt}\frac{\partial \ell}{\partial \dot q^i} = \frac{\partial \ell}{\partial q^i}\,, \tag{4.23}$$

which is the Euler-Lagrange equation.

(2.) Also, Euler's equations for solid body rotations emerge in the linear case, in which the role of the η^α is played by the components p, q, r of the rotations and the role of Ξ by the coupled external forces. Suppose the mass density distribution is given by $\rho(q_0)$ at points $q_0 \in \mathcal{B}$ inside the rigid body, which rotates with a time dependent orientation given by $q(t) = O(t)q_0$, where $O(t)$ is a time-dependent path in the three-dimensional rotation group $SO(3)$. Then we have $\dot q = \dot O(t)q_0$ and we have lifted the three-dimensional motion $(q, \dot q) \in T\mathbb{R}^3$ of the position and velocity of a point in the rigid body, to the motion $(O, \dot O) \in TSO(3)$ in the tangent space of the rotation group.

Euler's rigid body equations then emerge from Hamilton's variational principle for the following kinetic energy Lagrangian,

$$L(O, \dot O) = \frac12 \int_{\mathcal{B}} \rho(q_0)|\dot O(t)q_0|^2 d^3q_0$$

This Lagrangian is invariant under left multiplication by any fixed element of $SO(3)$, so it makes sense to transform to left invariant variable $\widehat\Omega := O^{-1}\dot O$. Note that the Lagrangian is not invariant under right rotations, because these would change the mass distribution, and $\rho(Oq_0) \neq \rho(q_0)$, by assumption. By left invariance, we have

$$\begin{aligned} L(O, \dot O) &= \frac12 \int_{\mathcal{B}} \rho(q_0)|O^{-1}\dot O(t)q_0|^2 d^3q_0 = \frac12 \int_{\mathcal{B}} \rho(q_0)|\widehat\Omega(t)q_0|^2 d^3q_0 \\ &= \frac12 \int_{\mathcal{B}} \rho(q_0)\mathrm{tr}\big((\widehat\Omega(t)q_0)^T\widehat\Omega(t)q_0\big)d^3q_0 = \frac12 \int_{\mathcal{B}} \rho(q_0)\widehat\Omega_{ij}(t)q_0^jq_0^k\widehat\Omega_{ki}(t)\,d^3q_0 = \frac12\widehat\Omega_{ij}(t)I^{jk}_{\mathcal{B}}\widehat\Omega_{ki}(t) = \ell(\widehat\Omega)\,, \end{aligned}$$

in which $I^{jk}_{\mathcal{B}}$ is the "moment of inertia" of the body. Upon identifying $\eta = \widehat\Omega$, Poincaré's equation (4.20) will recover Euler's rigid body equations. To see this easily, one should notice that $\widehat\Omega^T = -\widehat\Omega$ and $c^\gamma_{\alpha\beta} = \epsilon_{\alpha\beta\gamma}$ for this case, with $\epsilon_{123} = +1$. To write Euler's rigid body equations in their usual vector form, one may identify the Lie algebra $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by writing the "hat map", as $\widehat\Omega_{ij} = -\epsilon_{ijk}\Omega^k$. Then the Lagrangian $\ell(\widehat\Omega) \in \mathfrak{so}(3)$ may be written in terms of the vector $\mathbf{\Omega} \in \mathbb{R}^3$ as

$$\ell(\mathbf{\Omega}) = \frac12 \int_{\mathcal{B}} \rho(q_0)|\mathbf{\Omega} \times \boldsymbol{q}_0|^2\,d^3q_0 = \frac12\Omega^j\left(\int_{\mathcal{B}} \rho(q_0)\left(|q_0|^2Id - q_0\otimes q_0\right)_{jk}\,d^3q_0\right)\Omega^k =: \frac12\mathbf{\Omega}\cdot\mathcal{I}\mathbf{\Omega}\,, \tag{4.24}$$

in which the moment of inertia tensor $\mathcal{I}$ is defined in its classical form.

Exercise. Verify Poincaré's claims (1.) and (2.) above. ★

As a final comment, Poincaré remarked that these equations will be of special interest when the potential V is absent and the kinetic energy T only depends on the η, so Ξ vanishes. This remark was accurate, as we see with the rigid body example! And more remains to be done in the nonlinear case!

Exercise. The rigid body example. ★

4.2 Clebsch (1859): Constrained variations and transformation to Hamiltonian form

Although Poincaré [Po1901] used a version of what we would now call reduction by symmetry, let's now present an earlier approach, due to Clebsch [Clebsch1859], which introduces constrained variations into Hamilton's principle by imposing velocity maps corresponding in the deterministic case to the infinitesimal transformations of a Lie group. We will observe later that the velocity maps could also have been stochastic.

In what follows, we will introduce a few modern conventions which simplify the notation. For example, we will sum repeated indices over their range, without writing the summation symbol (Einstein's convention). Let $\mathfrak{g}$ denote a Lie algebra with elements $u^\alpha \in \mathfrak{g}$, where $\alpha = 1, 2, \dots, \dim(\mathfrak{g}) = r$. Let Q be a manifold and suppose the corresponding Lie group G acts transitively on Q. Namely, for any element $u \in \mathfrak{g}$ there corresponds a vector field defined on Q, denoted u_Q, which can be written in local coordinates as

$$u_Q(q) = u^i_Q(q)\frac{\partial}{\partial q^i} = A^i_\alpha(q)u^\alpha\frac{\partial}{\partial q^i}\,,$$

where A_α are vector fields on Q, $i = 1, ..., \dim(Q) = n$.

Let us consider the following constrained Hamilton's principle, $\delta S = 0$, with variation of the action integral $S(u, p, q)$ given by

$$0 = \delta S(u,p,q) = \delta\int_0^T \ell(u)dt + \left\langle p_i, \frac{dq^i}{dt} - u^i_Q(q)\right\rangle dt \quad \text{with} \quad u^i_Q(q) = A^i_\alpha(q)u^\alpha\,. \tag{4.25}$$

The variations are taken by embedding the paths $q^i(t,\epsilon)$ and $u^\alpha(t,\epsilon)$ into a family of nearby curves, parameterised by ϵ with $\epsilon = 0$ denoting the solution curves obtained from Hamilton's principle with variational operator given by

$$\delta q^i(t) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} q^i(t,\epsilon) \quad\text{and}\quad \delta u^\alpha(t) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} u^\alpha(t,\epsilon)\,. \tag{4.26}$$

Performing the variations in (4.25) yields

$$\begin{aligned} 0 = \delta S(u,p,q) = \int_0^T \Big\langle \frac{\partial \ell(u,q)}{\partial u^\alpha} - p_i A^i_\alpha(q)\,,\, \delta u^\alpha \Big\rangle + \Big\langle \delta p_i, \frac{dq^i}{dt} - A^i_\alpha(q) u^\alpha \Big\rangle \\ - \Big\langle \frac{dp_i}{dt} + p_j \frac{\partial A^j_\alpha(q)}{\partial q^i} u^\alpha - \frac{\partial \ell}{\partial q^i}\,,\, \delta q^i \Big\rangle\, dt + \big\langle p_i\,,\, \delta q^i \big\rangle \Big|_0^T . \end{aligned} \tag{4.27}$$

Therefore, if we choose variations which vanish at the endpoints, so that $\delta q^i(0) = 0 = \delta q^i(T)$, then we can ignore the endpoint terms arising from integration by parts in time, and deduce the conditions for Hamilton's principle to be satisfied, as

$$\begin{aligned} \delta u^\alpha :&\quad \frac{\partial \ell}{\partial u^\alpha} - p_i A^i_\alpha(q) = 0\,, \\ \delta p_i :&\quad \frac{dq^i}{dt} - A^i_\alpha(q) u^\alpha = 0\,, \\ \delta q^i :&\quad \frac{dp_i}{dt} + p_j \frac{\partial A^j_\alpha(q)}{\partial q^i} u^\alpha - \frac{\partial \ell}{\partial q^i} = 0\,. \end{aligned} \tag{4.28}$$

The map $m : T^*Q \to \mathfrak{g}^*$ given by

$$m_\alpha(q,p) = \frac{\partial \ell}{\partial u^\alpha} = p_i A^i_\alpha(q) \tag{4.29}$$

is called the *momentum map*. We will have much more to say about momentum maps, later. For now, we remark that inserting $\delta q^i = u^i_Q(q) = A^i_\alpha(q) u^\alpha$ into the endpoint term in (4.27) naturally contains the momentum map, as $\langle p_i A^i_\alpha(q)\,,\, u^\alpha \rangle$.

Lemma

4.2. *Hamilton's principle $\delta S = 0$ for the constrained action integral in (4.25) recovers Poincaré's equation (4.21) for right action (with an appropriate minus sign),*

$$\frac{d}{dt}\frac{\partial \ell}{\partial u^\alpha} = \Big(\mathrm{ad}^*_u \frac{\partial \ell}{\partial u} \Big)_\alpha + \big[\ell\big]_\alpha\,, \tag{4.30}$$

with $[\ell]$ denoting the standard directional derivative of $\ell(\eta, q)$ along $A^i_\alpha(q)$. That is,

$$\big[\ell\big]_\alpha := \sum_{i=1}^{n} \frac{\partial \ell}{\partial q^i} A^i_\alpha(q)\,. \tag{4.31}$$

Proof. For the proof, we keep in mind Poincaré's equation (4.7), rewritten in this notation as

$$[\,A_\alpha\,,\,A_\beta\,]^j = A^i_\alpha \frac{\partial A^j_\beta}{\partial q^i} - A^i_\beta \frac{\partial A^j_\alpha}{\partial q^i} = c^\gamma_{\alpha\beta}(q) A^j_\gamma\,.$$

We compute directly from the variational equations (4.28) that,

$$\begin{aligned}
\frac{d}{dt}\frac{\delta \ell}{\delta u^\alpha} &= \frac{d}{dt}\big(p_i A^i_\alpha(q)\big) = \left(\frac{dp_i}{dt}\right) A^i_\alpha(q) + p_i \frac{\partial A^i_\alpha}{\partial q^j}\frac{dq^j}{dt} \\
\text{By equation (4.28)} \quad &= -\,p_i \left(A^j_\alpha \frac{\partial A^i_\beta}{\partial q^j} - A^j_\beta \frac{\partial A^i_\alpha}{\partial q^j} \right) u^\beta + A^i_\alpha(q) \frac{\partial \ell}{\partial q^i} \\
\text{By equation (4.7)} \quad &= -\,p_i c^\gamma_{\alpha\beta} A^i_\gamma(q)\, u^\beta + A^i_\alpha(q) \frac{\partial \ell}{\partial q^i} \\
\text{By equation (4.28)} \quad &= -\,c^\gamma_{\alpha\beta} \frac{\partial \ell}{\partial u^\gamma}\, u^\beta + A^i_\alpha(q) \frac{\partial \ell}{\partial q^i} \\
&=: \left(\mathrm{ad}^*_u \frac{\partial \ell}{\partial u} \right)_\alpha + \big[\ell\big]_\alpha\,.
\end{aligned} \tag{4.32}$$

□

Legendre transformation to canonical Hamiltonian form Next, we will make the *Legendre transformation* from the Lagrangian $\ell : TM \to \mathbb{R}$ to the Hamiltonian $h : T^*M \to \mathbb{R}$ via

$$h(p,q) = \left\langle p_i, \frac{dq^i}{dt} \right\rangle - \ell(u,q) - \left\langle p_i, \frac{dq^i}{dt} - A^i_\alpha(q) u^\alpha \right\rangle = \big\langle p_i, A^i_\alpha(q) u^\alpha \big\rangle - \ell(u,q)\,, \tag{4.33}$$

with variations

$$\delta h(q,p) = \left\langle p_i A^i_\alpha(q) - \frac{\partial \ell}{\partial u^\alpha}\,,\, \delta u^\alpha \right\rangle + \big\langle A^i_\alpha(q) u^\alpha\,,\, \delta p_i \big\rangle + \left\langle p_j \frac{\partial A^j_\alpha(q)}{\partial q^i} u^\alpha - \frac{\partial \ell}{\partial q^i}\,,\, \delta q^i \right\rangle , \tag{4.34}$$

so that, upon identifying coefficients,

$$\frac{\partial h}{\partial u^\alpha} = 0 = p_i A^i_\alpha(q) - \frac{\partial \ell}{\partial u^\alpha}\,, \quad \frac{\partial h}{\partial p_i} = A^i_\alpha(q) u^\alpha \quad \text{and} \quad \frac{\partial h}{\partial q^i} = p_j \frac{\partial A^j_\alpha(q)}{\partial q^i} u^\alpha - \frac{\partial \ell}{\partial q^i}\,. \tag{4.35}$$

These three equations recover the equations in (4.28). The momentum map in (4.29) appears again in the first of these equations. The other two equations in (4.35) convert the corresponding equations in (4.28) into their Hamiltonian forms,

$$\begin{aligned}\frac{dq^i}{dt} &= \{q^i\,,\,h\}_{can} = \frac{\partial h}{\partial p_i} = A^i_\alpha(q) u^\alpha\,,\\ \frac{dp_i}{dt} &= \{p_i\,,\,h\}_{can} = -\frac{\partial h}{\partial q^i} = -p_j \frac{\partial A^j_\alpha(q)}{\partial q^i} u^\alpha + \frac{\partial \ell}{\partial q^i}\,.\end{aligned} \tag{4.36}$$

Here, the canonical Poisson bracket symbol $\{\,\cdot\,,\,\cdot\,\}_{can} : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ for smooth functions $f \in \mathcal{F} : T^*M \to \mathbb{R}$ denotes

$$\{f\,,\,h\}_{can} = \frac{\partial f}{\partial q^i}\frac{\partial h}{\partial p_i} - \frac{\partial h}{\partial q^i}\frac{\partial f}{\partial p_i}\,. \tag{4.37}$$

Legendre transformation to non-canonical Hamiltonian form Another Legendre transformation involving the momentum map in (4.29) may be constructed, by defining

$$\mathcal{H}(m,q) = \langle m_\alpha\,,\,u^\alpha\rangle - \ell(u,q)\,. \tag{4.38}$$

The variations of this Hamiltonian are given by

$$\delta\mathcal{H}(m,q) = \langle \delta m_\alpha\,,\,u^\alpha\rangle + \left\langle m_\alpha - \frac{\partial \ell}{\partial u^\alpha}\,,\,\delta u^\alpha\right\rangle - \left\langle \frac{\partial \ell}{\partial q}\,,\,\delta q\right\rangle \tag{4.39}$$

which implies the relations

$$\frac{\partial \mathcal{H}}{\partial u^\alpha} = 0 = m_\alpha - \frac{\partial \ell}{\partial u^\alpha}\,, \quad \frac{\partial \mathcal{H}}{\partial m_\alpha} = u^\alpha \quad \text{and} \quad \frac{\partial \mathcal{H}}{\partial q^i} = -\frac{\partial \ell}{\partial q^i}\,. \tag{4.40}$$

The Legendre transformation (4.38) involving the momentum map will allow us to transform the canonical Hamiltonian equations (4.36) into the following non-canonical form of the Euler-Poincaré equations in (4.20) and (4.21).

$$\begin{aligned}\frac{dm_\alpha}{dt} &= \{m_\alpha\,,\,\mathcal{H}\} = \{m_\alpha\,,\,m_\beta\}\frac{\partial \mathcal{H}}{\partial m_\beta} + \{m_\alpha\,,\,q^i\}\frac{\partial \mathcal{H}}{\partial q^i} = -\,c^\gamma_{\alpha\beta}(q)\,m_\gamma \frac{\partial \mathcal{H}}{\partial m_\beta} - A^i_\alpha(q)\frac{\partial \mathcal{H}}{\partial q^i}\,,\\ \frac{dq^i}{dt} &= \{q^i\,,\,\mathcal{H}\} = \{q^i\,,\,m_\beta\}\frac{\partial \mathcal{H}}{\partial m_\beta} = A^i_\beta(q)\frac{\partial \mathcal{H}}{\partial m_\beta}\end{aligned} \tag{4.41}$$

The corresponding Poisson bracket $\{\mathcal{F},\mathcal{H}\}(m,q)$ on the space of smooth functions of (m_α, q^i) is given by

$$\frac{d\mathcal{F}}{dt} = \begin{bmatrix} \partial\mathcal{F}/\partial m_\alpha \\ \partial\mathcal{F}/\partial q^i \end{bmatrix}^T \begin{bmatrix} dm_\alpha/dt \\ dq^i/dt \end{bmatrix} = \begin{bmatrix} \partial\mathcal{F}/\partial m_\alpha \\ \partial\mathcal{F}/\partial q^i \end{bmatrix}^T \begin{bmatrix} -\,c^\gamma_{\alpha\beta}(q)\,m_\gamma & -\,A^j_\alpha(q) \\ A^i_\beta(q) & 0 \end{bmatrix} \begin{bmatrix} \partial\mathcal{H}/\partial m_\beta \\ \partial\mathcal{H}/\partial q^j \end{bmatrix} dt = \{\mathcal{F},\mathcal{H}\}\,. \tag{4.42}$$

A straightforward proof that the bracket (4.42) satisfies the Jacobi identity and other defining conditions (bilinear, skew symmetric, Leibniz) for a valid Poisson bracket is available, by transforming back to the canonical variables, in which the corresponding Poisson bracket on the extended space $\{f,h\}_{can}(m(q,p),q,p)$ with $m_\alpha(q,p) = p_i A^i_\alpha(q)$ as defined in equation (4.29) is given by

$$\frac{df}{dt} = \begin{bmatrix} \partial f/\partial m_\alpha \\ \partial f/\partial q^i \\ \partial f/\partial p_i \end{bmatrix}^T \begin{bmatrix} dm_\alpha/dt \\ dq^i/dt \\ dp_i/dt \end{bmatrix} = \begin{bmatrix} \partial f/\partial m_\alpha \\ \partial f/\partial q^i \\ \partial f/\partial p_i \end{bmatrix}^T \begin{bmatrix} -\,c^\gamma_{\alpha\beta}(q)\,m_\gamma & -\,A^j_\alpha(q) & p_m\frac{\partial A^m_\alpha}{\partial q^j} \\ A^i_\beta(q) & 0 & \delta^i_j \\ -\,p_l\frac{\partial A^l_\beta}{\partial q^i} & -\,\delta^j_i & 0 \end{bmatrix} \begin{bmatrix} \partial h/\partial m_\beta \\ \partial h/\partial q^j \\ \partial h/\partial p_j \end{bmatrix} = \{f,h\}_{can}(m(q,p),q,p)\,. \tag{4.43}$$

When $h \to \mathcal{H}$ depends on p_i only via m_α, the last row and column of the Hamiltonian matrix may be ignored and (4.42) appears.

Remark

4.3. *When the infinitesimal generators in the velocity map are linear, so that $A^i_\alpha(q) = [A_\alpha]^i_j q^j$, then the Poisson bracket in (4.42) is Lie-Poisson, defined on the dual to a semidirect product Lie algebra action.*

The equations of motion in (4.41) may be rewritten in terms of Lie algebra operations, as

$$\frac{dm_\alpha}{dt} = \left(\mathrm{ad}^*_{\frac{\partial\mathcal{H}}{\partial m}} m\right)_\alpha - A^i_\alpha(q)\frac{\partial\mathcal{H}}{\partial q^i} \tag{4.44}$$

$$\frac{dq^i}{dt} = A^i_\beta(q)\frac{\partial\mathcal{H}}{\partial m_\beta} \tag{4.45}$$

where the $A^i_\beta(q)$ form a nonlinear realisation of the Lie algebra action, in terms of the vector fields

$$\mathcal{A}_\beta(q) := A^i_\beta(q)\frac{\partial}{\partial q^i} \quad \text{and} \quad \mathcal{A}^i(q) := A^i_\beta(q)\frac{\partial}{\partial m_\beta}$$

One may wonder whether we can also write the nonlinear action (4.42) in semidirect-product. This is promising:

$$\frac{d\mathcal{F}}{dt} = \frac{\partial\mathcal{F}}{\partial m_\alpha}\left(\mathrm{ad}^*_{\frac{\partial\mathcal{H}}{\partial m}} m\right)_\alpha - \frac{\partial\mathcal{F}}{\partial m_\alpha}A^i_\alpha(q)\frac{\partial\mathcal{H}}{\partial q^i} + \frac{\partial\mathcal{H}}{\partial m_\alpha}A^i_\alpha(q)\frac{\partial\mathcal{F}}{\partial q^i}$$

When $A^i_\alpha(q) = [A_\alpha]^i_j q^j$ is linear in q, this reduces to the usual linear semidirect-product action.

4.3 Original, H. Poincaré (1901) Translated literally

Sur une forme nouvelle des équations de la méchanique.[5]

C.R. Acad. Sci. **132,** 369-371.

Having had the opportunity to work on the rotational motion of hollow solid bodies filled with liquid, I have been led to cast the equations of mechanics into a new form that could be interesting to know. Assume there are n degrees of freedom and let $\{x^1, ..., x^n\}$ be the variables describing the state of the system. Let T and U be the kinetic and potential energy of the system.

Consider any continuous, transitive group (that is, its action covers the entire manifold). Let $X_i(f)$ be any infinitesimal transformation of this group such that[6]

$$X_i(f) = \sum_{\mu=1}^{n} X_i^\mu(x)\frac{\partial f}{\partial x^\mu} = X_i^1\frac{\partial f}{\partial x^1} + X_i^2\frac{\partial f}{\partial x^2} + \cdots + X_i^n\frac{\partial f}{\partial x^n}\,.$$

Since these transformations form a group, we must have

$$X_iX_k - X_kX_i = \sum_{s=1}^{r} c_{ik}{}^s X_s\,.$$

Since the group is transitive we can write

$$\dot{x}^\mu(t) = \frac{dx^\mu}{dt} = \sum_{i=1}^{r}\eta^i(t)X_i^\mu(x) = \eta^1(t)X_1^\mu + \eta^2(t)X_2^\mu + \cdots + \eta^r(t)X_r^\mu\,,$$

in such a way that we can go from the state $(x^1, \dots, x^n)$ of the system to a state $(x^1 + \dot{x}^1 \mathrm{d}t, \dots, x^n + \dot{x}^n \mathrm{d}t)$ by using the infinitesimal transformation of the group, $\sum_{i=1}^{r}\eta^i X_i(f)$.

T instead of being expressed as a function of the x and $\dot{x}$ can be written as a function of the η and x. If we increase the η and x by virtual displacements $\delta\eta$ and δx, respectively, there will be resulting increases in T and U

$$\delta T = \sum\frac{\delta T}{\delta\eta}\delta\eta + \sum\frac{\delta T}{\delta x}\delta x \quad\text{and}\quad \delta U = \sum\frac{\delta U}{\delta x}\delta x\,.$$

Since the group is transitive I will be able to write

$$\delta x^\mu = \xi^1X_1^\mu + \xi^2X_2^\mu + \cdots + \xi^rX_r^\mu\,,$$

[5] Translation of [Po1901] into English by D. D. Holm and J. Kirsten

[6] As a special case of the finite dimensional systems considered here, the set of infinitesimal transformations $\{X_\alpha\}$ may be restricted to be linear in x, *represented* by a set of r constant $n \times n$ matrices acting linearly on the set of n states $\{x\}$. Then, for example, $X_\alpha^i(x) = \sum_{j=1}^{n}[X_\alpha]^i_jx^j$. (Translator's note)

in such a way that we can go from the state x^μ of the system to the state $x^\mu + \delta x^\mu$ by using the infinitesimal transformation of the group $\delta x^\mu = \sum_{i=1}^r \xi^i X_i(x^\mu)$. I will then write[7]

$$\delta T - \delta U = \sum_{i=1}^{r} \frac{\delta T}{\delta \eta^i}\delta\eta^i + \sum_{\mu=1}^{n}\left(\frac{\delta T}{\delta x^\mu} - \frac{\delta U}{\delta x^\mu}\right)\delta x^\mu = \sum_{i=1}^{r} \frac{\delta T}{\delta \eta^i}\delta\eta^i + \sum_{i=1}^{r}\Omega_i \xi^i .$$

Next, let the Hamilton integral be

$$J = \int (T - U)\,\mathrm{d}t\,,$$

so we will have

$$\delta J = \int \left(\sum \frac{\delta T}{\delta \eta^i}\delta\eta^i + \sum \Omega_i \xi^i\right)\mathrm{d}t\,,$$

and can easily find

$$\delta\eta^i = \frac{d\xi^i}{dt} + \sum_{s,k=1}^{r} c_{sk}{}^i\,\eta^k \xi^s\,.$$

The principle of stationary action then gives

$$\frac{d}{dt}\frac{\delta T}{\delta \eta^s} = \sum c_{sk}{}^i\,\frac{\delta T}{\delta \eta^i}\eta^k + \Omega_s\,.$$

These equations encompass some particular cases:

1. The Lagrange equations, when the group is reduced to the transformations, all commuting amongst each other, which each shift one of the variables x by an infinitesimally small constant.

2. Also, Euler's equations for solid body rotations emerge, in which the role of the η_i is played by the components p, q, r of the rotations and the role of Ω by the coupled external forces.

These equations will be of special interest where the potential U is zero and the kinetic energy T only depends on the η in which case Ω vanishes.

[7] Here Poincaré's formula reveals that $\Omega_i = \sum_{\mu,\nu=1}^n \frac{\delta L}{\delta x^\mu}[X_i]^\mu_\nu x^\nu$ with $L = T - U$, or equivalently $\Omega = \frac{\partial L}{\partial x} \diamond x$ in modern notation.

5 Integrability of rotational motion on $SO(n)$: the rigid body

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

Proposition

5.1 (Manakov [Ma1976]). *Euler's equations for a rigid body on $SO(n)$ take the matrix commutator form,*

$$\frac{dM}{dt} = [\, M \,,\, \Omega \,] \quad \textit{with} \quad M = \mathbb{A}\Omega + \Omega\mathbb{A}\,, \tag{5.1}$$

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 (5.1) follows as the Euler–Lagrange equations for Hamilton's principle $\delta S = 0$ with $S = \int l\, dt$ for the Lagrangian

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

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

$$\delta\Omega = \delta(O^{-1}\dot{O}) = \Xi^{\cdot} + [\,\Omega\,,\,\Xi\,] = \Xi^{\cdot} + \mathrm{ad}_\Omega\,\Xi\,, \quad \text{with} \quad \Xi = (O^{-1}\delta O) \tag{5.2}$$

for $\delta\Omega$ now leads to

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

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

$$\delta S \;=\; \frac{1}{2}\int_a^b \mathrm{tr}\big(\Xi\,(\,\dot{M} + \Omega M - M\Omega\,)\big)\, dt = \frac{1}{2}\int_a^b \big\langle\,\dot{M} - \mathrm{ad}^*_\Omega M\,,\,\Xi\big\rangle\, dt = 0\,.$$

Invoking stationarity for arbitrary Ξ and noting that $\mathrm{ad}^* = -\mathrm{ad}$ for semisimple Lie algebra $\mathfrak{so}(n)$ implies commutator form (5.1). $\square$

5.2 Matrix Euler–Poincaré equations

Manakov's commutator form of the rigid-body equations in (5.1) recalls much earlier work by Poincaré [Po1901], who also noticed that the matrix commutator form of Euler's rigid-body equations suggests an additional mathematical structure going back to Lie's theory of groups of transformations depending continuously on parameters. In particular, Poincaré [Po1901] remarked that the commutator form of Euler's rigid-body equations would make sense for any Lie algebra, not just for $so(3)$. The proof of Manakov's commutator form (5.1) by Hamilton's principle is essentially the same as Poincaré's proof in [Po1901]. The important feature is that $\mathrm{ad}^* = -\mathrm{ad}$ for semisimple Lie algebras such as $so(n)$.

Exercise. Prove that $\mathrm{ad}^* = -\mathrm{ad}$ for semisimple Lie algebras. ★

Theorem

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

$$\frac{dM}{dt} = [\,M\,,\,\Omega\,] \quad \textit{with} \quad M = \frac{\delta l}{\delta \Omega}\,, \tag{5.3}$$

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

Proof. The proof here is the same as the proof of Manakov's commutator formula via Hamilton's principle, modulo replacing $O^{-1}\dot{O} \in so(n)$ with $g^{-1}\dot{g} \in \mathfrak{g}$, and invoking $\mathrm{ad}^* = -\mathrm{ad}$ for semisimple Lie algebras. □

Remark

5.3.
Poincaré's observation leading to the matrix Euler–Poincaré Equation (5.3) was reported in two pages with no references [Po1901]. The proof above shows that the matrix Euler–Poincaré equations possess a natural variational principle. Note that if $\Omega = g^{-1}\dot{g} \in \mathfrak{g}$, then $M = \delta l/\delta\Omega \in \mathfrak{g}^*$, where the dual is defined in terms of the matrix trace pairing.

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

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

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

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

for the evolution of the $n \times n$ skew-symmetric matrices M, 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.

Proposition

5.4. *Isospectrality allows the quadratic rigid-body dynamics (5.1) 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

$$\begin{aligned} 0 = \frac{d}{dt}(M\psi - \lambda\psi) &= \dot{M}\psi + M\dot{\psi} - \lambda\dot{\psi} \\ (\text{By } \dot{\psi} = -\,\Omega\psi) &= \dot{M}\psi - M\Omega\psi + \Omega\lambda\psi \\ (\text{By } M\psi = \lambda\psi) &= \dot{M}\psi - M\Omega\psi + \Omega M\psi = (\dot{M} + [\,\Omega\,,\,M\,])\psi\,. \end{aligned}$$

This recovers (5.1) as $\dot{M} + [\,\Omega\,,\,M\,] = 0$. □

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

Manakov [Ma1976] observed that Equations (5.1) may be "deformed" into

$$\frac{d}{dt}(M + \lambda A) = [(M + \lambda A), (\Omega + \lambda B)]\,, \tag{5.4}$$

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:

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

- 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 (5.4) 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 $n = 4$.

Remark

5.5.
This result generalises considerably. For example, Manakov's method determines the solution for all the algebraically solvable rigid bodies on $SO(n)$. The moments of inertia of these bodies possess only $2n-3$ parameters. (Recall that in Manakov's case for $SO(4)$ the moment of inertia possesses only five parameters.)

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

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

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

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

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

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

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

If we decide that time is short, we will go directly to Section 8

6 Action principles on Lie algebras

6.1 The Euler–Poincaré theorem

In the notation for the AD, Ad and ad actions of Lie groups and Lie algebras, Hamilton's principle (that the equations of motion arise from stationarity of the action) for Lagrangians defined on Lie algebras may be expressed as follows. This is the Euler–Poincaré theorem [Po1901].

Theorem

6.1 (Euler–Poincaré theorem)**.** *Stationarity*

$$\delta S(\xi) = \delta \int_a^b l(\xi)\, dt = 0 \tag{6.1}$$

of an action

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

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

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi}\,, \tag{6.2}$$

for variations of the left-invariant Lie algebra element

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

that are restricted to the form

$$\delta\xi = \dot{\eta} + \mathrm{ad}_\xi\, \eta\,, \tag{6.3}$$

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

Exercise. What is the solution to the Euler–Poincaré Equation (6.2) in terms of $\mathrm{Ad}^*_{g(t)}$?

Hint: Take a look at the earlier equation (3.15). ★

Remark

6.2. *Such variations are defined for any Lie algebra.*

Proof. A direct computation proves Theorem 6.1. Later, we will explain the source of the constraint (6.3) on the form of the variations on the Lie algebra. One verifies the statement of the theorem by computing with a nondegenerate pairing $\langle \cdot \,, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$,

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

upon integrating by parts. The last term vanishes, by the endpoint conditions, $\eta(b) = \eta(a) = 0$.
Since $\eta(t) \in \mathfrak{g}$ is otherwise arbitrary, (6.1) is equivalent to

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

which recovers the Euler–Poincaré Equation (6.2) in the statement of the theorem. □

Corollary

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

Proof. Consider the endpoint terms $\langle \frac{\delta l}{\delta \xi}\,,\,\eta\rangle\big|_a^b$ arising in the variation δS in (6.1) and note that this implies for any time $t \in [a,b]$ that

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

when the Euler–Poincaré Equations (6.2) are satisfied. $\square$

Corollary

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

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

as a consequence of the Euler–Poincaré Equation (6.2).

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

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

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

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

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

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

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

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

The constancy of this quantity under the Euler–Poincaré dynamics in (6.2) is verified, upon taking the time derivative and using the coadjoint motion relation (3.14) in Proposition 3.5. $\square$

Remark

6.5. *The form of the variation in (6.3) arises directly by*

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

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

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

Namely, one computes,

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

On taking the difference, the terms with cross derivatives cancel and one finds the variational formula (6.3),

$$
\xi\,' - \dot{\eta} = [\,\xi\,,\,\eta\,] \quad \textit{with} \quad [\,\xi\,,\,\eta\,] := \xi\,\eta - \eta\,\xi = \mathrm{ad}_\xi\,\eta\,. \tag{6.4}
$$

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

Example

6.6 (Euler–Poincaré equation for $SE(3)$)**.** *The Euler–Poincaré Equation (6.2) for* $SE(3)$ *is equivalent to*

$$
\left(\frac{d}{dt}\frac{\delta l}{\delta \xi}\,,\ \frac{d}{dt}\frac{\delta l}{\delta \alpha} \right) = \mathrm{ad}^*_{(\xi\,,\,\alpha)} \left(\frac{\delta l}{\delta \xi}\,,\ \frac{\delta l}{\delta \alpha} \right)\,. \tag{6.5}
$$

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

6.2 Hamilton–Pontryagin principle

Formula (6.4) for the variation of the vector $\xi = g^{-1}\dot{g} \in \mathfrak{g}$ may be imposed as a constraint in Hamilton's principle and thereby provide an immediate derivation of the Euler–Poincaré Equation (6.2). This constraint is incorporated into the following theorem.

Theorem

6.7 (Hamilton–Pontryagin principle). *The Euler–Poincaré equation*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} \tag{6.6}$$

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

$$\delta S(\xi, g, \dot{g}) = \delta \int_a^b l(\xi, g, \dot{g})\, dt = 0, \tag{6.7}$$

for a constrained action

$$\begin{aligned} S(\xi, g, \dot{g}) &= \int_a^b l(\xi, g, \dot{g})\, dt \\ &= \int_a^b \Big[l(\xi) + \langle\, \mu\,,\, (g^{-1}\dot{g} - \xi)\,\rangle \Big]\, dt\,. \end{aligned} \tag{6.8}$$

Proof. The variations of S in formula (6.8) are given by

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

Substituting $\delta(g^{-1}\dot{g})$ from (6.4) into the last term produces

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

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

$$\frac{\delta l}{\delta \xi} = \mu\,, \quad g^{-1}\dot{g} = \xi\,, \quad \dot{\mu} = \mathrm{ad}^*_\xi\,\mu\,. \tag{6.9}$$

$\square$

Remark

6.8 (Interpreting the variational formulas (6.9)).

The first formula in (6.9) is the fibre derivative needed in the Legendre transformation $\mathfrak{g} \mapsto \mathfrak{g}^*$, for passing to the Hamiltonian formulation. The second is the reconstruction formula for obtaining the solution curve $g(t) \in G$ on the Lie group G given the solution $\xi(t) = g^{-1}\dot{g} \in \mathfrak{g}$. The third formula in (6.9) is the Euler–Poincaré equation on $\mathfrak{g}^*$. The interpretation of Noether's theorem in Corollary 6.4 transfers to the Hamilton–Pontryagin variational principle as preservation of the quantity

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

under the Euler–Poincaré dynamics.

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

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

6.3 Clebsch approach to Euler–Poincaré

The Hamilton–Pontryagin (HP) Theorem 6.7 elegantly delivers the three key formulas in (6.9) needed for deriving the Lie–Poisson Hamiltonian formulation of the Euler–Poincaré equation. Perhaps surprisingly, the HP theorem accomplishes this without invoking any properties of how the invariance group of the Lagrangian G acts on the configuration space M.

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

Proposition

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

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} \tag{6.10}$$

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

$$\delta S(\xi, q, \dot q, p) = \delta \int_a^b l(\xi, q, \dot q, p)\, dt = 0, \tag{6.11}$$

for an action constrained by the reconstruction formula

$$\begin{aligned} S(\xi, q, \dot q, p) &= \int_a^b l(\xi, q, \dot q, p)\, dt \\ &= \int_a^b \Big[l(\xi) + \Big\langle\!\Big\langle\, p\,,\, \dot q + \pounds_\xi q \,\Big\rangle\!\Big\rangle \Big]\, dt\,, \end{aligned} \tag{6.12}$$

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

Remark

6.10. *The Lagrange multiplier p in the second term of (6.12) imposes the constraint*

$$\dot q + \pounds_\xi q = 0\,. \tag{6.13}$$

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

6.4 Recalling the definition of the Lie derivative

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

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

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

and as for the spatial angular velocity,

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

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

6.4.1 Right-invariant velocity vector

The Lie derivative $\pounds_\xi$ appearing in the reconstruction relation $\dot q = -\pounds_\xi q$ in (6.13) is defined via the Lie group operation on the configuration space exactly as for free rotation. For example, one computes the tangent vectors to the motion induced by the group operation acting from the left as $q(t) = g(t)q(0)$ by differentiating with respect to time t,

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

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

6.4.2 Left-invariant velocity vector

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

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

in which the quantity

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

is the left-invariant tangent vector.

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

Exercise. Compute the time derivatives and thus the forms of the right- and left-invariant velocity vectors for the group operations by the inverse $q(t) = q(0)g^{-1}(t)$ and $q(t) = g^{-1}(t)q(0)$. Observe the equivalence (up to a sign) of these velocity vectors with the vectors ξ_R and ξ_L, respectively. Note that the reconstruction formula (6.13) arises from the latter choice. ★

6.5 Clebsch Euler–Poincaré principle

Let us first define the concepts and notation that will arise in the course of the proof of Proposition 6.9.

Definition

6.11 (The diamond operation $\diamond$)**.** *The diamond operation* ($\diamond$) *in Equation (6.17) is defined as minus the dual of the Lie derivative with respect to the pairing induced by the variational derivative in* q, *namely,*

$$\Big\langle\, p \diamond q \,,\, \xi \,\Big\rangle = \Big\langle\!\!\Big\langle\, p \,,\, -\pounds_\xi q \,\Big\rangle\!\!\Big\rangle . \tag{6.14}$$

Definition

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

$$\Big\langle\!\!\Big\langle\, \pounds_\xi^T p \,,\, q \,\Big\rangle\!\!\Big\rangle = \Big\langle\!\!\Big\langle\, p \,,\, \pounds_\xi q \,\Big\rangle\!\!\Big\rangle . \tag{6.15}$$

Proof. The variations of the action integral

$$S(\xi, q, \dot q, p) = \int_a^b \Big[l(\xi) + \Big\langle\!\!\Big\langle\, p \,,\, \dot q + \pounds_\xi q \,\Big\rangle\!\!\Big\rangle \Big]\, dt \tag{6.16}$$

from formula (6.12) are given by

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

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

$$
\frac{\delta l}{\delta \xi} = p\diamond q\,, \quad \dot q = -\,\pounds_\xi q\,, \quad \dot p = \pounds_\xi^T p\,. \tag{6.17}
$$

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

$$
\langle\,\cdot\,,\,\cdot\,\rangle : \mathfrak{g}^*\times\mathfrak{g}\mapsto\mathbb{R} \quad\text{and}\quad \langle\!\langle\,\cdot\,,\,\cdot\,\rangle\!\rangle : T^*M\times TM\mapsto\mathbb{R}\,. \tag{6.18}
$$

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

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

$$
\begin{aligned}
\Big\langle \frac{d}{dt}\frac{\delta l}{\delta \xi}\,,\,\eta \Big\rangle &= \frac{d}{dt}\Big\langle \frac{\delta l}{\delta \xi}\,,\,\eta \Big\rangle\,, \\
\text{[Definition of } \diamond] &= \frac{d}{dt}\Big\langle p\diamond q\,,\,\eta \Big\rangle = \frac{d}{dt}\Big\langle\!\Big\langle p\,,\,-\pounds_\eta q \Big\rangle\!\Big\rangle\,, \\
\text{[Equations (6.17)]} &= \Big\langle\!\Big\langle \pounds_\xi^T p\,,\,-\pounds_\eta q \Big\rangle\!\Big\rangle + \Big\langle\!\Big\langle p\,,\,\pounds_\eta \pounds_\xi q \Big\rangle\!\Big\rangle\,, \\
\text{[Transpose, } \diamond \text{ and ad]} &= \Big\langle\!\Big\langle p\,,\,-\pounds_{[\xi,\eta]} q \Big\rangle\!\Big\rangle = \Big\langle p\diamond q\,,\,\mathrm{ad}_\xi\eta \Big\rangle\,, \\
\text{[Definition of ad}^*] &= \Big\langle \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi}\,,\,\eta \Big\rangle\,.
\end{aligned}
$$

This is the Euler–Poincaré Equation (6.10). □

Exercise. Show that the diamond operation defined in Equation (6.14) is antisymmetric,

$$\Big\langle\, p \diamond q \,,\, \xi \,\Big\rangle = - \Big\langle\, q \diamond p \,,\, \xi \,\Big\rangle . \tag{6.19}$$

★

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

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

Compute the variations of the action integral

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

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

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

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

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

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

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

6.6 Lie–Poisson Hamiltonian formulation

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

$$l(p,q,\dot{q},\xi)=l(\xi)+\left\langle\!\!\left\langle\, p\,,\,\dot{q}+\pounds_\xi q\,\right\rangle\!\!\right\rangle$$

in the action (6.16) to the Hamiltonian,

$$H(p,q)=\left\langle\!\!\left\langle\, p\,,\,\dot{q}\,\right\rangle\!\!\right\rangle-l(p,q,\dot{q},\xi)=\left\langle\!\!\left\langle\, p\,,\,-\pounds_\xi q\,\right\rangle\!\!\right\rangle-l(\xi)\,,$$

whose variations are given by

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

These variational derivatives recover Equations (6.17) in canonical Hamiltonian form,

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

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

$$\frac{\delta l}{\delta \xi} = p \diamond q \,. \tag{6.20}$$

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

$$\frac{d\mu}{dt} = \{\mu, h\} = \mathrm{ad}^*_{\delta h/\delta \mu}\, \mu \,, \tag{6.21}$$

with

$$\mu = p \diamond q = \frac{\delta l}{\delta \xi} \,, \quad h(\mu) = \langle \mu, \xi \rangle - l(\xi) \,, \quad \xi = \frac{\delta h}{\delta \mu} \,. \tag{6.22}$$

By Equation (6.22), the evolution of a smooth real function $f : \mathfrak{g}^* \to \mathbb{R}$ is governed by

$$\begin{aligned}
\frac{df}{dt} &= \left\langle \frac{\delta f}{\delta \mu} \,,\, \frac{d\mu}{dt} \right\rangle \\
&= \left\langle \frac{\delta f}{\delta \mu} \,,\, \mathrm{ad}^*_{\delta h/\delta \mu}\, \mu \right\rangle \\
&= \left\langle \mathrm{ad}_{\delta h/\delta \mu} \frac{\delta f}{\delta \mu} \,,\, \mu \right\rangle \\
&= -\left\langle \mu \,,\, \left[\frac{\delta f}{\delta \mu} \,,\, \frac{\delta h}{\delta \mu} \right] \right\rangle \\
&=: \big\{ f, h \big\} \,.
\end{aligned} \tag{6.23}$$

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

7 Worked Example: Continuum spin chain equations

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

The Lie–Poisson Hamiltonian formulation of the Euler–Poincaré Equation (6.10) for this problem will be derived via the Legendre Transformation by following calculations similar to those done previously for the rigid body. To emphasise the systematic nature of the Legendre transformation from the Euler–Poincaré picture to the Lie–Poisson picture, we will lay out the procedure in well-defined steps.

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

$$S = \int_a^b\int_{-\infty}^{\infty}\ell(\Omega,\Xi)\,dx\,dt\,, \tag{7.1}$$

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

$$\Omega(t,x) = g^{-1}\partial_t g(t,x) \quad \text{and} \quad \Xi(t,x) = g^{-1}\partial_x g(t,x)\,, \tag{7.2}$$

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

$$\begin{aligned} \ell &= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{Tr}\Big(\big[g^{-1}\partial_t g, g^{-1}\partial_x g\big]^2\Big)dx \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{Tr}\Big(\big[\Omega,\,\Xi\big]^2\Big)dx\,, \end{aligned} \tag{7.3}$$

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 $\mathbb{R}^3$.

7.1 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,

$$\partial_t\Xi(t,x) - \partial_x\Omega(t,x) = \Xi\,\Omega - \Omega\,\Xi = [\Xi,\,\Omega] =: -\,\mathrm{ad}_\Omega\Xi\,. \tag{7.4}$$

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

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

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

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

$$
\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \Omega} = \mathrm{ad}^*_\Omega \frac{\delta \ell}{\delta \Omega} - \frac{\partial}{\partial x} \frac{\delta \ell}{\delta \Xi} + \mathrm{ad}^*_\Xi \frac{\delta \ell}{\delta \Xi} \,. \tag{7.5}
$$

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

Exercise. Use Equation (3.14) in Proposition 3.5 to show that the Euler–Poincaré Equation (7.5) is a ***conservation law*** for spin angular momentum $\Pi = \delta\ell/\delta\Omega$,

$$
\frac{\partial}{\partial t}\left(\mathrm{Ad}^*_{g(t,x)^{-1}} \frac{\delta l}{\delta \Omega} \right) = - \frac{\partial}{\partial x}\left(\mathrm{Ad}^*_{g(t,x)^{-1}} \frac{\delta l}{\delta \Xi} \right) . \tag{7.6}
$$

★

7.2 Hamiltonian formulation of the continuum spin chain

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

$$
h(\Pi, \Xi) = \Big\langle \Pi \,,\, \Omega \Big\rangle - \ell(\Omega, \Xi) \,. \tag{7.7}
$$

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

$$
\begin{aligned}
\partial_t \Pi &= \mathrm{ad}^*_{\delta h/\delta\Pi}\, \Pi + \partial_x \frac{\delta h}{\delta \Xi} - \mathrm{ad}^*_{\Xi}\, \frac{\delta h}{\delta \Xi} \,, \\
\partial_t \Xi &= \partial_x \frac{\delta h}{\delta \Pi} - \mathrm{ad}_{\delta h/\delta\Pi}\, \Xi \,.
\end{aligned}
\tag{7.8}
$$

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,

$$
\frac{\partial}{\partial t} \begin{bmatrix} \Pi \\ \Xi \end{bmatrix} = \begin{bmatrix} \mathrm{ad}^*_{\square}\Pi & (\mathrm{div} - \mathrm{ad}^*_{\Xi})\square \\ (\mathrm{grad} - \mathrm{ad}_{\square})\Xi & 0 \end{bmatrix} \begin{bmatrix} \delta h/\delta\Pi \\ \delta h/\delta\Xi \end{bmatrix}
\tag{7.9}
$$

The boxes $\square$ in Equation (9.45) indicate how the ad and ad* operations are applied in the matrix multiplication. For example,

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

so each matrix entry acts on its corresponding vector component.[8]

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 (9.45). 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 (9.45) 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,

$$\frac{\partial}{\partial t}\begin{bmatrix}\Pi_\alpha\\ \Xi^\alpha_i\end{bmatrix} = B_{\alpha\beta}\begin{bmatrix}\delta h/\delta\Pi_\beta\\ \delta h/\delta\Xi^\beta_j\end{bmatrix}, \tag{7.10}$$

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

$$B_{\alpha\beta} = \begin{bmatrix}-\Pi_\kappa\, t^\kappa_{\alpha\beta} & \delta^\beta_\alpha\partial_j + t^{\,\beta}_{\alpha\kappa}\Xi^\kappa_j\\ \delta^\alpha_\beta\partial_i - t^{\,\alpha}_{\beta\kappa}\Xi^\kappa_i & 0\end{bmatrix}. \tag{7.11}$$

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 (**??**), the Lie–Poisson Hamiltonian matrix in (7.11) may be rewritten for the $so(3)$ case in $\mathbb{R}^3$ vector form as

$$\frac{\partial}{\partial t}\begin{bmatrix}\boldsymbol{\Pi}\\ \boldsymbol{\Xi}_i\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Pi}\times & \partial_j + \boldsymbol{\Xi}_j\times\\ \partial_i + \boldsymbol{\Xi}_i\times & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta\boldsymbol{\Pi}\\ \delta h/\delta\boldsymbol{\Xi}_j\end{bmatrix}. \tag{7.12}$$

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

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

[8]This is the lower right corner of the Hamiltonian matrix for a perfect complex fluid [Ho2002, GBRa2008]. It also appears in the Lie–Poisson brackets for Yang–Mills fluids [GiHoKu1982] and for spin glasses [HoKu1988].

That the equations have the same form is to be expected because of the exchange symmetry under $t \leftrightarrow x$ and $\boldsymbol{\Omega} \leftrightarrow \boldsymbol{\Xi}$. Perhaps less expected is that the heavy-top form reappears.

For $G = SO(3)$ and the Lagrangian $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ in one spatial dimension $\ell(\boldsymbol{\Omega},\, \boldsymbol{\Xi})$ the Euler–Poincaré equation and its Hamiltonian form are given in terms of vector operations in $\mathbb{R}^3$, as follows. First, the Euler–Poincaré Equation (7.5) becomes

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \boldsymbol{\Omega}} = -\,\boldsymbol{\Omega} \times \frac{\delta \ell}{\delta \boldsymbol{\Omega}} - \frac{\partial}{\partial x}\frac{\delta \ell}{\delta \boldsymbol{\Xi}} - \,\boldsymbol{\Xi} \times \frac{\delta \ell}{\delta \boldsymbol{\Xi}}\,. \tag{7.13}$$

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 (7.13) 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***, 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 (7.5) becomes

$$\frac{\partial}{\partial t}\mathbb{A}\boldsymbol{\Omega} + \boldsymbol{\Omega}\times\mathbb{A}\boldsymbol{\Omega} + \frac{\partial}{\partial x}\mathbb{B}\boldsymbol{\Xi} + \boldsymbol{\Xi}\times\mathbb{B}\boldsymbol{\Xi} = 0\,. \tag{7.14}$$

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 (7.14) recovers the ***chiral model***, or ***sigma model***, which is completely integrable.

Hamiltonian structures The Hamiltonian structures of these equations on $so(3)^*$ are obtained from the Legendre-transform relations

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

Hence, the Euler–Poincaré Equation (7.5) becomes

$$\frac{\partial}{\partial t}\boldsymbol{\Pi} = \boldsymbol{\Pi}\times\frac{\delta h}{\delta\boldsymbol{\Pi}} + \frac{\partial}{\partial x}\frac{\delta h}{\delta\boldsymbol{\Xi}} + \boldsymbol{\Xi}\times\frac{\delta h}{\delta\boldsymbol{\Xi}}\,, \tag{7.15}$$

and the auxiliary Equation (7.16) becomes

$$\frac{\partial}{\partial t}\boldsymbol{\Xi} = \frac{\partial}{\partial x}\frac{\delta h}{\delta\boldsymbol{\Pi}} + \boldsymbol{\Xi}\times\frac{\delta h}{\delta\boldsymbol{\Pi}}\,, \tag{7.16}$$

which recovers the Lie–Poisson structure in Equation (7.12).

Finally, the reconstruction equations may be expressed using the hat map as

$$\begin{aligned}\partial_t O(t,x) &= O(t,x)\widehat{\Omega}(t,x)\quad\text{and}\\ \partial_x O(t,x) &= O(t,x)\widehat{\Xi}(t,x)\,.\end{aligned} \tag{7.17}$$

Remark

7.1. *The Euler–Poincaré equations for the continuum spin chain discussed here and their Lie–Poisson Hamiltonian formulation provide a framework for systematically investigating three-dimensional orientation dynamics along a one-dimensional strand. These partial differential equations are interesting in their own right and they have many possible applications. For an idea of where the applications of these equations could lead, consult [SiMaKr1988,EGHPR2010].*

Exercise. Write the Euler–Poincaré equations of the continuum spin chain for $SE(3)$, in which each point is both rotating and translating. Recall that

$$\left(\frac{d}{dt}\frac{\delta l}{\delta \xi}\,,\,\frac{d}{dt}\frac{\delta l}{\delta \alpha}\right) = \mathrm{ad}^*_{(\xi\,,\,\alpha)}\left(\frac{\delta l}{\delta \xi}\,,\,\frac{\delta l}{\delta \alpha}\right)\,. \tag{7.18}$$

Apply formula (7.18) to express the space-time Euler–Poincaré Equation (7.5) for $SE(3)$ in vector form.

Complete the computation of the Lie–Poisson Hamiltonian form for the continuum spin chain on $SE(3)$. ★

Exercise. Let the set of 2×2 matrices M_i with $i=1,2,3$ satisfy the defining relation for the symplectic Lie group $Sp(2)$,

$$M_iJM_i^T = J \quad \text{with} \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{7.19}$$

The corresponding elements of its Lie algebra $m_i = \dot{M}_iM_i^{-1} \in sp(2)$ satisfy $(Jm_i)^T = Jm_i$ for each $i=1,2,3$. Thus, $\mathsf{X}_i = Jm_i$ satisfying $\mathsf{X}_i^T = \mathsf{X}_i$ is a set of three symmetric 2×2 matrices. Define $\mathsf{X} = J\dot{M}M^{-1}$ with time derivative $\dot{M} = \partial M(t,x)/\partial t$ and $\mathsf{Y} = JM'M^{-1}$ with space derivative $M' = \partial M(t,x)/\partial x$. Then show that

$$\mathsf{X}' = \dot{\mathsf{Y}} + [\mathsf{X},\mathsf{Y}]_J\,, \tag{7.20}$$

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 (7.20) on $\mathfrak{sp}(2)$. ★

7.3 The $SO(3)$ G-Strand system in $\mathbb{R}^3$ vector form

By using the hat map, $\mathfrak{so}(3) \to \mathbb{R}^3$, the matrix G-Strand system for $SO(3)$ may be written in $\mathbb{R}^3$ vector form by following the analogy with the Euler rigid body [Ma1976] in standard notation,

$$\begin{aligned} \partial_t \Pi + \Omega \times \Pi - \partial_s \Xi - \Gamma \times \Xi = 0\,, \\ \partial_t \Gamma - \partial_s \Omega - \Gamma \times \Omega = 0\,, \end{aligned} \tag{7.21}$$

where $\Omega := O^{-1}\partial_t O \in \mathfrak{so}(3)$ and $\Pi := \partial\ell/\partial\Omega \in \mathfrak{so}(3)^*$ are the body angular velocity and momentum, while $\Gamma := O^{-1}\partial_s O \in \mathfrak{so}(3)$ and $\Xi = -\,\partial\ell/\partial\Gamma \in \mathfrak{so}(3)^*$ are the body angular strain and stress. These G-Strand equations for $\mathfrak{g} = \mathfrak{so}(3)$ equations may be expressed in Lie–Poisson Hamiltonian form in terms of vector operations in $\mathfrak{so}(3) \times \mathbb{R}^3$ as,

$$\frac{\partial}{\partial t}\begin{bmatrix}\Pi\\ \Gamma\end{bmatrix} = \begin{bmatrix}\Pi\times & \partial_s + \Gamma\times\\ \partial_s + \Gamma\times & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta\Pi = \Omega\\ \delta h/\delta\Gamma = \Xi\end{bmatrix}. \tag{7.22}$$

This Hamiltonian matrix yields a ***Lie–Poisson bracket*** defined on the dual of the semidirect-product Lie algebra $\mathfrak{so}(3)\circledS\mathbb{R}^3$ with a two-cocycle given by ∂_s. Namely,

$$\begin{aligned} \{f,\,h\} &= \int \begin{bmatrix}\delta f/\delta\Pi\,, & \delta f/\delta\Gamma\end{bmatrix} \cdot \begin{bmatrix}\Pi\times & \partial_s + \Gamma\times\\ \partial_s + \Gamma\times & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta\Pi\\ \delta h/\delta\Gamma\end{bmatrix} ds \\ &= \int -\Pi\cdot\frac{\delta f}{\delta\Pi}\times\frac{\delta h}{\delta\Pi} - \Gamma\cdot\left(\frac{\delta f}{\delta\Pi}\times\frac{\delta h}{\delta\Gamma} - \frac{\delta h}{\delta\Pi}\times\frac{\delta f}{\delta\Gamma}\right) + \frac{\delta f}{\delta\Pi}\partial_s\frac{\delta h}{\delta\Gamma} + \frac{\delta f}{\delta\Gamma}\partial_s\frac{\delta h}{\delta\Pi}\;ds. \end{aligned} \tag{7.23}$$

Dual variables are Π dual to $\mathfrak{so}(3)$ and Γ dual to $\mathbb{R}^3$. For more information about Lie–Poisson brackets, see [MaRa1999].

The $\mathbb{R}^3$ G-Strand equations (7.21) combine two classic ODEs due separately to Euler and Kirchhoff into a single PDE system. The $\mathbb{R}^3$ vector representation of $\mathfrak{so}(3)$ implies that $\mathrm{ad}^*_\Omega\Pi = -\,\Omega\times\Pi = -\mathrm{ad}_\Omega\Pi$, so the corresponding Euler–Poincaré equation has a ZCR. To find its integrability conditions, we set

$$L := \lambda^2 A + \lambda\Pi + \Gamma \quad \text{and} \quad M := \lambda^2 B + \lambda\Xi + \Omega\,, \tag{7.24}$$

and compute the conditions in terms of Π Ω, Ξ, Γ and the constant vectors A and B that are required to write the vector system (7.21) in zero-curvature form,

$$\partial_t L - \partial_s M - L \times M = 0\,. \tag{7.25}$$

By direct substitution of (7.24) into (7.25) and equating the coefficient of each power of λ to zero, one finds

$$
\begin{aligned}
&\lambda^4 : A \times B = 0\\
&\lambda^3 : A \times \Xi - B \times \Pi = 0\\
&\lambda^2 : A \times \Omega - B \times \Gamma + \Pi \times \Xi = 0\\
&\lambda^1 : \Pi \times \Omega + \Gamma \times \Xi = \partial_t \Pi - \partial_s \Xi \quad \text{(EP equation)}\\
&\lambda^0 : \Gamma \times \Omega = \partial_t \Gamma - \partial_s \Omega \quad \text{(compatibility)}
\end{aligned}
\tag{7.26}
$$

where A and B are taken as constant nonzero vectors. These imply the following relationships

$$
\begin{aligned}
&\lambda^4 : A = \alpha B\\
&\lambda^3 : A \times (\Xi - \Pi/\alpha) = 0 \Longrightarrow \Xi - \Pi/\alpha = \beta A\\
&\lambda^2 : A \times (\Omega - \Gamma/\alpha) = \Xi \times \Pi = \beta A \times \Pi
\end{aligned}
\tag{7.27}
$$

We solve equations (7.27) for the diagnostic variables Ξ and Ω, as

$$
\Xi = \frac{1}{\alpha}\Pi + \beta A \quad \text{and} \quad \Omega = \frac{\Gamma}{\alpha} + \beta\Pi + \gamma A\,, \tag{7.28}
$$

where α is real constant and β, γ at this point is arbitrary. Hence, we have proved,

Theorem

7.2. [ZCR formulation for $SO(3)$ G-Strand]
The $SO(3)$ matrix G-Strand system of equations in (7.21), may be expressed as a zero curvature representation (ZCR, or Lax pair),

$$
\partial_t L - \partial_s M - L \times M = 0\,, \quad \text{or equivalently} \quad \partial_t L - \partial_s M - [L, M] = 0\,, \tag{7.29}
$$

with $L, M \in \mathfrak{so}(3) \simeq \mathbb{R}^3$ given by

$$
L := \lambda^2 A + \lambda\Pi + \Gamma \quad \text{and} \quad M := \lambda^2 B + \lambda\Xi + \Omega\,, \tag{7.30}
$$

where A and B are constant nonzero vectors (diagonal matrices), and the diagnostic variables $\Xi, \Omega \in \mathfrak{so}(3) \simeq \mathbb{R}^3$, are defined in (7.28).

Equation (7.29) is the compatibility condition for a pair of *linear* equations with 3×3 antisymmetric matrix operators, L, M,

$$
\psi_s = L \times \psi \quad \text{and} \quad \psi_t = M \times \psi\,, \quad \text{or equivalently} \quad \psi_s = L\psi \quad \text{and} \quad \psi_t = M\psi
$$

Admission of a ZCR leading to its equivalent linear system is a key step in identifying completely integrable Hamiltonian PDEs.

8 EPDiff and Shallow Water Waves

Figure 3: This section is about using EPDiff to model unidirectional shallow water wave trains and their interactions in one dimension.

8.1 The Euler-Poincaré equation for EPDiff

Exercise. (Worked exercise: Deriving the Euler-Poincaré equation for EPDiff in one dimension)

The EPDiff(H^1) equation is written on the real line ($\mathbb{R}$) in terms of its velocity u and its momentum $m = \delta l/\delta u$ in one spatial dimension as

$$m_t + um_x + 2mu_x = 0\,, \quad \text{where} \quad m = u - u_{xx} \tag{8.1}$$

where subscripts denote partial derivatives in x and t, and we assume $u(x)$ vanishes as $|x| \to \infty$.

The EPDiff($\mathbb{R}$) equation (8.1) for the H^1 norm of the velocity u is obtained from the Euler-Poincaré reduction theorem for a right-invariant Lagrangian, when one defines the Lagrangian to be half the square of the H^1 norm $\|u\|_{H^1}$ of the vector field of velocity $u = \dot{g}g^{-1} \in \mathfrak{X}(\mathbb{R})$ on the real line $\mathbb{R}$ with $g \in \text{Diff}(\mathbb{R})$. Namely,

$$l(u) = \frac{1}{2}\|u\|^2_{H^1} = \frac{1}{2}\int_{-\infty}^{\infty} u^2 + u_x^2 \, dx\,.$$

(Recall that $u(x)$ vanishes as $|x| \to \infty$.)

(A) Derive the EPDiff equation on the real line in terms of its velocity u and its momentum $m = \delta l/\delta u = u - u_{xx}$ in one spatial dimension for this Lagrangian.
Hint: Prove a Lemma first, that $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \text{ad}_u\eta$ with $\eta = \delta g g^{-1}$.

(B) Use the Clebsch constrained Hamilton's principle

$$S(u,p,q) = \int l(u)\,dt + \sum_{a=1}^{N}\int p_a(t)\big(\dot{q}_a(t) - u(q(_at),t)\big)dt$$

to derive the peakon singular solution $m(x,t)$ of EPDiff as a momentum map in terms of canonically conjugate variables $q_a(t)$ and $p_a(t)$, with $a = 1, 2, \dots, N$.

★

Answer.

(A) ***Lemma***

The definition of velocity $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \mathrm{ad}_u\eta$ with $\eta = \delta g g^{-1}$.

Proof. Write $u = \dot{g}g^{-1}$ and $\eta = g'g^{-1}$ in natural notation and express the partial derivatives $\dot{g} = \partial g/\partial t$ and $g' = \partial g/\partial \epsilon$ using the right translations as

$$\dot{g} = u \circ g \quad \text{and} \quad g' = \eta \circ g\,.$$

By the chain rule, these definitions have mixed partial derivatives

$$\dot{g}' = u' = \nabla u \cdot \eta \quad \text{and} \quad \dot{g}' = \dot{\eta} = \nabla \eta \cdot u\,.$$

The difference of the mixed partial derivatives implies the desired formula,

$$u' - \dot{\eta} = \nabla u \cdot \eta - \nabla \eta \cdot u = -\,[u, \eta] =: -\mathrm{ad}_u\eta\,,$$

so that

$$u' = \dot{\eta} - \mathrm{ad}_u\eta\,.$$

In 3D, this becomes

$$\delta \mathbf{u} = \dot{\mathbf{v}} - \mathrm{ad}_{\mathbf{u}}\mathbf{v}\,. \tag{8.2}$$

This formula may be rederived as follows. We write $\mathbf{u} = \dot{g}g^{-1}$ and $\mathbf{v} = g'g^{-1}$ in natural notation and express the partial derivatives $\dot{g} = \partial g/\partial t$ and $g' = \partial g/\partial \epsilon$ using the right translations as

$$\dot{g} = \mathbf{u} \circ g \quad \text{and} \quad g' = \mathbf{v} \circ g\,.$$

To compute the mixed partials, consider the chain rule for say $\mathbf{u}(g(t,\epsilon)\mathbf{x}_0)$ and set $\mathbf{x}(t,\epsilon) = g(t,\epsilon)\cdot\mathbf{x}_0$. Then,

$$\mathbf{u}' = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \epsilon} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \cdot g'(t,\epsilon)\mathbf{x}_0 = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \cdot g'g^{-1}\mathbf{x} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \cdot \mathbf{v}(\mathbf{x})\,.$$

The chain rule for $\dot{\mathbf{v}}$ gives a similar formula with $\mathbf{u}$ and $\mathbf{v}$ exchanged. Thus, the chain rule gives two expressions for the mixed partial derivative $\dot{g}'$ as

$$\dot{g}' = \mathbf{u}' = \nabla \mathbf{u} \cdot \mathbf{v} \quad \text{and} \quad \dot{g}' = \dot{\mathbf{v}} = \nabla \mathbf{v} \cdot \mathbf{u}\,.$$

The difference of the mixed partial derivatives then implies the desired formula (8.2), since

$$\mathbf{u}' - \dot{\mathbf{v}} = \nabla \mathbf{u} \cdot \mathbf{v} - \nabla \mathbf{v} \cdot \mathbf{u} = -\,[\mathbf{u}, \mathbf{v}] = -\,\mathrm{ad}_{\mathbf{u}}\mathbf{v}\,.$$

$\square$

The EPDiff(H^1) equation on $\mathbb{R}$. The EPDiff(H^1) equation is written on the real line in terms of its velocity u and its momentum $m = \delta l/\delta u$ in one spatial dimension as

$$m_t + u m_x + 2 m u_x = 0\,, \quad \text{where} \quad m = u - u_{xx} \tag{8.3}$$

where subscripts denote partial derivatives in x and t.

Proof. This equation is derived from the variational principle with $l(u) = \frac{1}{2}\|u\|^2_{H^1}$ as follows.

$$\begin{aligned}
0 = \delta S &= \delta \int l(u) dt = \frac{1}{2}\,\delta \iint u^2 + u_x^2\, dx\, dt \\
&= \iint (u - u_{xx})\, \delta u\, dx\, dt =: \iint m\, \delta u\, dx\, dt \\
&= \iint m\, (\eta_t - \mathrm{ad}_u \eta)\, dx\, dt \\
&= \iint m\, (\eta_t + u\eta_x - \eta u_x)\, dx\, dt \\
&= -\iint (m_t + (um)_x + m u_x)\, \eta\, dx\, dt \\
&= -\iint (m_t + \mathrm{ad}^*_u m)\, \eta\, dx\, dt\,,
\end{aligned}$$

where $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \mathrm{ad}_u \eta$ with $\eta = \delta g g^{-1}$. □

Exercise. Follow this proof in 3D and write out the resulting Euler-Poincaré equation. ★

(B) The constrained Clebsch action integral is given as

$$S(u, p, q) = \int l(u)\, dt + \sum_{a=1}^{N} \int p_a(t)\big(\dot{q}_a(t) - u(q_a(t), t)\big) dt$$

whose variation in u is gotten by inserting a delta function, so that

$$\begin{aligned} 0 = \delta S \;&=\; \int \left(\frac{\delta l}{\delta u} - \sum_{a=1}^{N} p_a \delta(x - q_a(t)) \right) \delta u \, dx \, dt \\ &\quad - \int \left(\dot{p}_a(t) + \frac{\partial u}{\partial q_a} \, p_a(t) \right) \delta q_a - \delta p_a \Big(\dot{q}_a(t) - u(q_a(t), t) \Big) dt \,. \end{aligned}$$

The singular momentum solution $m(x,t)$ of EPDiff(H^1) is written as the cotangent-lift momentum map

$$m(x,t) = \delta l / \delta u = \sum_{a=1}^{N} p_a(t) \delta(x - q_a(t)) \tag{8.4}$$

Inserting this solution into the Legendre transform

$$h(m) = \langle \, m, u \, \rangle - l(u)$$

yields the conserved energy

$$e = \frac{1}{2} \int m(x,t) u(x,t) \, dx = \frac{1}{2} \sum_{a=1}^{N} \int p_a(t) \delta(x - q_a(t)) u(x,t) \, dx = \frac{1}{2} \sum_{a=1}^{N} p_a(t) u(q_a(t), t) \tag{8.5}$$

Consequently, the variables (q_a, p_a) satisfy equations,

$$\dot{q}_a(t) = u(q_a(t), t) \,, \qquad \dot{p}_a(t) = - \frac{\partial u}{\partial q_a} \, p_a(t) \,, \tag{8.6}$$

with the pulse-train solution for velocity

$$u(q_a, t) = \sum_{b=1}^{N} p_b K(q_a, q_b) = \frac{1}{2} \sum_{b=1}^{N} p_b e^{-|q_a - q_b|} \tag{8.7}$$

where $K(x,y) = \frac{1}{2} e^{-|x-y|}$ is the Green's function kernel for the Helmholtz operator $1 - \partial_x^2$. Each pulse in the pulse-train solution for velocity (8.7) has a sharp peak. For that reason, these pulses are called *peakons*. In fact, equations (8.6) are Hamilton's canonical equations with Hamiltonian obtained from equations (8.5) for energy and (8.7) for velocity, as given in [CaHo1993],

$$H_N = \frac{1}{2} \sum_{a,b=1}^{N} p_a \, p_b \, K(q_a, q_b) = \frac{1}{4} \sum_{a,b=1}^{N} p_a \, p_b \, \mathrm{e}^{-|q_a - q_b|} \,. \tag{8.8}$$

The first canonical equation in eqn (8.6) implies that the peaks at the positions $x = q^a(t)$ in the peakon-train solution (8.7) move with the flow of the fluid velocity u at those positions, since $u(q^a(t), t) = \dot{q}^a(t)$. This means the positions $q^a(t)$ are ***Lagrangian coordinates*** frozen into the flow of EPDiff. Thus, the singular solution obtained from the cotangent-lift momentum map (8.4) is the map from Lagrangian coordinates to Eulerian coordinates (that is, the ***Lagrange-to-Euler map***) for the momentum.

▲

Remark

8.1 (Solution behaviour of EPDiff(H^1)). *The peakon-train solutions of EPDiff are an* ***emergent phenomenon****. A wave train of peakons emerges in solving the initial-value problem for the EPDiff equation (8.3) for essentially any spatially confined initial condition. A numerical simulation of the solution behaviour for EPDiff(H^1) given in Figure 4 shows the emergence of a wave train of peakons from a Gaussian initial condition.*

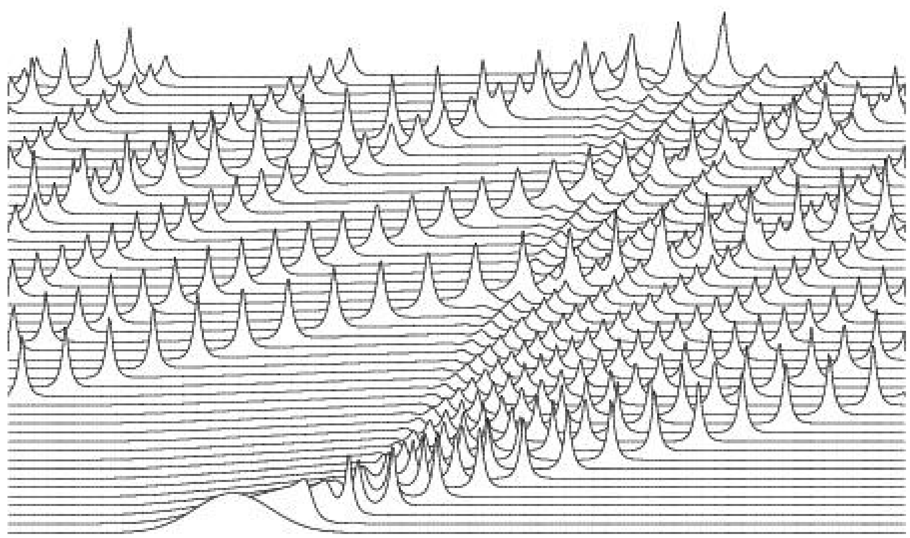


Figure 4: Under the evolution of the EPDiff(H^1) equation (8.3), an ordered ***wave train of peakons*** emerges from a smooth localized initial condition (a Gaussian). The spatial profiles at successive times are offset in the vertical to show the evolution. The peakon wave train eventually wraps around the periodic domain, thereby allowing the leading peakons to overtake the slower peakons from behind in collisions that conserve momentum and preserve the peakon shape but cause phase shifts in the positions of the peaks, as discussed in [CaHo1993].

Exercise. Compute the Lie–Poisson Hamiltonian form of the EPDiff equation (8.3). ★

Answer.

Lie-Poisson Hamiltonian form of EPDiff. In terms of m, the conserved energy Hamiltonian for the EPDiff equation (8.3) is obtained by Legendre transforming the kinetic-energy Lagrangian $l(u)$, as

$$h(m) = \Big\langle m\,, u \Big\rangle - l(u)\,.$$

Thus, the Hamiltonian depends on m, as

$$h(m) = \frac{1}{2}\int m(x)K(x-y)m(y)\,dxdy\,,$$

which also reveals the geodesic nature of the EPDiff equation (8.3) and the role of $K(x,y)$ in the kinetic energy metric on the Hamiltonian side.

The corresponding ***Lie-Poisson bracket*** for EPDiff as a Hamiltonian evolution equation is given by,

$$\partial_t m = \big\{ m, h \big\} = -\mathrm{ad}^*_{\delta h/\delta m} m = -\left(\partial_x m + m\partial_x\right)\frac{\delta h}{\delta m} \quad \text{and} \quad \frac{\delta h}{\delta m} = u\,,$$

which recovers the starting equation and indicates some of its connections with fluid equations on the Hamiltonian side. For any two smooth functionals f, h of m in the space for which the solutions of EPDiff exist, this Lie-Poisson bracket may be expressed as,

$$\big\{ f, h \big\} = -\int \frac{\delta f}{\delta m}\left(\partial_x m + m\partial_x\right)\frac{\delta h}{\delta m} dx = -\int m\left[\frac{\delta f}{\delta m}\,, \frac{\delta h}{\delta m}\right] dx$$

where $[\cdot\,,\,\cdot]$ denotes the Lie algebra bracket of vector fields. That is,

$$\left[\frac{\delta f}{\delta m}\,, \frac{\delta h}{\delta m}\right] = \frac{\delta f}{\delta m}\partial_x\frac{\delta h}{\delta m} - \frac{\delta h}{\delta m}\partial_x\frac{\delta f}{\delta m}\,.$$

▲

Exercise. What is the Casimir for this Lie Poisson bracket? What does it mean from the viewpoint of coadjoint orbits? What is the 3D version of this Lie Poisson bracket? Does it have a Casimir in 3D? ★

8.2 The CH equation is bi-Hamiltonian

The completely integrable CH equation for unidirectional shallow water waves first derived in [CaHo1993],

$$m_t + um_x + 2mu_x = \underbrace{-c_0 u_x + \gamma u_{xxx}}_{\text{Linear Dispersion}}\,, \qquad m = u - \alpha^2 u_{xx}\,, \qquad u = K * m \quad \text{with} \quad K(x,y) = \frac{1}{2}e^{-|x-y|}\,. \tag{8.9}$$

This equation describes shallow water dynamics as completely integrable soliton motion at quadratic order in the asymptotic expansion for unidirectional shallow water waves on a free surface under gravity.

The term bi-Hamiltonian means the equation may be written in two compatible Hamiltonian forms, namely as

$$m_t = -B_2 \frac{\delta H_1}{\delta m} = -B_1 \frac{\delta H_2}{\delta m} \tag{8.10}$$

with

$$\begin{aligned} H_1 &= \frac{1}{2}\int (u^2 + \alpha^2 u_x^2)\, dx\,, \quad \text{and} \quad B_2 = \partial_x m + m\partial_x + c_0\partial_x + \gamma\,\partial_x^3 \\ H_2 &= \frac{1}{2}\int u^3 + \alpha^2 u\, u_x^2 + c_0 u^2 - \gamma\, u_x^2\, dx\,, \quad \text{and} \quad B_1 = \partial_x - \alpha^2\partial_x^3\,. \end{aligned} \tag{8.11}$$

These bi-Hamiltonian forms restrict properly to those for KdV when $\alpha^2 \to 0$, and to those for EPDiff when $c_0, \gamma \to 0$. Compatibility of B_1 and B_2 is assured, because $(\partial_x m + m\partial_x)$, ∂_x and ∂_x^3 are all mutually compatible Hamiltonian operators. That is, any linear combination of these operators defines a Poisson bracket,

$$\{f, h\}(m) = -\int \frac{\delta f}{\delta m}(c_1 B_1 + c_2 B_2)\frac{\delta h}{\delta m}\, dx\,, \tag{8.12}$$

as a bilinear skew-symmetric operation which satisfies the Jacobi identity. Moreover, no further deformations of these Hamiltonian operators involving higher order partial derivatives would be compatible with B_2, as shown in [Ol2000]. This was already known in the literature for KdV, whose bi-Hamilton structure has $B_1 = \partial_x$ and B_2 the same as CH.

8.3 Magri's theorem

As we shall see, because equation (8.9) is bi-Hamiltonian, it has an infinite number of conservation laws. These laws can be constructed by defining the transpose operator $\mathcal{R}^T = B_1^{-1}B_2$ that leads from the variational derivative of one conservation law to the next, according to

$$\frac{\delta H_n}{\delta m} = \mathcal{R}^T \frac{\delta H_{n-1}}{\delta m}\,, \quad n = -1, 0, 1, 2, \dots . \tag{8.13}$$

The operator $\mathcal{R}^T = B_1^{-1}B_2$ recursively takes the variational derivative of H_{-1} to that of H_0, to that of H_1, to then that of H_2. The next steps are not so easy for the integrable CH hierarchy, because each application of the recursion operator introduces an additional convolution integral into the sequence. Correspondingly, the recursion operator $\mathcal{R} = B_2B_1^{-1}$ leads to a hierarchy of commuting flows, defined by $K_{n+1} = \mathcal{R}K_n$, for $n = 0, 1, 2, \dots$,

$$m_t^{(n+1)} = K_{n+1}[m] = -B_1\frac{\delta H_n}{\delta m} = -\,B_2\frac{\delta H_{n-1}}{\delta m} = B_2B_1^{-1}K_n[m]\,. \tag{8.14}$$

The first three flows in the "positive hierarchy" when $c_0, \gamma \to 0$ are

$$m_t^{(1)} = 0\,, \quad m_t^{(2)} = -\,m_x\,, \quad m_t^{(3)} = -\,(m\partial + \partial m)u\,, \tag{8.15}$$

the third being EPDiff. The next flow is too complicated to be usefully written here. However, by construction, all of these flows commute with the other flows in the hierarchy, so they each conserve H_n for $n = 0, 1, 2, \dots$.

The recursion operator can also be continued for negative values of n. The conservation laws generated this way do not introduce convolutions, but care must be taken to ensure the conserved densities are integrable. All the Hamiltonian densities in the negative hierarchy are expressible in terms of m only and do not involve u. Thus, for instance, the first few Hamiltonians in the negative hierarchy of EPDiff are given by

$$H_0 = \int_{-\infty}^{\infty} m\,dx\,, \qquad H_{-1} = \int_{-\infty}^{\infty} \sqrt{m}\,dx\,, \tag{8.16}$$

and

$$H_{-2} = \frac{1}{2}\int_{-\infty}^{\infty}\left[\frac{\alpha^2}{4}\frac{m_x^2}{m^{5/2}} - \frac{2}{\sqrt{m}}\right]. \tag{8.17}$$

The flow defined by (8.14) for these is thus,

$$m_t^{(0)} = -B_1\frac{\delta H_{-1}}{\delta m} = -\,B_2\frac{\delta H_{-2}}{\delta m} = -(\partial - \alpha^2\partial^3)\left(\frac{1}{2\sqrt{m}}\right). \tag{8.18}$$

This flow is similar to the Dym equation,

$$u_{xxt} = \partial^3\left(\frac{1}{2\sqrt{u_{xx}}}\right). \tag{8.19}$$

8.4 Equation (8.9) is isospectral

The isospectral eigenvalue problem associated with equation (8.9) may be found by using the recursion relation of the bi-Hamiltonian structure, following the standard technique of [GeDo1979]. Let us introduce a spectral parameter λ and multiply by λ^n the $n-$th step of the recursion relation (8.14), then summing yields

$$B_1 \sum_{n=0}^{\infty} \lambda^n \frac{\delta H_n}{\delta m} = \lambda B_2 \sum_{n=0}^{\infty} \lambda^{(n-1)} \frac{\delta H_{n-1}}{\delta m}\,, \tag{8.20}$$

or, by introducing

$$\psi^2(x,t;\lambda) := \sum_{n=-1}^{\infty} \lambda^n \frac{\delta H_n}{\delta m}\,, \tag{8.21}$$

one finds, formally,

$$B_1 \psi^2(x,t;\lambda) = \lambda B_2 \psi^2(x,t;\lambda)\,. \tag{8.22}$$

This is a *third order* eigenvalue problem for the squared-eigenfunction ψ^2, which turns out to be equivalent to a *second order* Sturm-Liouville problem for ψ. It is straightforward to show that if ψ satisfies

$$\lambda \left(\frac{1}{4} - \alpha^2 \partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi\,, \tag{8.23}$$

then ψ^2 is a solution of (8.22). Now, assuming λ will be independent of time, we seek, in analogy with the KdV equation, an evolution equation for ψ of the form,

$$\psi_t = a\psi_x + b\psi\,, \tag{8.24}$$

where a and b are functions of u and its derivatives to be determined by the requirement that the compatibility condition $\psi_{xxt} = \psi_{txx}$ between (8.23) and (8.24) implies (8.9). Cross differentiation shows

$$b = -\frac{1}{2}a_x\,, \quad \text{and} \quad a = -\,(\lambda + u)\,. \tag{8.25}$$

Consequently,

$$\psi_t = -\,(\lambda + u)\psi_x + \frac{1}{2}u_x\psi\,, \tag{8.26}$$

is the desired evolution equation for ψ.

Summary of the isospectral property of equation (8.9) Thus, according to the standard Gelfand-Dorfman theory of [GeDo1979] for obtaining the isospectral problem for equation via the squared-eigenfunction approach, its bi-Hamiltonian property implies that the nonlinear shallow water wave equation (8.9) arises as a compatibility condition for two linear equations. These are the *isospectral eigenvalue problem*,

$$\lambda\left(\frac{1}{4}-\alpha^2\partial_x^2\right)\psi=\left(\frac{c_0}{4}+\frac{m(x,t)}{2}+\gamma\,\partial_x^2\right)\psi\,, \tag{8.27}$$

and the *evolution equation* for the eigenfunction ψ,

$$\psi_t=-(u+\lambda)\,\psi_x+\frac{1}{2}u_x\,\psi\,.$$

Compatibility of these linear equations ($\psi_{xxt}=\psi_{txx}$) together with isospectrality

$$d\lambda/dt=0\,,$$

imply equation (8.9). Consequently, the nonlinear water wave equation (8.9) admits the IST method for the solution of its initial value problem, just as the KdV equation does. In fact, the isospectral problem for equation (8.9) restricts to the isospectral problem for KdV (i.e., the Schrödinger equation) when $\alpha^2\to 0$.

Dispersionless case In the dispersionless case $c_0=0=\gamma$, the shallow water equation (8.9) becomes the 1D geodesic equation EPDiff(H^1) in (8.3)

$$m_t+um_x+2mu_x=0\,,\qquad m=u-\alpha^2u_{xx}\,, \tag{8.28}$$

and the spectrum of its eigenvalue problem (8.27) becomes *purely disctrete*. The traveling wave solutions of 1D EPDiff (8.28) in this dispersionless case are the "peakons," described by the reduced, or collective, solutions (8.6) for EPDiff equation (8.3) with traveling waves

$$u(x,t)=c\,K(x-ct)=c\,e^{-|x-ct|/\alpha}\,.$$

In this case, the EPDiff equation (8.3) may also be written as a conservation law for momentum,

$$\partial_t m=-\partial_x\Big(um+\frac{1}{2}u^2-\frac{\alpha^2}{2}u_x^2\Big)\,. \tag{8.29}$$

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 (8.27) in the dispersionless case $c_0=0=\Gamma$ has purely discrete spectrum on the real line and the $N-$soliton solutions for this equation have the peakon form,

$$u(x,t)=\sum_{i=1}^{N}p_i(t)e^{-|x-q_i(t)|/\alpha}\,. \tag{8.30}$$

Here $p_i(t)$ and $q_i(t)$ satisfy the finite dimensional geodesic motion equations obtained as canonical Hamiltonian equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}\,, \tag{8.31}$$

when the Hamiltonian is given by,

$$H = \frac{1}{2}\sum_{i,j=1}^{N} p_i\, p_j\, e^{-|q_i - q_j|/\alpha}\,. \tag{8.32}$$

Thus, the CH peakons turn out to be an integrable subcase of the pulsons.

Integrability of the $N-$peakon dynamics One may verify integrability of the $N-$peakon dynamics by substituting the $N-$peakon solution (8.30) (which produces the sum of delta functions in (8.4) for the momentum map m) into the isospectral problem (8.27). This substitution reduces (8.27) to an $N \times N$ matrix eigenvalue problem.

In fact, the canonical equations (8.31) for the peakon Hamiltonian (8.32) may be written directly in Lax matrix form,

$$\frac{dL}{dt} = [L, A] \quad \Longleftrightarrow \quad L(t) = U(t)L(0)U^{\dagger}(t)\,, \tag{8.33}$$

with $A = \dot{U}U^{\dagger}(t)$ and $UU^{\dagger} = Id$. Explicitly, L and A are $N \times N$ matrices with entries

$$L_{jk} = \sqrt{p_j p_k}\,\phi(q_j - q_k)\,, \quad A_{jk} = -2\sqrt{p_j p_k}\,\phi'(q_j - q_k)\,. \tag{8.34}$$

Here $\phi'(x)$ denotes derivative with respect to the argument of the function ϕ, given by $\phi(x) = e^{-|x|/2\alpha}$. The Lax matrix L in (8.34) 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 independent, nontrivial and in involution. Hence, the canonically Hamiltonian $N-$peakon dynamics (8.31) is integrable.

Exercise. Show that the peakon Hamiltonian H_N in (8.32) is expressed as a function of the invariants of the matrix L, as

$$H_N = -\mathrm{tr}\, L^2 + 2(\mathrm{tr}\, L)^2\,. \tag{8.35}$$

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 no difficulties arise, either due to the nonanalyticity of $\phi(x)$, or the sign in the square-roots in the Lax matrices L and A. ★

8.5 Steepening Lemma and peakon formation

We now address the mechanism for the formation of the peakons, by showing that initial conditions exist for which the solution of the EPDiff(H^1) equation,

$$\partial_t m + u m_x + 2 u_x m = 0 \quad \text{with} \quad m = u - \alpha^2 u_{xx}\,, \tag{8.36}$$

can develop a vertical slope in its velocity $u(t, x)$, in finite time. The mechanism turns out to be associated with inflection points of negative slope, such as occur on the leading edge of a rightward propagating velocity profile. In particular, we have the following steepening lemma.

Lemma

8.2 (Steepening Lemma).
Suppose the initial profile of velocity $u(0,x)$ has an inflection point at $x=\overline{x}$ to the right of its maximum, and otherwise it decays to zero in each direction sufficiently rapidly for the Hamiltonian H_1 in equation (8.11) to be finite. Then the negative slope at the inflection point will become vertical in finite time.

Proof. Consider the evolution of the slope at the inflection point. Define $s=u_x(\overline{x}(t),t)$. Then the EPDiff(H^1) equation (8.36), rewritten as,

$$(1-\alpha^2\partial^2)(u_t+uu_x)=-\,\partial\Big(u^2+\frac{\alpha^2}{2}u_x^2\Big)\,, \tag{8.37}$$

yields an equation for the evolution of s. Namely, using $u_{xx}(\overline{x}(t),t)=0$ leads to

$$\frac{ds}{dt}=-\frac{1}{2}s^2+\frac{1}{2}\int_{-\infty}^{\infty}\mathrm{sgn}(\overline{x}-y)e^{-|\overline{x}-y|}\partial_y\left(u^2+\frac{1}{2}u_y^2\right)dy\,. \tag{8.38}$$

Integrating by parts and using the inequality $a^2+b^2\geq 2ab$, for any two real numbers a and b, leads to

$$\begin{aligned}\frac{ds}{dt} &= -\frac{1}{2}s^2-\frac{1}{2}\int_{-\infty}^{\infty}e^{-|\overline{x}-y|}\left(u^2+\frac{1}{2}u_y^2\right)dy+u^2(\overline{x}(t),t)\\ &\leq -\frac{1}{2}s^2+2u^2(\overline{x}(t),t)\,.\end{aligned} \tag{8.39}$$

Then, provided $u^2(\overline{x}(t),t)$ remains finite, say less than a number $M/4$, we have

$$\frac{ds}{dt}=-\frac{1}{2}s^2+\frac{M}{2}\,, \tag{8.40}$$

which implies, for negative slope initially $s\leq-\sqrt{M}$, that

$$s\leq\sqrt{M}\coth\left(\sigma+\frac{t}{2}\sqrt{M}\right)\,, \tag{8.41}$$

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 (8.40) exists is verified in general by a Sobolev inequality. In fact, $M=8H_1$, since

$$\max_{x\in\mathbb{R}}u^2(x,t)\leq\int_{-\infty}^{\infty}(u^2+u_x^2)\,dx=2H_1=const\,. \tag{8.42}$$

$\square$

Remark

8.3. If the initial condition is antisymmetric, then the inflection point at $u = 0$ is fixed and $d\overline{x}/dt = 0$, due to the symmetry $(u, x) \to (-u, -x)$ admitted by equation (8.9). In this case, $M = 0$ and no matter how small $|s(0)|$ (with $s(0) < 0$) verticality $s \to -\infty$ develops at $\overline{x}$ in finite time.

The steepening lemma indicates that traveling wave solutions of EPDiff(H^1) in (8.36) must not have the usual sech^2 shape, since inflection points with sufficiently negative slope can lead to unsteady changes in the shape of the profile if inflection points are present. In fact, numerical simulations show that the presence of an inflection point in any confined initial velocity distribution is the ***mechanism*** for the formation of the peakons. Namely. the initial (positive) velocity profile "leans" to the right and steepens, then produces a peakon which is taller than the initial profile, so it propagates away to the right. This leaves a profile behind with an inflection point of negative slope; so the process repeats, thereby producing a train of peakons with the tallest and fastest ones moving rightward in order of height. This discrete process of peakon creation corresponds to the discreteness of the isospectrum for the eigenvalue problem (8.27) 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.

9 The Euler-Poincaré framework: fluid dynamics à la [HoMaRa1998a]

9.1 Left and right momentum maps for CH

As mentioned earlier, the singular momentum solution in equation (8.4) $m(x,t)$ of EPDiff(H^1) is a cotangent-lift momentum map

$$m(x,t) = \delta l/\delta u = \sum_{a=1}^{N} p_a(t)\delta(x - q_a(t)) \tag{9.1}$$

Here, we will consider the n-dimensional version of the momentum map for solutions of the $n-$dimensional EPDiff equation

$$\frac{\partial}{\partial t}\mathbf{m} \;+\; \underbrace{\mathbf{u}\cdot\nabla\mathbf{m}}_{\text{convection}} \;+\; \underbrace{\nabla\mathbf{u}^T\cdot\mathbf{m}}_{\text{stretching}} \;+\; \underbrace{\mathbf{m}\,(\operatorname{div}\mathbf{u})}_{\text{expansion}} = 0\,. \tag{9.2}$$

In coordinates x^i, $i = 1, 2, \dots, n$, using the summation convention, and writing $\mathbf{m} = m_i dx^i \otimes d^n x$ (regarding $\mathbf{m}$ as a one-form density) and $\mathbf{u} = u^i\partial/\partial x^i$ (regarding $\mathbf{u}$ as a vector field), the EPDiff equations read

$$\frac{\partial}{\partial t}m_i + u^j\frac{\partial m_i}{\partial x^j} + m_j\frac{\partial u^j}{\partial x^i} + m_i\frac{\partial u^j}{\partial x^j} = 0\,. \tag{9.3}$$

The EPDiff equations can also be written concisely as

$$\frac{\partial m}{\partial t} + \pounds_u m = 0, \tag{9.4}$$

where $\pounds_u m$ denotes the Lie derivative of the momentum one form density $m = \mathbf{m}\cdot d\mathbf{x}\otimes d^n x$ with respect to the velocity vector field $u = \mathbf{u}\cdot\nabla$. The n-dimensional version of the momentum map (9.1) is given by

$$\mathbf{m}(\mathbf{x},t) = \sum_{a=1}^{N}\int \mathbf{P}^a(s,t)\,\delta\big(\,\mathbf{x} - \mathbf{Q}^a(s,t)\,\big)ds. \tag{9.5}$$

These solutions are vector-valued functions supported in $\mathbb{R}^n$ on a set of N surfaces (or curves) of codimension $(n-k)$ for $s\in\mathbb{R}^k$ with $k<n$. They may, for example, be supported on sets of points (vector peakons, $k=0$), one-dimensional filaments (strings, $k=1$), or two-dimensional surfaces (sheets, $k=2$) in three dimensions. *One of the main results of geometric mechhanics is the theorem stating that the singular solution ansatz (9.5) is an equivariant momentum map. This result has helped to organize the theory and and its study has suggested many new avenues of exploration.*

Substitution of the solution ansatz (9.5) into the EPDiff equations (9.2) implies the following integro-partial-differential equations (IPDEs) for the evolution of such strings and sheets,

$$
\begin{aligned}
\frac{\partial}{\partial t}\mathbf{Q}^a(s,t) &= \sum_{b=1}^{N}\int \mathbf{P}^b(s',t)\, G(\mathbf{Q}^a(s,t)-\mathbf{Q}^b(s',t)\,)\,ds'\,, \\
\frac{\partial}{\partial t}\mathbf{P}^a(s,t) &= -\sum_{b=1}^{N}\int \big(\mathbf{P}^a(s,t)\cdot\mathbf{P}^b(s',t)\big)\,\frac{\partial}{\partial \mathbf{Q}^a(s,t)}G\big(\mathbf{Q}^a(s,t)-\mathbf{Q}^b(s',t)\big)\,ds'\,.
\end{aligned}
\tag{9.6}
$$

Importantly for the interpretation of these solutions, the coordinates $s\in\mathbb{R}^k$ turn out to be Lagrangian coordinates. The velocity field corresponding to the momentum solution ansatz (9.5) is given by

$$
\mathbf{u}(\mathbf{x},t) = G*\mathbf{m} = \sum_{b=1}^{N}\int \mathbf{P}^b(s',t)\,G\big(\,\mathbf{x}-\mathbf{Q}^b(s',t)\,\big)ds'\,,\quad \mathbf{u}\in\mathbb{R}^n\,. \tag{9.7}
$$

When evaluated along the curve $\mathbf{x}=\mathbf{Q}^a(s,t)$, the velocity satisfies,

$$
\mathbf{u}(\mathbf{x},t)\Big|_{\mathbf{x}=\mathbf{Q}^a(s,t)} = \sum_{b=1}^{N}\int \mathbf{P}^b(s',t)\,G\big(\,\mathbf{Q}^a(s,t)-\mathbf{Q}^b(s',t)\,\big)ds' = \frac{\partial \mathbf{Q}^a(s,t)}{\partial t}\,. \tag{9.8}
$$

Thus, the lower-dimensional support sets defined on $\mathbf{x}=\mathbf{Q}^a(s,t)$ and parameterized by coordinates $s\in\mathbb{R}^k$ move with the fluid velocity. Moreover, equations (9.6) for the evolution of these support sets are canonical Hamiltonian equations,

$$
\frac{\partial}{\partial t}\mathbf{Q}^a(s,t) = \frac{\delta H_N}{\delta \mathbf{P}^a}\,,\qquad \frac{\partial}{\partial t}\mathbf{P}^a(s,t) = -\,\frac{\delta H_N}{\delta \mathbf{Q}^a}\,. \tag{9.9}
$$

The Hamiltonian function $H_N:(\mathbb{R}^n\times\mathbb{R}^n)^N\to\mathbb{R}$ is,

$$
H_N = \frac{1}{2}\iint\sum_{a\,,\,b=1}^{N}\big(\mathbf{P}^a(s,t)\cdot\mathbf{P}^b(s',t)\big)\,G\big(\mathbf{Q}^a(s,t)-\mathbf{Q}^b(s',t)\big)\,ds\,ds'\,. \tag{9.10}
$$

This is the Hamiltonian for geodesic motion on the cotangent bundle of a set of curves $\mathbf{Q}^a(s,t)$ with respect to the metric given by G. This dynamics was investigated numerically in [?, ?] to which we refer for more details of the solution properties.

The map that implements the canonical $(\mathbf{Q},\mathbf{P})$ variables in terms of singular solutions is a (cotangent bundle) momentum map. Such momentum maps are Poisson maps; so the canonical Hamiltonian nature of the dynamical equations for $(\mathbf{Q},\mathbf{P})$ fits into a general theory which also provides a framework for suggesting other avenues of investigation.

9.1.1 The momentum map for left action

Theorem

9.1. *The momentum ansatz* (9.5) *for measure-valued solutions of the* EPDiff *equation* (9.2)*, defines an equivariant momentum map*

$$\mathbf{J}_{\text{Sing}} : T^* \operatorname{Emb}(S, \mathbb{R}^n) \to \mathfrak{X}(\mathbb{R}^n)^*$$

*that we will call the **singular solution momentum map**.*

We shall explain the notation used in this statement in the course of the proof. Right away, however, we note that the sense of "defines" is quite simple, namely expressing $\mathbf{m}$ in terms of $\mathbf{Q}, \mathbf{P}$ (which are, in turn, functions of s) can be regarded as a map from the space of $(\mathbf{Q}(s), \mathbf{P}(s))$ to the space of $\mathbf{m}$'s.

Proof. The strategy of the proof is, as follows. Consider the group $\mathfrak{G} = \operatorname{Diff}$ of diffeomorphisms of the space M in which the EPDiff equations are operating, concretely in our case $\mathbb{R}^n$. Let it act on M by composition on the *left*. Namely for $\eta \in \operatorname{Diff}(\mathbb{R}^n)$, we let

$$\eta \cdot \mathbf{Q} = \eta \circ \mathbf{Q}. \tag{9.11}$$

Now lift this action to the cotangent bundle $T^* \operatorname{Emb}(S, \mathbb{R}^n)$ in the standard way. This lifted action is a symplectic (and hence Poisson) action and has an equivariant momentum map. This momentum map will turn out to be the ansatz (9.5).

First let us recall the general formula. Namely, the momentum map is the map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ ($\mathfrak{g}^*$ denotes the dual of the Lie algebra $\mathfrak{g}$ of $\mathfrak{G}$) defined by

$$\mathbf{J}(\alpha_q) \cdot \xi = \langle \alpha_q, \xi_Q(q) \rangle \,, \tag{9.12}$$

where $\alpha_q \in T_q^*Q$ and $\xi \in \mathfrak{g}$, where ξ_Q is the infinitesimal generator of the action of $\mathfrak{G}$ on Q associated to the Lie algebra element ξ, and where $\langle \alpha_q, \xi_Q(q) \rangle$ is the natural pairing of an element of T_q^*Q with an element of T_qQ.

Now we apply formula (9.12) to the special case in which the group $\mathfrak{G}$ is the diffeomorphism group $\operatorname{Diff}(\mathbb{R}^n)$, the manifold Q is $\operatorname{Emb}(S, \mathbb{R}^n)$ and where the action of the group on $\operatorname{Emb}(S, \mathbb{R}^n)$ is given by (9.11). The sense in which the Lie algebra of $\mathfrak{G} = \operatorname{Diff}$ is the space $\mathfrak{g} = \mathfrak{X}$ of vector fields is well-understood. Hence, its dual is naturally regarded as the space of one-form densities. The momentum map is thus a map $\mathbf{J} : T^* \operatorname{Emb}(S, \mathbb{R}^n) \to \mathfrak{X}^*$.

To calculate $\mathbf{J}$ given by (9.12), we first work out the infinitesimal generators. Let $X \in \mathfrak{X}$ be a Lie algebra element. By differentiating the action (9.11) with respect to η in the direction of X at the identity element we find that the infinitesimal generator is given by

$$X_{\operatorname{Emb}(S,\mathbb{R}^n)}(\mathbf{Q}) = X \circ \mathbf{Q}.$$

Thus, taking α_q to be the cotangent vector $(\mathbf{Q}, \mathbf{P})$, equation (9.12) gives

$$\begin{aligned}\langle \mathbf{J}(\mathbf{Q}, \mathbf{P}), X\rangle &= \langle (\mathbf{Q}, \mathbf{P}), X \circ \mathbf{Q}\rangle \\ &= \int_S P_i(s) X^i(\mathbf{Q}(s)) d^k s.\end{aligned}$$

On the other hand, note that the right hand side of (9.5) (again with the index a suppressed, and with t suppressed as well), when paired with the Lie algebra element X is

$$\begin{aligned}\left\langle \int_S \mathbf{P}(s)\, \delta\left(\mathbf{x} - \mathbf{Q}(s)\right) d^k s, X \right\rangle &= \int_{\mathbb{R}^n} \int_S \left(P_i(s)\, \delta\left(\mathbf{x} - \mathbf{Q}(s)\right) d^k s\right) X^i(\mathbf{x}) d^n x \\ &= \int_S P_i(s) X^i(\mathbf{Q}(s) d^k s.\end{aligned}$$

This shows that the expression given by (9.5) is equal to $\mathbf{J}$ and so the result is proved.

$\square$

This proof has shown the following basic fact.

Corollary

9.2. *The singular solution momentum map defined by the singular solution ansatz, namely,*

$$\mathbf{J}_{\text{Sing}} : T^* \operatorname{Emb}(S, \mathbb{R}^n) \to \mathfrak{X}(\mathbb{R}^n)^*$$

is a Poisson map from the canonical Poisson structure on $T^* \operatorname{Emb}(S, \mathbb{R}^n)$ *to the Lie-Poisson structure on* $\mathfrak{X}(\mathbb{R}^n)^*$.

This is perhaps the most basic property of the singular solution momentum map.

9.1.2 The momentum map for right action

There is another group that acts on $\operatorname{Emb}(S, \mathbb{R}^n)$, namely the group $\operatorname{Diff}(S)$ of diffeomorphisms of S, which acts on the *right*, while $\operatorname{Diff}(\mathbb{R}^n)$ acted by composition on the *left* (and this gave rise to our singular solution momentum map, $\mathbf{J}_{\text{Sing}}$). As explained above, the action of $\operatorname{Diff}(S)$ from the right gives us the momentum map $\mathbf{J}_S : T^* \operatorname{Emb}(S, \mathbb{R}^n) \to \mathfrak{X}(S)^*$.

The momentum map $\mathbf{J}_S$ and the Kelvin circulation theorem. The momentum map $\mathbf{J}_{\rm Sing}$ involves $\mathrm{Diff}(\mathbb{R}^n)$, the left action of the diffeomorphism group on the space of embeddings $\mathrm{Emb}(S,\mathbb{R}^n)$ by smooth maps of the target space $\mathbb{R}^n$, namely,

$$\mathrm{Diff}(\mathbb{R}^n): \ \mathbf{Q}\cdot\eta = \eta\circ\mathbf{Q}, \tag{9.13}$$

where, recall, $\mathbf{Q}: S\to\mathbb{R}^n$. As above, the cotangent bundle $T^*\,\mathrm{Emb}(S,\mathbb{R}^n)$ is identified with the space of pairs of maps $(\mathbf{Q},\mathbf{P})$, with $\mathbf{Q}: S\to\mathbb{R}^n$ and $\mathbf{P}: S\to T^*\mathbb{R}^n$.

However, there is another momentum map $\mathbf{J}_S$ associated with the *right action* of the diffeomorphism group of S on the embeddings $\mathrm{Emb}(S,\mathbb{R}^n)$ by smooth maps of the "Lagrangian labels" S (fluid particle relabeling by $\eta: S\to S$). This action is given by

$$\mathrm{Diff}(S): \ \mathbf{Q}\cdot\eta = \mathbf{Q}\circ\eta\,. \tag{9.14}$$

The infinitesimal generator of this right action is

$$X_{\mathrm{Emb}(S,\mathbb{R}^n)}(\mathbf{Q}) = \frac{d}{dt}\Big|_{t=0}\mathbf{Q}\circ\eta_t = T\mathbf{Q}\circ X. \tag{9.15}$$

where $X\in\mathfrak{X}$ is tangent to the curve η_t at $t=0$. Thus, again taking $N=1$ (so we suppress the index a) and also letting α_q in the momentum map formula (9.12) be the cotangent vector $(\mathbf{Q},\mathbf{P})$, one computes $\mathbf{J}_S$:

$$\begin{aligned}
\langle \mathbf{J}_S(\mathbf{Q},\mathbf{P}), X\rangle &= \langle (\mathbf{Q},\mathbf{P}), T\mathbf{Q}\cdot X\rangle \\
&= \int_S P_i(s)\frac{\partial Q^i(s)}{\partial s^m}X^m(s)\,d^ks \\
&= \int_S X\Big(\mathbf{P}(s)\cdot d\mathbf{Q}(s)\Big)\,d^ks \\
&= \left(\int_S \mathbf{P}(s)\cdot d\mathbf{Q}(s)\otimes\,d^ks\,, X(s)\right) \\
&= \langle\,\mathbf{P}\cdot d\mathbf{Q}\,, X\,\rangle\,.
\end{aligned}$$

Consequently, the momentum map formula (9.12) yields

$$\mathbf{J}_S(\mathbf{Q},\mathbf{P}) = \mathbf{P}\cdot d\mathbf{Q}\,, \tag{9.16}$$

with the indicated pairing of the one-form density $\mathbf{P}\cdot d\mathbf{Q}$ with the vector field X. We have set things up so that the following is true.

Proposition

9.3. *The momentum map $\mathbf{J}_S$ is preserved by the evolution equations* (9.9) *for $\mathbf{Q}$ and $\mathbf{P}$.*

Proof. It is enough to notice that the Hamiltonian H_N in equation (9.10) is invariant under the cotangent lift of the action of $\mathrm{Diff}(S)$; it merely amounts to the invariance of the integral over S under reparametrization; that is, the change of variables formula; keep in mind that $\mathbf{P}$ includes a density factor. ■

This result is similar to the Kelvin-Noether theorem for circulation Γ of an ideal fluid, which may be written as $\Gamma = \oint_{c(s)} D(s)^{-1}\mathbf{P}(s)\cdot d\mathbf{Q}(s)$ for each Lagrangian circuit $c(s)$, where D is the mass density and $\mathbf{P}$ is again the canonical momentum density. This similarity should come as no surprise, because the Kelvin-Noether theorem for ideal fluids arises from invariance of Hamilton's principle under fluid parcel relabeling by the *same* right action of the diffeomorphism group, as in (9.14). Note that, being an equivariant momentum map, the map $\mathbf{J}_S$, as with $\mathbf{J}_{\text{Sing}}$, is also a Poisson map. That is, substituting the canonical Poisson bracket into relation (9.16); that is, the relation $\mathbf{M}(\mathbf{x}) = \sum_i P_i(\mathbf{x})\nabla Q^i(\mathbf{x})$ yields the Lie-Poisson bracket on the space of $\mathbf{M}$'s. We use the different notations $\mathbf{m}$ and $\mathbf{M}$ because these quantitites are analogous to the body and spatial angular momentum for rigid body mechanics. In fact, the quantity $\mathbf{m}$ given by the solution Ansatz; specifically, $\mathbf{m} = \mathbf{J}_{\text{Sing}}(\mathbf{Q},\mathbf{P})$ gives the singular solutions of the EPDiff equations, while $\mathbf{M}(\mathbf{x}) = \mathbf{J}_S(\mathbf{Q},\mathbf{P}) = \sum_i P_i(\mathbf{x})\nabla Q^i(\mathbf{x})$ is a conserved quantity. In the language of fluid mechanics, the expression of $\mathbf{m}$ in terms of $(\mathbf{Q},\mathbf{P})$ is an example of a "Clebsch representation," which expresses the solution of the EPDiff equations in terms of canonical variables that evolve by standard canonical Hamilton equations. This has been known in the case of fluid mechanics for more than 100 years. For modern discussions of the Clebsch representation for ideal fluids, see, for example, [?, ?].

One more remark is in order; namely the special case in which $S = M$ is of course allowed. In this case, $\mathbf{Q}$ corresponds to the map η itself and $\mathbf{P}$ just corresponds to its conjugate momentum. The quantity $\mathbf{m}$ corresponds to the spatial (dynamic) momentum density (that is, right translation of $\mathbf{P}$ to the identity), while $\mathbf{M}$ corresponds to the conserved "body" momentum density (that is, left translation of $\mathbf{P}$ to the identity).

9.1.3 A dual pair of momentum maps for left and right action

We now assemble both momentum maps into one figure as follows:

$$\begin{tikzcd} & T^*\,\mathrm{Emb}(S,M) \arrow[dl, "\mathbf{J}_{\text{Sing}}"'] \arrow[dr, "\mathbf{J}_S"] & \\ \mathfrak{X}(M)^* & & \mathfrak{X}(S)^* \end{tikzcd}$$

These maps have the formal dual pair property, namely that the kernel of the derivatives of each map at a given point are symplectic orthogonals of one another. The proof of this aspect of the dual pair properties is beyond the scope of this course. For that proof, see [GBVi2012].

9.2 Left and right momentum maps for fluids

The basic idea for the description of fluid dynamics by the action of diffeomorphisms is sketched in Fig 5.

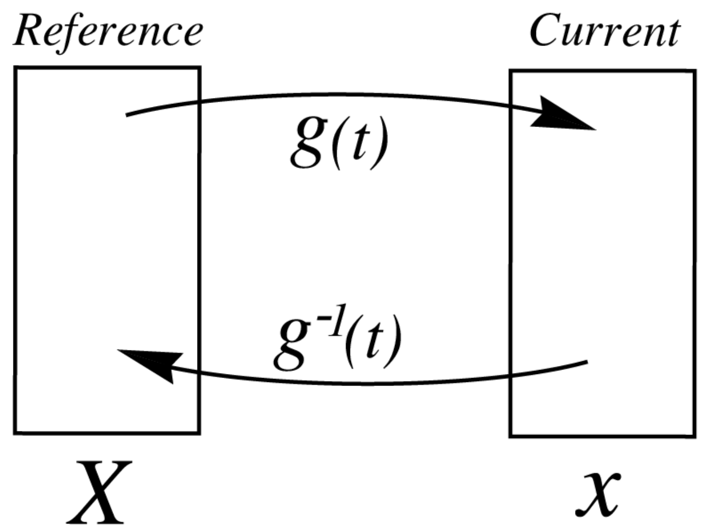


Figure 5: The forward and inverse group actions $g(t)$ and $g^{-1}(t)$ that represent ideal fluid flow are sketched here.

The forward and inverse maps sketched in Fig 5 represent ideal fluid flow by *left* group action of $g_t \in \mathrm{Diff}$ on reference ($X \in M$) and current ($x \in M$) coordinates. They are denoted as,

$$g_t : x(t,X) = g_t X \quad \text{and} \quad g_t^{-1} : X(t,x) = g_t^{-1} x\,, \tag{9.17}$$

so that taking time derivatives yields

$$\dot{x}(t,X) = \dot{g}_t X = (\dot{g}_t g_t^{-1})x = \pounds_u x =: u(x,t) = u_t \circ g_t X\,, \tag{9.18}$$

and

$$\dot{X}(t,x) = (T_x g_t^{-1})(\dot{g}_t g_t^{-1} x) = T_x X \cdot u = \pounds_u X =: V(X,t) = V_t \circ g_t^{-1} x\,. \tag{9.19}$$

Here $u = \dot{g}_t g_t^{-1}$ is called the *Eulerian* velocity, and $V = \mathrm{Ad}_{g_t^{-1}} u$ is called the *convective* velocity. For $O_t \in SO(3)$, these correspond to the *spatial* angular velocity $\omega = \dot{O}_t O_t^{-1}$ and the *body* angular velocity $\Omega = \mathrm{Ad}_{O_t^{-1}}\omega = O_t^{-1}\dot{O}_t$. We shall mainly deal with the Eulerian fluid velocity in these notes.

Exercise. Use the Clebsch method to compute the momentum maps for the left group actions in (9.17). ★

9.3 The Euler-Poincaré framework for ideal fluids [HoMaRa1998a]

Almost all fluid models of interest admit the following general assumptions. These assumptions form the basis of the Euler-Poincaré theorem for ideal fluids that we shall state later in this section, after introducing the notation necessary for dealing geometrically with the reduction of Hamilton's Principle from the material (or Lagrangian) picture of fluid dynamics, to the spatial (or Eulerian) picture. This theorem was first stated and proved in [HoMaRa1998a], to which we refer for additional details, as well as for abstract definitions and proofs.

Basic assumptions underlying the Euler-Poincaré theorem for continua

- There is a *right* representation of a Lie group G on the vector space V and G acts in the natural way on the *right* on $TG \times V^*$: $(U_g, a)h = (U_g h, ah)$.

- The Lagrangian function $L : TG \times V^* \to \mathbb{R}$ is right G–invariant.[9]

- In particular, if $a_0 \in V^*$, define the Lagrangian $L_{a_0} : TG \to \mathbb{R}$ by $L_{a_0}(U_g) = L(U_g, a_0)$. Then L_{a_0} is right invariant under the lift to TG of the right action of G_{a_0} on G, where G_{a_0} is the isotropy group of a_0.

- Right G–invariance of L permits one to define the Lagrangian on the Lie algebra $\mathfrak{g}$ of the group G. Namely, $\ell : \mathfrak{g} \times V^* \to \mathbb{R}$ is defined by,
$$\ell(u, a) = L\big(U_g g^{-1}(t), a_0 g^{-1}(t)\big) = L(U_g, a_0)\,,$$
where $u = U_g g^{-1}(t)$ and $a = a_0 g^{-1}(t)$, Conversely, this relation defines for any $\ell : \mathfrak{g} \times V^* \to \mathbb{R}$ a function $L : TG \times V^* \to \mathbb{R}$ that is right G–invariant, up to relabeling of a_0.

- For a curve $g(t) \in G$, let $u(t) := \dot{g}(t)g(t)^{-1}$ and define the curve $a(t)$ as the unique solution of the linear differential equation with time dependent coefficients $\dot{a}(t) = -\, a(t)u(t) = \pounds_u a(t)$, where the right action of an element of the Lie algebra $u \in \mathfrak{g}$ on an advected quantity $a \in V^*$ is denoted by concatenation from the right. The solution with initial condition $a(0) = a_0 \in V^*$ can be written as $a(t) = a_0 g(t)^{-1}$.

[9]For fluid dynamics, right G–invariance of the Lagrangian function L is traditionally called "particle relabeling symmetry."

Notation for reduction of Hamilton's Principle by symmetries

- Let $\mathfrak{g}(\mathcal{D})$ denote the space of vector fields on $\mathcal{D}$ of some fixed differentiability class. These vector fields are endowed with the ***Lie bracket*** given in components by (summing on repeated indices)

$$[\mathbf{u},\mathbf{v}]^i = u^j\frac{\partial v^i}{\partial x^j} - v^j\frac{\partial u^i}{\partial x^j} =: -\left(\mathrm{ad}_{\mathbf{u}}\,\mathbf{v}\right)^i . \tag{9.20}$$

  The notation $\mathrm{ad}_{\mathbf{u}}\,\mathbf{v} := -[\mathbf{u},\,\mathbf{v}]$ formally denotes the adjoint action of the *right* Lie algebra of $\mathrm{Diff}(\mathcal{D})$ on itself. This Lie algebra is given by the smooth right-invariant vector fields, $\mathfrak{g} = \mathfrak{X}$.

- Identify the Lie algebra of vector fields $\mathfrak{g}$ with its dual $\mathfrak{g}^*$ by using the L^2 pairing

$$\langle \mathbf{u},\mathbf{v}\rangle = \int_{\mathcal{D}} \mathbf{u}\cdot\mathbf{v}\,dV . \tag{9.21}$$

- Let $\mathfrak{g}(\mathcal{D})^*$ denote the geometric dual space of $\mathfrak{g}(\mathcal{D})$, that is, $\mathfrak{g}(\mathcal{D})^* := \Lambda^1(\mathcal{D})\otimes \mathrm{Den}(\mathcal{D})$. This is the space of one–form densities on $\mathcal{D}$. If $\mathbf{m}\otimes dV \in \Lambda^1(\mathcal{D})\otimes\mathrm{Den}(\mathcal{D})$, then the pairing of $\mathbf{m}\otimes dV$ with $\mathbf{u}\in\mathfrak{g}(\mathcal{D})$ is given by the L^2 pairing,

$$\langle \mathbf{m}\otimes dV,\mathbf{u}\rangle = \int_{\mathcal{D}} \mathbf{m}\cdot\mathbf{u}\,dV \tag{9.22}$$

  where $\mathbf{m}\cdot\mathbf{u}$ is the standard contraction of a one–form $\boldsymbol{m}$ with a vector field $\mathbf{u}$.

- For $\mathbf{u}\in\mathfrak{g}(\mathcal{D})$ and $\mathbf{m}\otimes dV\in\mathfrak{g}(\mathcal{D})^*$, the dual of the adjoint representation is defined by

$$\langle \mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV),\mathbf{v}\rangle = \int_{\mathcal{D}} \mathbf{m}\cdot\mathrm{ad}_{\mathbf{u}}\mathbf{v}\,dV = -\int_{\mathcal{D}}\mathbf{m}\cdot[\mathbf{u},\mathbf{v}]\,dV \tag{9.23}$$

  and its expression is

$$\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV) = (\pounds_{\mathbf{u}}\mathbf{m} + (\mathrm{div}_{dV}\,\mathbf{u})\mathbf{m})\otimes dV = \pounds_{\mathbf{u}}(\mathbf{m}\otimes dV)\,, \tag{9.24}$$

  where $\mathrm{div}_{dV}\mathbf{u}$ is the divergence of $\mathbf{u}$ relative to the measure dV, that is, $\pounds_{\mathbf{u}}dV = (\mathrm{div}_{dV}\mathbf{u})dV$. Hence, $\mathrm{ad}^*_{\mathbf{u}}$ coincides with the Lie-derivative $\pounds_{\mathbf{u}}$ for one-form densities.

- If $\mathbf{u} = u^j\partial/\partial x^j$, $\mathbf{m} = m_i dx^i$, then the one–form factor in the preceding formula for $\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m}\otimes dV)$ has the ***coordinate expression***

$$\Big(\mathrm{ad}^*_{\mathbf{u}}\,\mathbf{m}\Big)_i dx^i = \left(u^j\frac{\partial m_i}{\partial x^j} + m_j\frac{\partial u^j}{\partial x^i} + (\mathrm{div}_{dV}\,\mathbf{u})m_i\right)dx^i = \left(\frac{\partial}{\partial x^j}(u^j m_i) + m_j\frac{\partial u^j}{\partial x^i}\right)dx^i . \tag{9.25}$$

  The last equality assumes that the divergence is taken relative to the standard measure $dV = d^n\mathbf{x}$ in $\mathbb{R}^n$. (On a Riemannian manifold the metric divergence needs to be used.)

Definition

9.4. The ***representation space*** V^* of $\mathrm{Diff}(\mathcal{D})$ in continuum mechanics is often some subspace of the tensor field densities on $\mathcal{D}$, denoted as $\mathfrak{T}(\mathcal{D})\otimes\mathrm{Den}(\mathcal{D})$, and the representation is given by pull back. It is thus a *right* representation of $\mathrm{Diff}(\mathcal{D})$ on $\mathfrak{T}(\mathcal{D})\otimes\mathrm{Den}(\mathcal{D})$. The right action of the Lie algebra $\mathfrak{g}(\mathcal{D})$ on V^* is denoted as ***concatenation from the right***. That is, we denote

$$a\mathbf{u} := \pounds_{\mathbf{u}} a\,,$$

which is the Lie derivative of the tensor field density a along the vector field $\mathbf{u}$.

Definition

9.5. The ***Lagrangian of a continuum mechanical system*** is a function

$$L : T\,\mathrm{Diff}(\mathcal{D}) \times V^* \to \mathbb{R}\,,$$

which is right invariant relative to the tangent lift of right translation of $\mathrm{Diff}(\mathcal{D})$ on itself and pull back on the tensor field densities. Invariance of the Lagrangian L induces a function $\ell : \mathfrak{g}(\mathcal{D}) \times V^* \to \mathbb{R}$ given by

$$\ell(\mathbf{u}, a) = L(\mathbf{u}\circ\eta, \eta^* a) = L(\mathbf{U}, a_0)\,,$$

where $\mathbf{u} \in \mathfrak{g}(\mathcal{D})$ and $a \in V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, and where $\eta^* a$ denotes the pull back of a by the diffeomorphism η and $\mathbf{u}$ is the Eulerian velocity. That is,

$$\mathbf{U} = \mathbf{u}\circ\eta \quad \text{and} \quad a_0 = \eta^* a\,. \tag{9.26}$$

The evolution of a is by right action, given by the equation

$$\dot{a} = -\,\pounds_{\mathbf{u}}\, a = -\,a\mathbf{u}. \tag{9.27}$$

The solution of this equation, for the initial condition a_0, is

$$a(t) = \eta_{t*} a_0 = a_0 g^{-1}(t)\,, \tag{9.28}$$

where the lower star denotes the push forward operation and η_t is the flow of $\mathbf{u} = \dot{g}g^{-1}(t)$.

Definition

9.6. ***Advected Eulerian quantities*** are defined in continuum mechanics to be those variables which are Lie transported by the flow of the Eulerian velocity field. Using this standard terminology, equation (9.27), or its solution (9.28) states that the tensor field density $a(t)$ (which may include mass density and other Eulerian quantities) is advected.

Remark

9.7 (Dual tensors)**.** As we mentioned, typically $V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$ for continuum mechanics. On a general manifold, tensors of a given type have natural duals. For example, symmetric covariant tensors are dual to symmetric contravariant tensor densities, the pairing being given by the integration of the natural contraction of these tensors. Likewise, k–forms are naturally dual to $(n-k)$–forms, the pairing being given by taking the integral of their wedge product.

Definition

9.8. The ***diamond operation*** $\diamond$ between elements of V and V^* produces an element of the dual Lie algebra $\mathfrak{g}(\mathcal{D})^*$ and is defined as

$$\langle b \diamond a, \mathbf{w} \rangle = -\int_{\mathcal{D}} b \cdot \pounds_{\mathbf{w}}\, a \, , \tag{9.29}$$

where $b \cdot \pounds_{\mathbf{w}}\, a$ denotes the contraction, as described above, of elements of V and elements of V^* and $\mathbf{w} \in \mathfrak{g}(\mathcal{D})$. (These operations do *not* depend on a Riemannian structure.)

For a path $\eta_t \in \mathrm{Diff}(\mathcal{D})$, let $\mathbf{u}(x,t)$ be its Eulerian velocity and consider the curve $a(t)$ with initial condition a_0 given by the equation

$$\dot{a} + \pounds_{\mathbf{u}} a = 0. \tag{9.30}$$

Let the Lagrangian $L_{a_0}(\mathbf{U}) := L(\mathbf{U}, a_0)$ be right-invariant under $\mathrm{Diff}(\mathcal{D})$. We can now state the Euler–Poincaré Theorem for Continua of [HoMaRa1998a].

Theorem

9.9 (Euler–Poincaré Theorem for Continua.). *Consider a path η_t in $\mathrm{Diff}(\mathcal{D})$ with Lagrangian velocity $\mathbf{U}$ and Eulerian velocity $\mathbf{u}$. The following are equivalent:*

i *Hamilton's variational principle*

$$\delta \int_{t_1}^{t_2} L\left(X, \mathbf{U}_t(X), a_0(X)\right)\, dt = 0 \tag{9.31}$$

holds, for variations $\delta\eta_t$ vanishing at the endpoints.

ii *η_t satisfies the Euler–Lagrange equations for L_{a_0} on $\mathrm{Diff}(\mathcal{D})$.*

iii *The constrained variational principle in Eulerian coordinates*

$$\delta \int_{t_1}^{t_2} \ell(\mathbf{u}, a)\, dt = 0 \tag{9.32}$$

holds on $\mathfrak{g}(\mathcal{D}) \times V^*$, using variations of the form

$$\delta \mathbf{u} = \frac{\partial \mathbf{w}}{\partial t} + [\mathbf{u}, \mathbf{w}] = \frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\,\mathbf{u}} \mathbf{w}\,, \qquad \delta a = -\pounds_{\mathbf{w}}\, a, \tag{9.33}$$

where $\mathbf{w}_t = \delta\eta_t \circ \eta_t^{-1}$ vanishes at the endpoints.

iv *The Euler–Poincaré equations for continua*

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \mathbf{u}} = -\,\mathrm{ad}^*_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a = -\pounds_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a\,, \tag{9.34}$$

hold, with auxiliary equations $(\partial_t + \pounds_{\mathbf{u}})a = 0$ for each advected quantity $a(t)$. The $\diamond$ operation defined in (9.29) needs to be determined on a case by case basis, depending on the nature of the tensor $a(t)$. The variation $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ is a one–form density and we have used relation (9.24) in the last step of equation (9.34).

We refer to [HoMaRa1998a] for the proof of this theorem in the abstract setting. We shall see some of the features of this result in the concrete setting of continuum mechanics shortly.

Discussion of the Euler-Poincaré equations

The following string of equalities shows *directly* that **iii** is equivalent to **iv**:

$$\begin{aligned} 0 &= \delta \int_{t_1}^{t_2} l(\mathbf{u}, a) dt = \int_{t_1}^{t_2} \left(\frac{\delta l}{\delta \mathbf{u}} \cdot \delta \mathbf{u} + \frac{\delta l}{\delta a} \cdot \delta a \right) dt \\ &= \int_{t_1}^{t_2} \left[\frac{\delta l}{\delta \mathbf{u}} \cdot \left(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\, \mathbf{w} \right) - \frac{\delta l}{\delta a} \cdot \pounds_{\mathbf{w}}\, a \right] dt \\ &= \int_{t_1}^{t_2} \mathbf{w} \cdot \left[- \frac{\partial}{\partial t} \frac{\delta l}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}} \frac{\delta l}{\delta \mathbf{u}} + \frac{\delta l}{\delta a} \diamond a \right] dt \,. \end{aligned} \tag{9.35}$$

The rest of the proof follows essentially the same track as the proof of the pure Euler-Poincaré theorem, modulo slight changes to accomodate the advected quantities.

In the absence of dissipation, most Eulerian fluid equations[10] can be written in the EP form in equation (9.34),

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} = \frac{\delta \ell}{\delta a} \diamond a \,, \quad \text{with} \quad \big(\partial_t + \pounds_{\mathbf{u}}\big) a = 0 \,. \tag{9.36}$$

Equation (9.36) is ***Newton's Law***: The Eulerian time derivative of the momentum density $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ (a one-form density dual to the velocity $\mathbf{u}$) is equal to the force density $(\delta\ell/\delta a) \diamond a$, with the $\diamond$ operation defined in (9.29). Thus, Newton's Law is written in the Eulerian fluid representation as,[11]

$$\frac{d}{dt}\Big|_{Lag} \mathbf{m} := \big(\partial_t + \pounds_{\mathbf{u}}\big)\mathbf{m} = \frac{\delta \ell}{\delta a} \diamond a \,, \quad \text{with} \quad \frac{d}{dt}\Big|_{Lag} a := \big(\partial_t + \pounds_{\mathbf{u}}\big) a = 0 \,. \tag{9.37}$$

- The left side of the EP equation in (9.37) describes the fluid's dynamics due to its kinetic energy. A fluid's kinetic energy typically defines a norm for the Eulerian fluid velocity, $KE = \frac{1}{2}\|\mathbf{u}\|^2$. The left side of the EP equation is the ***geodesic*** part

[10]Exceptions to this statement are certain multiphase fluids, and complex fluids with active internal degrees of freedom such as liquid crystals. These require a further extension, not discussed here.

[11]In coordinates, a one-form density takes the form $\mathbf{m} \cdot d\mathbf{x} \otimes dV$ and the EP equation (9.34) is given neumonically by

$$\frac{d}{dt}\Big|_{Lag} \big(\mathbf{m} \cdot d\mathbf{x} \otimes dV\big) = \underbrace{\frac{d\mathbf{m}}{dt}\Big|_{Lag} \cdot d\mathbf{x} \otimes dV}_{\text{Advection}} + \underbrace{\mathbf{m} \cdot d\mathbf{u} \otimes dV}_{\text{Stretching}} + \underbrace{\mathbf{m} \cdot d\mathbf{x} \otimes (\nabla \cdot \mathbf{u}) dV}_{\text{Expansion}} = \frac{\delta \ell}{\delta a} \diamond a$$

with $\frac{d}{dt}\Big|_{Lag} d\mathbf{x} := \big(\partial_t + \pounds_{\mathbf{u}}\big) d\mathbf{x} = d\mathbf{u} = \mathbf{u}_{,j} dx^j$, upon using commutation of Lie derivative and exterior derivative. Compare this formula with the definition of $\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV)$ in equation (9.25).

of its evolution, with respect to this norm. See [Ar1966, Ar1979, ArKh1998] for discussions of this interpretation of ideal incompressible flow and references to the literature. However, in a gravitational field, for example, there will also be dynamics due to potential energy. And this dynamics will by governed by the right side of the EP equation.

- The right side of the EP equation in (9.37) modifies the geodesic motion. Naturally, the right side of the EP equation is also a geometrical quantity. The diamond operation $\diamond$ represents the dual of the Lie algebra action of vectors fields on the tensor a. Here $\delta\ell/\delta a$ is the dual tensor, under the natural pairing (usually, L^2 pairing) $\langle\,\cdot\,,\cdot\,\rangle$ that is induced by the variational derivative of the Lagrangian $\ell(\mathbf{u}, a)$. The diamond operation $\diamond$ is defined in terms of this pairing in (9.29). For the L^2 pairing, this is integration by parts of (minus) the Lie derivative in (9.29).

- The quantity a is typically a tensor (e.g., a density, a scalar, or a differential form) and we shall sum over the various types of tensors a that are involved in the fluid description. The second equation in (9.37) states that each tensor a is carried along by the Eulerian fluid velocity $\mathbf{u}$. Thus, a is for fluid "attribute," and its Eulerian evolution is given by minus its Lie derivative, $-\pounds_{\mathbf{u}}a$. That is, a stands for the set of fluid attributes that each Lagrangian fluid parcel carries around (advects), such as its buoyancy, which is determined by its individual salt, or heat content, in ocean circulation.

- Many examples of how equation (9.37) arises in the dynamics of continuous media are given in [HoMaRa1998a]. The EP form of the Eulerian fluid description in (9.37) is analogous to the classical dynamics of rigid bodies (and tops, under gravity) in body coordinates. Rigid bodies and tops are also governed by Euler-Poincaré equations, as Poincaré showed in a two-page paper with no references, over a century ago [Po1901]. For modern discussions of the EP theory, see, e.g., [MaRa1999], or [HoMaRa1998a].

Exercise. State the EP theorem and write the EP equations for the *convective* velocity. ★

9.4 Corollary of the EP theorem: the Kelvin-Noether circulation theorem

Corollary

9.10 (Kelvin-Noether Circulation Theorem.). Assume $\mathbf{u}(x,t)$ satisfies the Euler–Poincaré equations for continua:

$$\frac{\partial}{\partial t}\left(\frac{\delta \ell}{\delta \mathbf{u}}\right) = -\pounds_{\mathbf{u}}\left(\frac{\delta \ell}{\delta \mathbf{u}}\right) + \frac{\delta \ell}{\delta a}\diamond a$$

and the quantity a satisfies the ***advection relation***

$$\frac{\partial a}{\partial t} + \pounds_{\mathbf{u}} a = 0. \tag{9.38}$$

Let η_t be the flow of the Eulerian velocity field $\mathbf{u}$, that is, $\mathbf{u} = (d\eta_t/dt)\circ\eta_t^{-1}$. Define the advected fluid loop $\gamma_t := \eta_t \circ \gamma_0$ and the circulation map $I(t)$ by

$$I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}}\,. \tag{9.39}$$

In the circulation map $I(t)$ the advected mass density D_t satisfies the push forward relation $D_t = \eta_* D_0$. This implies the advection relation (9.38) with $a = D$, namely, the continuity equation,

$$\partial_t D + \operatorname{div} D\mathbf{u} = 0\,.$$

Then the map $I(t)$ satisfies the ***Kelvin circulation relation***,

$$\frac{d}{dt} I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta a}\diamond a\;. \tag{9.40}$$

Both an abstract proof of the Kelvin-Noether Circulation Theorem and a proof tailored for the case of continuum mechanical systems are given in [HoMaRa1998a]. We provide a version of the latter below.

Proof. First we change variables in the expression for $I(t)$:

$$I(t) = \oint_{\gamma_t} \frac{1}{D_t}\frac{\delta l}{\delta \mathbf{u}} = \oint_{\gamma_0} \eta_t^*\left[\frac{1}{D_t}\frac{\delta l}{\delta \mathbf{u}}\right] = \oint_{\gamma_0} \frac{1}{D_0}\eta_t^*\left[\frac{\delta l}{\delta \mathbf{u}}\right].$$

Next, we use the Lie derivative formula, namely

$$\frac{d}{dt}\left(\eta_t^*\alpha_t\right) = \eta_t^*\left(\frac{\partial}{\partial t}\alpha_t + \pounds_{\mathbf{u}}\alpha_t\right)\,,$$

applied to a one–form density α_t. This formula gives

$$\begin{aligned}
\frac{d}{dt}I(t) &= \frac{d}{dt}\oint_{\gamma_0}\frac{1}{D_0}\eta_t^*\left[\frac{\delta l}{\delta\mathbf{u}}\right] \\
&= \oint_{\gamma_0}\frac{1}{D_0}\frac{d}{dt}\left(\eta_t^*\left[\frac{\delta l}{\delta\mathbf{u}}\right]\right) \\
&= \oint_{\gamma_0}\frac{1}{D_0}\eta_t^*\left[\frac{\partial}{\partial t}\left(\frac{\delta l}{\delta\mathbf{u}}\right) + \pounds_{\mathbf{u}}\left(\frac{\delta l}{\delta\mathbf{u}}\right)\right].
\end{aligned}$$

By the Euler–Poincaré equations (9.34), this becomes

$$\frac{d}{dt}I(t) = \oint_{\gamma_0}\frac{1}{D_0}\eta_t^*\left[\frac{\delta l}{\delta a}\diamond a\right] = \oint_{\gamma_t}\frac{1}{D_t}\left[\frac{\delta l}{\delta a}\diamond a\right],$$

again by the change of variables formula. $\square$

Corollary

9.11. Since the last expression holds for every loop γ_t, we may write it as

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right)\frac{1}{D}\frac{\delta l}{\delta\mathbf{u}} = \frac{1}{D}\frac{\delta l}{\delta a}\diamond a\,. \tag{9.41}$$

Remark

9.12. The Kelvin-Noether theorem is called so here because its derivation relies on the invariance of the Lagrangian L under the particle relabeling symmetry, and Noether's theorem is associated with this symmetry. However, the result (9.40) is the ***Kelvin circulation theorem***: the circulation integral $I(t)$ around any fluid loop (γ_t, moving with the velocity of the fluid parcels $\mathbf{u}$) is invariant under the fluid motion. These two statements are equivalent. We note that ***two velocities*** appear in the integrand $I(t)$: the fluid velocity $\mathbf{u}$ and $D^{-1}\delta\ell/\delta\mathbf{u}$. The latter velocity is the momentum density $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ divided by the mass density D. These two velocities are the basic ingredients for performing modeling and analysis in any ideal fluid problem. One simply needs to put these ingredients together in the Euler-Poincaré theorem and its corollary, the Kelvin-Noether theorem.

9.5 The Hamiltonian formulation of ideal fluid dynamics

Legendre transform Taking the Legendre-transform of the Lagrangian $l(u,a)$: $\mathfrak{g}\times V\to\mathbb{R}$ yields the Hamiltonian $h(m,a):\mathfrak{g}^*\times V\to\mathbb{R}$, given by

$$h(m,a)=\Big\langle m\,,\,u\Big\rangle-l(u,a)\,. \tag{9.42}$$

Differentiating the Hamiltonian determines its partial derivatives:

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

The middle term vanishes because $m-\delta l/\delta u=0$ defines m. These derivatives allow one to rewrite the Euler–Poincaré equation for continua in (9.34) solely in terms of momentum m and advected quantities a as

$$\begin{aligned}
\partial_t m &= -\,\mathrm{ad}^*_{\delta h/\delta m}\,m-\frac{\delta h}{\delta a}\diamond a\,,\\
\partial_t a &= -\,\pounds_{\delta h/\delta m}\,a\,.
\end{aligned}\tag{9.43}$$

Hamiltonian equations The corresponding Hamiltonian equation for any functional of $f(m,a)$ is then

$$\begin{aligned}
\frac{d}{dt}f(m,a) &= \Big\langle \partial_t m\,,\,\frac{\delta f}{\delta m}\Big\rangle+\Big\langle\partial_t a\,,\,\frac{\delta f}{\delta a}\Big\rangle\\
&= -\Big\langle \mathrm{ad}^*_{\delta h/\delta m}\,m+\frac{\delta h}{\delta a}\diamond a\,,\,\frac{\delta f}{\delta m}\Big\rangle-\Big\langle\pounds_{\delta h/\delta m}\,a\,,\,\frac{\delta f}{\delta a}\Big\rangle\\
&= -\Big\langle m\,,\,\Big[\frac{\delta f}{\delta m},\,\frac{\delta h}{\delta m}\Big]\Big\rangle+\Big\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a}-\pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\Big\rangle\\
&=:\ \{f\,,\,h\}(m,a)\,,
\end{aligned}\tag{9.44}$$

which is plainly antisymmetric under the exchange $f\leftrightarrow h$. Assembling these equations into Hamiltonian form gives, symbolically,

$$\frac{\partial}{\partial t}\begin{bmatrix} m\\ a\end{bmatrix}=-\begin{bmatrix}\mathrm{ad}^*_{\square}m & \square\diamond a\\ \pounds_{\square}a & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta m\\ \delta h/\delta a\end{bmatrix} \tag{9.45}$$

The boxes $\Box$ in Equation (9.45) indicate how the various operations are applied in the matrix multiplication. For example,

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

so each matrix entry acts on its corresponding vector component.

Remark

9.13. *The expression*

$$\{f\,,\,h\}(m,a) = -\left\langle m\,,\,\left[\frac{\delta f}{\delta m},\,\frac{\delta h}{\delta m}\right]\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\right\rangle$$

in (9.44) *defines the Lie-Poisson bracket on the dual to the semidirect-product Lie algebra $\mathfrak{X}$ⓈV^* with Lie bracket*

$$\mathrm{ad}_{(u,\alpha)}\,(\overline{u},\,\overline{\alpha}) = \big(\mathrm{ad}_u\overline{u},\,\pounds^T_u\overline{\alpha} - \pounds^T_{\overline{u}}\alpha\big)$$

The coordinates are velocity vector field $u\in\mathfrak{X}$ dual to momentum density $m\in\mathfrak{X}^*$ and $\alpha\in V^*$ dual to the vector space of advected quantities $a\in V$.

Proof. We check that

$$\begin{aligned}
\frac{df}{dt}(m,a) = \{f\,,\,h\}(m,a) &= \left\langle m\,,\,\mathrm{ad}_{\frac{\delta f}{\delta m}}\frac{\delta h}{\delta m}\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\right\rangle \\
&= -\left\langle \mathrm{ad}^*_{\frac{\delta h}{\delta m}}m\,,\,\frac{\delta f}{\delta m}\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a}\right\rangle + \left\langle -\pounds_{\delta h/\delta m}a\,,\,\frac{\delta f}{\delta a}\right\rangle \\
&= -\left\langle \mathrm{ad}^*_{\frac{\delta h}{\delta m}}m + \frac{\delta h}{\delta a}\diamond a\,,\,\frac{\delta f}{\delta m}\right\rangle - \left\langle \pounds_{\delta h/\delta m}a\,,\,\frac{\delta f}{\delta a}\right\rangle
\end{aligned}$$

Note that the angle brackets refer to different types of pairings. This should cause no confusion. $\Box$

10 Worked Example: Euler–Poincaré theorem for GFD

Figure 6 shows a screen shot of numerical simulations of damped and driven geophysical fluid dynamics (GFD) equations of the type studied in this section, taken from `http://www.youtube.com/watch?v=ujBi9Ba8hqs&feature=youtu.be`. The variations in space and time of the driving and damping by the Sun are responsible for the characteristic patterns of the flow. The nonlinear GFD equations in the absence of damping and driving are formulated in this section by using the Euler–Poincaré theorem.

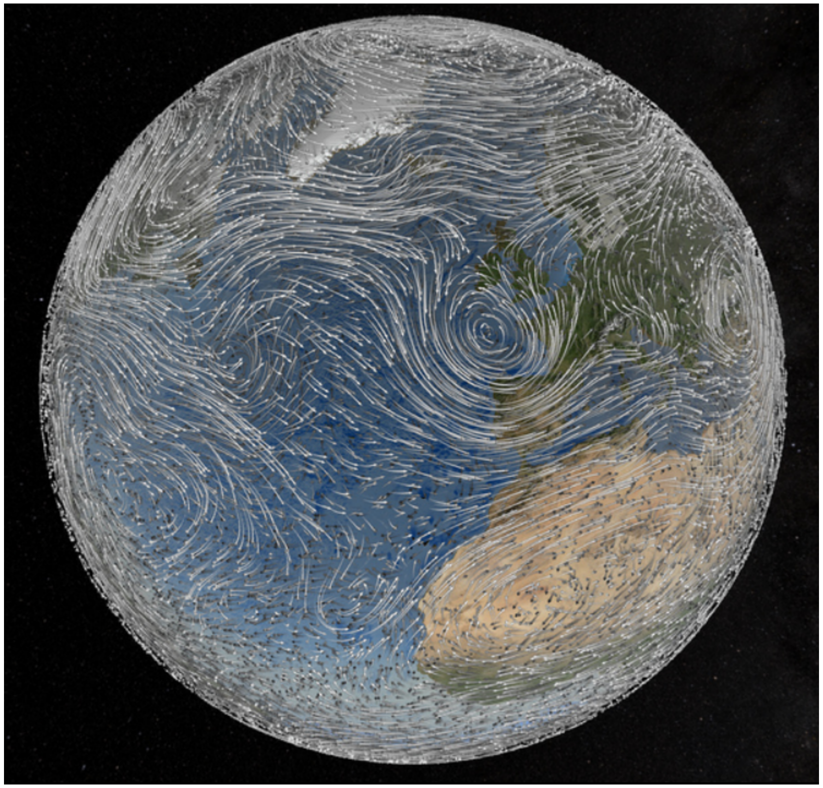


Figure 6: Atmospheric flows on Earth (wind currents) are driven by the Sun and its interaction with the surface and they are damped primarily by friction with the surface.

10.1 Variational Formulae in Three Dimensions

We compute explicit formulae for the variations δa in the cases that the set of tensors a is drawn from a set of scalar fields and densities on $\mathbb{R}^3$. We shall denote this symbolically by writing

$$a \in \{b, D\, d^3x\}\,. \tag{10.1}$$

We have seen that invariance of the set a in the Lagrangian picture under the dynamics of $\mathbf{u}$ implies in the Eulerian picture that

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) a = 0\,,$$

where $\pounds_{\mathbf{u}}$ denotes Lie derivative with respect to the velocity vector field $\mathbf{u}$. Hence, for a fluid dynamical Eulerian action $\mathfrak{S} = \int dt\ \ell(\mathbf{u}; b, D)$, the advected variables b and D satisfy the following Lie-derivative relations,

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) b = 0, \quad\text{or}\quad \frac{\partial b}{\partial t} = -\,\mathbf{u}\cdot\nabla\, b\,, \tag{10.2}$$

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) D\, d^3x = 0, \quad\text{or}\quad \frac{\partial D}{\partial t} = -\,\nabla\cdot(D\mathbf{u})\,. \tag{10.3}$$

In fluid dynamical applications, the advected Eulerian variables b and $D\,d^3x$ represent the buoyancy b (or specific entropy, for the compressible case) and volume element (or mass density) $D\,d^3x$, respectively. According to Theorem 9.9, equation (9.32), the variations of the tensor functions a at fixed $\mathbf{x}$ and t are also given by Lie derivatives, namely $\delta a = -\,\pounds_{\mathbf{w}}\, a$, or

$$\begin{aligned} \delta b &= -\,\pounds_{\mathbf{w}}\, b = -\,\mathbf{w}\cdot\nabla\, b\,, \\ \delta D\ d^3x &= -\,\pounds_{\mathbf{w}}\,(D\,d^3x) = -\,\nabla\cdot(D\mathbf{w})\ d^3x\,. \end{aligned} \tag{10.4}$$

Hence, Hamilton's principle (9.32) with this dependence yields

$$\begin{aligned} 0 &= \delta\int dt\ \ell(\mathbf{u}; b, D) \\ &= \int dt\ \left[\frac{\delta \ell}{\delta \mathbf{u}}\cdot\delta\mathbf{u} + \frac{\delta \ell}{\delta b}\ \delta b + \frac{\delta \ell}{\delta D}\ \delta D\right] \\ &= \int dt\ \left[\frac{\delta \ell}{\delta \mathbf{u}}\cdot\Big(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\,\mathbf{w}\Big) - \frac{\delta \ell}{\delta b}\ \mathbf{w}\cdot\nabla\, b - \frac{\delta \ell}{\delta D}\ \Big(\nabla\cdot(D\mathbf{w})\Big)\right] \\ &= \int dt\ \mathbf{w}\cdot\left[-\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}}\,\frac{\delta \ell}{\delta \mathbf{u}} - \frac{\delta \ell}{\delta b}\ \nabla\, b + D\ \nabla\frac{\delta \ell}{\delta D}\right] \\ &= -\int dt\ \mathbf{w}\cdot\left[\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)\frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta b}\ \nabla\, b - D\ \nabla\frac{\delta \ell}{\delta D}\right], \end{aligned} \tag{10.5}$$

where we have consistently dropped boundary terms arising from integrations by parts, by invoking natural boundary conditions. Specifically, we may impose $\hat{\mathbf{n}} \cdot \mathbf{w} = 0$ on the boundary, where $\hat{\mathbf{n}}$ is the boundary's outward unit normal vector and $\mathbf{w} = \delta\eta_t \circ \eta_t^{-1}$ vanishes at the endpoints.

10.2 Euler–Poincaré framework for GFD

The Euler–Poincaré equations for continua (9.34) may now be summarized in vector form for advected Eulerian variables a in the set (10.1). We adopt the notational convention of the circulation map I in equations (9.39) and (9.40) that a one form density can be made into a one form (no longer a density) by dividing it by the mass density D and we use the Lie-derivative relation for the continuity equation $(\partial/\partial t + \pounds_{\mathbf{u}})Dd^3x = 0$. Then, the Euclidean components of the Euler–Poincaré equations for continua in equation (10.5) are expressed in Kelvin theorem form (9.41) with a slight abuse of notation as

$$\Big(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\Big)\Big(\frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}} \cdot d\mathbf{x}\Big) + \frac{1}{D}\frac{\delta \ell}{\delta b}\nabla b \cdot d\mathbf{x} - \nabla\Big(\frac{\delta \ell}{\delta D}\Big) \cdot d\mathbf{x} = 0\,, \tag{10.6}$$

in which the variational derivatives of the Lagrangian ℓ are to be computed according to the usual physical conventions, i.e., as Fréchet derivatives. Formula (10.6) is the Kelvin–Noether form of the equation of motion for ideal continua. Hence, we have the explicit Kelvin theorem expression, cf. equations (9.39) and (9.40),

$$\frac{d}{dt}\oint_{\gamma_t(\mathbf{u})} \frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}} \cdot d\mathbf{x} = -\oint_{\gamma_t(\mathbf{u})} \frac{1}{D}\frac{\delta \ell}{\delta b}\nabla b \cdot d\mathbf{x}\,, \tag{10.7}$$

where the curve $\gamma_t(\mathbf{u})$ moves with the fluid velocity $\mathbf{u}$. Then, by Stokes' theorem, the Euler equations generate circulation of $\mathbf{v} := (D^{-1}\delta l/\delta \mathbf{u})$ whenever the gradients ∇b and $\nabla(D^{-1}\delta l/\delta b)$ are not collinear. The corresponding ***conservation of potential vorticity*** q on fluid parcels is given by

$$\frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q = 0\,, \quad \text{where} \quad q = \frac{1}{D}\nabla b \cdot \mathrm{curl}\left(\frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}}\right). \tag{10.8}$$

This is also called ***PV convection***. Equations (10.6-10.8) embody most of the panoply of equations for GFD. The vector form of equation (10.6) is,

$$\underbrace{\Big(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\Big)\Big(\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}}\Big) + \frac{1}{D}\frac{\delta l}{\delta u^j}\nabla u^j}_{\textbf{Geodesic Nonlinearity: Kinetic energy}} \quad = \underbrace{\nabla\frac{\delta l}{\delta D} - \frac{1}{D}\frac{\delta l}{\delta b}\nabla b}_{\textbf{Potential energy}} \tag{10.9}$$

In geophysical applications, the Eulerian variable D represents the frozen-in volume element and b is the buoyancy. In this case, **Kelvin's theorem** is

$$\frac{dI}{dt} = \int\int_{S(t)} \nabla\left(\frac{1}{D}\frac{\delta l}{\delta b}\right) \times \nabla b \cdot d\mathbf{S}\,,$$

with circulation integral

$$I = \oint_{\gamma(t)} \frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} \cdot d\mathbf{x}\,.$$

10.3 Euler's Equations for a Rotating Stratified Ideal Incompressible Fluid

The Lagrangian. In the Eulerian velocity representation, we consider Hamilton's principle for fluid motion in a three dimensional domain with action functional $S = \int l\,dt$ and Lagrangian $l(\mathbf{u}, b, D)$ given by

$$l(\mathbf{u}, b, D) = \int \rho_0 D(1+b)\left(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - gz\right) - p(D-1)\,d^3x\,, \tag{10.10}$$

where $\rho_{tot} = \rho_0 D(1+b)$ is the total mass density, ρ_0 is a dimensional constant and $\mathbf{R}$ is a given function of $\mathbf{x}$. This variations at fixed $\mathbf{x}$ and t of this Lagrangian are the following,

$$\begin{aligned}
\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} &= \rho_0(1+b)(\mathbf{u}+\mathbf{R})\,, \quad \frac{\delta l}{\delta b} = \rho_0 D\left(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\right),\\
\frac{\delta l}{\delta D} &= \rho_0(1+b)\left(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\right) - p\,, \quad \frac{\delta l}{\delta p} = -\,(D-1)\,.
\end{aligned} \tag{10.11}$$

Hence, from the Euclidean component formula (10.9) for Hamilton principles of this type and the fundamental vector identity,

$$(\mathbf{b}\cdot\nabla)\mathbf{a} + a_j\nabla b^j = -\,\mathbf{b}\times(\nabla\times\mathbf{a}) + \nabla(\mathbf{b}\cdot\mathbf{a})\,, \tag{10.12}$$

we find the motion equation for an Euler fluid in three dimensions,

$$\frac{d\mathbf{u}}{dt} - \mathbf{u}\times\operatorname{curl}\mathbf{R} + g\hat{\mathbf{z}} + \frac{1}{\rho_0(1+b)}\nabla p = 0\,, \tag{10.13}$$

where $\operatorname{curl}\mathbf{R} = 2\boldsymbol{u}(\mathbf{x})$ is the Coriolis parameter (i.e., twice the local angular rotation frequency). In writing this equation, we have used advection of buoyancy,

$$\frac{\partial b}{\partial t} + \mathbf{u}\cdot\nabla b = 0,$$

from equation (10.2). The pressure p is determined by requiring preservation of the constraint $D = 1$, for which the continuity equation (10.3) implies $\operatorname{div}\mathbf{u} = 0$. The Euler motion equation (10.13) is Newton's Law for the acceleration of a fluid due to three forces: Coriolis, gravity and pressure gradient. The dynamic balances among these three forces produce the many circulatory flows of geophysical fluid dynamics. The ***conservation of potential vorticity*** q on fluid parcels for these Euler GFD flows is given by

$$\frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q = 0\,, \quad \text{where, on using } D = 1\,, \quad q = \nabla b \cdot \operatorname{curl}\big(\mathbf{u} + \mathbf{R}\big). \tag{10.14}$$

Semidirect-product Lie-Poisson bracket for compressible ideal fluids.

1. Compute the Legendre transform for the Lagrangian,

$$l(\mathbf{u}, b, D) : \mathfrak{X} \times \Lambda^0 \times \Lambda^3 \mapsto \mathbb{R}$$

whose advected variables satisfy the auxiliary equations,

$$\frac{\partial b}{\partial t} = -\,\mathbf{u}\cdot\nabla\, b\,, \qquad \frac{\partial D}{\partial t} = -\,\nabla\cdot(D\mathbf{u})\,.$$

2. Compute the Hamiltonian, assuming the Legendre transform is a linear invertible operator on the velocity $\mathbf{u}$. For definiteness in computing the Hamiltonian, assume the Lagrangian is given by

$$l(\mathbf{u}, b, D) = \int D\Big(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - e(D, b)\Big)\, d^3x\,, \tag{10.15}$$

with prescribed function $\mathbf{R}(\mathbf{x})$ and specific internal energy $e(D, b)$ satisfying the First Law of Thermodynamics,

$$de = \frac{p}{D^2}dD + T db\,,$$

where p is pressure, T temperature.

3. Find the semidirect-product Lie-Poisson bracket for the Hamiltonian formulation of these equations.
4. Does this Lie-Poisson bracket have Casimirs? If so, what are the corresponding symmetries and momentum maps?
5. Write the equations of motion and confirm their Kelvin-Noether circulation theorem.
6. Use the Kelvin-Noether circulation theorem for this theory to determine its potential vorticity and obtain the corresponding conservation laws. Write these conservation laws explicitly.

★

References

[AbMa1978] Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, 2nd ed. Reading, MA: Addison-Wesley.

[Ar1966] Arnold V.I. [1966], Sur la géometrie differentialle des groupes de Lie de dimiension infinie et ses applications à l'hydrodynamique des fluids parfaits. *Ann. Inst. Fourier, Grenoble* **16**, 319-361.

[Ar1979] Arnold V.I. [1979], *Mathematical Methods of Classical Mechanics*. Volume 60 of *Graduate Texts in Mathematics*. Springer.

[ArKh1998] Arnold, V. I and Khesin, B. A. [1998] *Topological Methods in Hydrodynamics*. Springer: New York.

[BlMaZe2009] Bloch, A. M., Marsden, J. E. and Zenkov, D. V. [2009] Quasivelocities and symmetries in non-holonomic systems, *Dynamical Systems*, **24**:2, 187–222.

[Bo1911] Bohlin, K. [1911] Note sur le problème des deux corps et sur une intégration nouvelle dans le problème des trois corps *Bulletin Astronomique*, Serie I, **28**, 113-119.

[CaHo1993] Camassa, R. and Holm, D. D. [1993] An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.* **71**, 1661-64.

[Clebsch1859] Clebsch, A. [1859] Ueber die Integration der hydrodynamischen Gleichungen. Journal für die reine und angewandte Mathematik, 56: 1–10.

[GBRa2008] Gay-Balmaz, F. and Ratiu, T.S. [2008] The geometric structure of complex fluids. *Adv. in Appl. Math.* **42** (2), 176–275.

[GBVi2012] Gay-Balmaz, F. and C. Vizman, [2012] Dual pairs in fluid dynam Ann Glob Anal Geom 41: 1–24.

[GeDo1979] Gelfand, I. M. and Dorfman, I. Ya. R. [1979] *Funct. Anal. Appl.* **13**, 248.

[GiHoKu1982] Gibbons, J., Holm, D. D. and Kupershmidt, B. A. [1982] Gauge-invariant Poisson brackets for chromohydrodynamics. *Phys. Lett. A* **90**, 281–283.

[Ha1904] Hamel, G. [1904] Die Lagrange–Eulersche Gleichungen der Mechanik, *Z. Math. Phys.* **50** 1–57.

[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.

[Ho2011GM1] Holm, D. D. [2011]
Geometric Mechanics I: Dynamics and Symmetry, Second edition, World Scientific: Imperial College Press, Singapore, .

[Ho2011GM2] Holm, D. D. [2011]
Geometric Mechanics II: Rotating, Translating & Rolling, Second edition, World Scientific: Imperial College Press, Singapore, .

[Ho2011] Holm, D. D. [2011]
Applications of Poisson geometry to physical problems, *Geometry & Topology Monographs* **17**, 221–384.

[HoKu1988] Holm, D. D. and Kupershmidt, B. A. [1988] The analogy between spin glasses and Yang–Mills fluids. *J. Math Phys.* **29**, 21–30.

[HoMaRa1998a] Holm, D. D., Marsden, J. E. and Ratiu, T. S. [1998a] The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. in Math.* **137**, 1–81.

[HoScSt2009] Holm, D. D., Schmah, T. and Stoica, C. [2009]
Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions,
Oxford University Press.

[HoVi2012] Holm, D. D. and Vizman, C. [2012] Dual pairs in resonances, *J. Geom. Mech.* **4** (3) 297-311.

[Ma1976] Manakov, S. V. [1976]
Note on the integration of Euler's equations for the dynamics of an n-dimensional rigid body.
Funct. Anal. Appl. **10**, 328–329.

[Marle2013] Marle, C.-M. [2013] On Henri Poincarés note: sur une forme nouvelle des equations de la mécanique. *J. Geom. Symmetry Phys.* 29: 1–38; arXiv:1302.4096v1 [math].

[MaRa1999] J.E. Marsden, T.S. Ratiu [1999] *Introduction to Mechanics and Symmetry*, second edition, Texts Appl. Math., **17**, Springer-Verlag, New York.

[Ol2000] P. J. Olver [2000] *Applications of Lie Groups to Differential Equations*. Springer: New York.

[Po1901] H. Poincaré, Sur une forme nouvelle des équations de la méchanique,
C.R. Acad. Sci. **132** (1901) 369-371.
English translation in [Ho2011GM2], Appendix D.

[RaTuSbSoTe2005] Ratiu, T. S., Tudoran, R., Sbano, L., Sousa Dias, E. and Terra, G. [2005]
A crash course in geometric mechanics.
In *Geometric Mechanics and Symmetry: The Peyresq Lectures*,
edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306.
Cambridge: Cambridge University Press.