M3-4-5 A34 Lecture Notes: Dynamics, Symmetry and Integrability

Professor Darryl D Holm   Spring Term 2018
Imperial College London   d.holm@ic.ac.uk
`http://www.ma.ic.ac.uk/~dholm/`

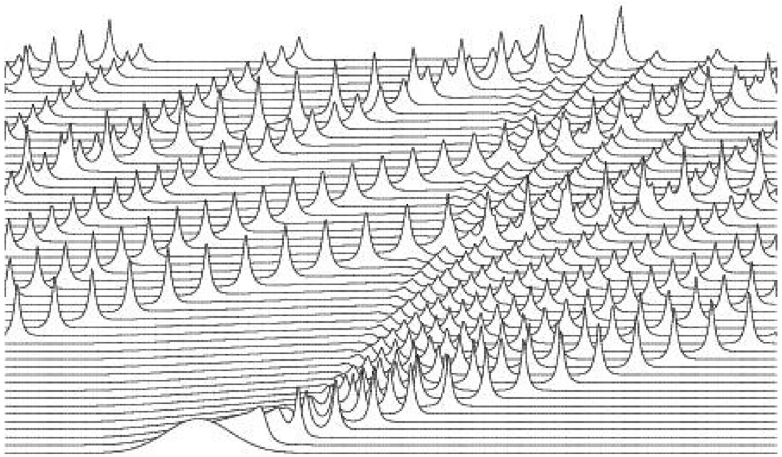
Course meets $\boxed{\text{11am-12pm Mon @ Hux 408 \& 1-3pm Fri @ Hux 658}}$

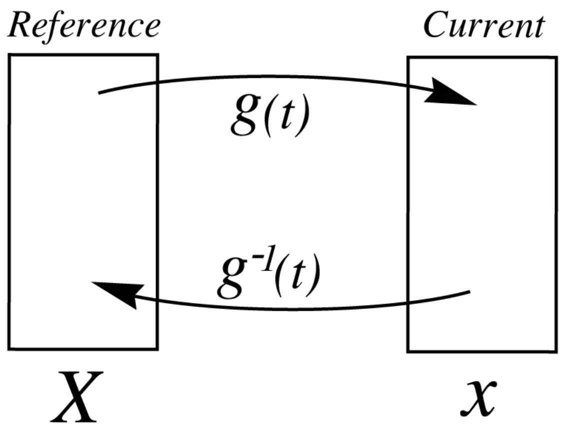
Three assessed homeworks at 3-week intervals, Exam taken mainly from these.

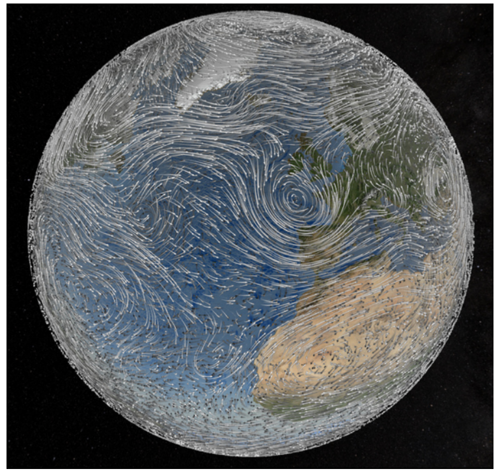
References:

[HoSmSt2009] *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, $2nd$ 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 calculus 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 on Lie groups. The main topics are shallow water waves, ideal incompressible fluid dynamics and geophysical fluid dynamics (GFD).

Three worked examples that illustrate the course material in simpler settings are given in full detail in the course notes. These will be assigned as outside reading and then discussed in Q&A sessions in class.

Lecture 1, Friday 12 Jan 2018: Introduction to the Course

1 What is Geometric Mechanics?

1.1 Geometric Mechanics is a framework for understanding dynamics

$$
\begin{tikzcd}
 & & L : TM \to \mathbb{R} \arrow[rr, leftrightarrow, "\textbf{LT}"] \arrow[dl, "\text{LagHvp}"] \arrow[dd, "TM/G"] \arrow[ddll, dash, bend right] & & H : T^*M \to \mathbb{R} \arrow[dl, "\text{HamHvp}"] \arrow[dd, "\mathbf{J}"] \arrow[ddr, dash] & \\
 & \text{EL eqns} \arrow[rr, Leftrightarrow] \arrow[dd, "TM/G"] & & \text{Ham eqns} \arrow[dd, "\mathbf{J}"] & & \\
\text{Noether's Theorem} \arrow[rr, dash] & & \ell : \mathfrak{g} \to \mathbb{R} \arrow[rr, leftrightarrow, "\textbf{LT}"] \arrow[dl, "\text{EPvp}"] & & h : \mathfrak{g}^* \to \mathbb{R} \arrow[r, dash] \arrow[dl, "\text{LPvp}"] & \text{Noether's Theorem} \\
 & \text{EP eqns} \arrow[rr, Leftrightarrow] & & \text{LP eqns} & &
\end{tikzcd}
$$

Figure 1: Diagram for Geometric Mechanics: Much of the remaining lecture will involve parsing this diagram.

2 What is Geometric Mechanics?

- Geometric mechanicswas introduced in Poincaré [1901], in a 2-page paper!
  - GM is a powerful formalism for understanding dynamical systems whose Lagrangian and Hamiltonian are invariant under the transformations of the configuration manifold M by a **Lie group** G. That is, :$G \times M \to M$
  - **Examples of its applications** range from the simple finite dimensional dynamics of the freely rotating rigid body to the infinite dimensional dynamics of the ideal fluid equations.
  - For a historical review and basic references, see, e.g., Holm, Marsden, Ratiu [1998] *Adv in Maths*, as well as various textbook introductions to geometric mechanics and background references.
- One of the **main approaches** of geometric mechanics is the **method of reduction** of the motion equations of a mechanical system by a Lie group symmetry, G, in either its Lagrangian formulation on the tangent space TM of a configuration manifold M, or its Hamiltonian formulation on the cotangent space T^*M.
- The **method of reduction by symmetry**
  - yields reduced Lagrangian and Hamiltonian formulations of the Euler-Poincaré equations governing the dynamics of the **momentum map** $J : T^*M \to \mathfrak{g}^*$, where $\mathfrak{g}^*$ is the dual Lie algebra of the Lie symmetry group G.
- In general terms, Lie group reduction by symmetry simplifies the motion equations of a mechanical system with symmetry by transforming them into new dynamical variables in $\mathfrak{g}^*$ which are invariant under the same Lie group symmetries as the Lagrangian and Hamiltonian of the dynamics.
- More specifically, **on the Lagrangian side**, the new invariant variables under the Lie symmetries are obtained from **Noether's theorem**, via the **tangent lift** of the infinitesimal action of the Lie symmetry group on the configuration manifold.
- The unreduced Euler–Lagrange equations are replaced by equivalent **Euler-Poincaré equations** expressed in the new invariant variables in $\mathfrak{g}^*$, plus an auxiliary **reconstruction equation**, which restores the information in the tangent space of the configuration space lost in transforming to group invariant dynamical variables.
- **On the Hamiltonian side**, after a **Legendre transformation**, equivalent new invariant variables in $\mathfrak{g}^*$ are defined by the momentum map $J : T^*M \to \mathfrak{g}^*$ from the phase space T^*M of the original system on the configuration manifold M to the dual $\mathfrak{g}^*$ of the Lie symmetry algebra $\mathfrak{g} \simeq T_eG$, via the cotangent lift of the infinitesimal action of the Lie symmetry group on the configuration manifold.

- The **cotangent lift momentum map** is an equivariant Poisson map which reformulates the canonical Hamiltonian flow equations in phase space as noncanonical Lie-Poisson equations governing flow of the momentum map on an orbit of the coadjoint action of the Lie symmetry group on the dual of its Lie algebra $\mathfrak{g}^*$, plus an auxiliary reconstruction equation for lifting the Lie group reduced coadjoint motion back to phase space T^*M.
- Thus, **Lie symmetry reduction** yields coadjoint motion of the corresponding momentum map.
- The **dimension of the dynamical system reduces**, because its solutions are restricted to remain on certain subspaces of the original phase space, called coadjoint orbits.
- These subspaces are **coadjoint orbits** of the action of the group G on $\mathfrak{g}^*$, the dual space of its Lie algebra $\mathfrak{g}$, with respect to a certain pairing.

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

- Coadjoint orbits lie on level sets of the distinguished smooth functions $C \in \mathcal{F} : \mathfrak{g}^* \to \mathbb{R}$ of the symmetry-reduced dual Lie algebra variables $\mu \in \mathfrak{g}^*$ called **Casimir functions**.
- The Casimir functions are **conserved quantities**. Indeed, Casimir functions have null Lie-Poisson brackets $\{C, F\}(\mu) = 0$ with any other functions $F \in \mathcal{F}(\mathfrak{g}^*)$, including the reduced Hamiltonian $h(\mu)$.
- Furthermore, level sets of the Casimirs, on which the coadjoint orbits lie, are **symplectic manifolds** which provide the framework on which geometric mechanics is constructed.
- These symplectic manifolds have many applications in physics, as well as in symplectic geometry, whenever Lie symmetries are present.
- In particular, coadjoint motion of the momentum map $J(t) = \mathrm{Ad}^*_{g(t)} J(0)$ for a solution curve $g(t) \in C(G)$ takes place on the intersections of level sets of the Casimirs with level sets of the Hamiltonian.

• What is Complete Integrability?

In the Geometric Mechanics framework, Integrability solves Hamiltonian dynamical systems, by using Reduction by Symmetry, Noether's Theorem and Momentum Maps. These methods apply well for evolutionary ordinary and partial differential equations (i.e., for both ODEs and PDEs). In our lectures on Integrability, we will focus on instructive, illustrative examples. In particular, we will study integrability of ODEs for rigid body motion and of PDEs for nonlinear shallow water waves, such as the KdV equation. Both of these famous classes of integrable systems represent geodesic motions on Lie groups. Following Poincaré [1901], we will use transformation theory to represent motions on a given configuration manifold as time-dependent curves on a Lie group that acts transitively on that configuration manifold. For example, the rotation of the rigid body is lifted to $SO(3)$ and the propagation of nonlinear waves in one dimension is lifted to the Lie group, Diff($\mathbb{R}$), of smooth invertible transformation of the real line.

Contents

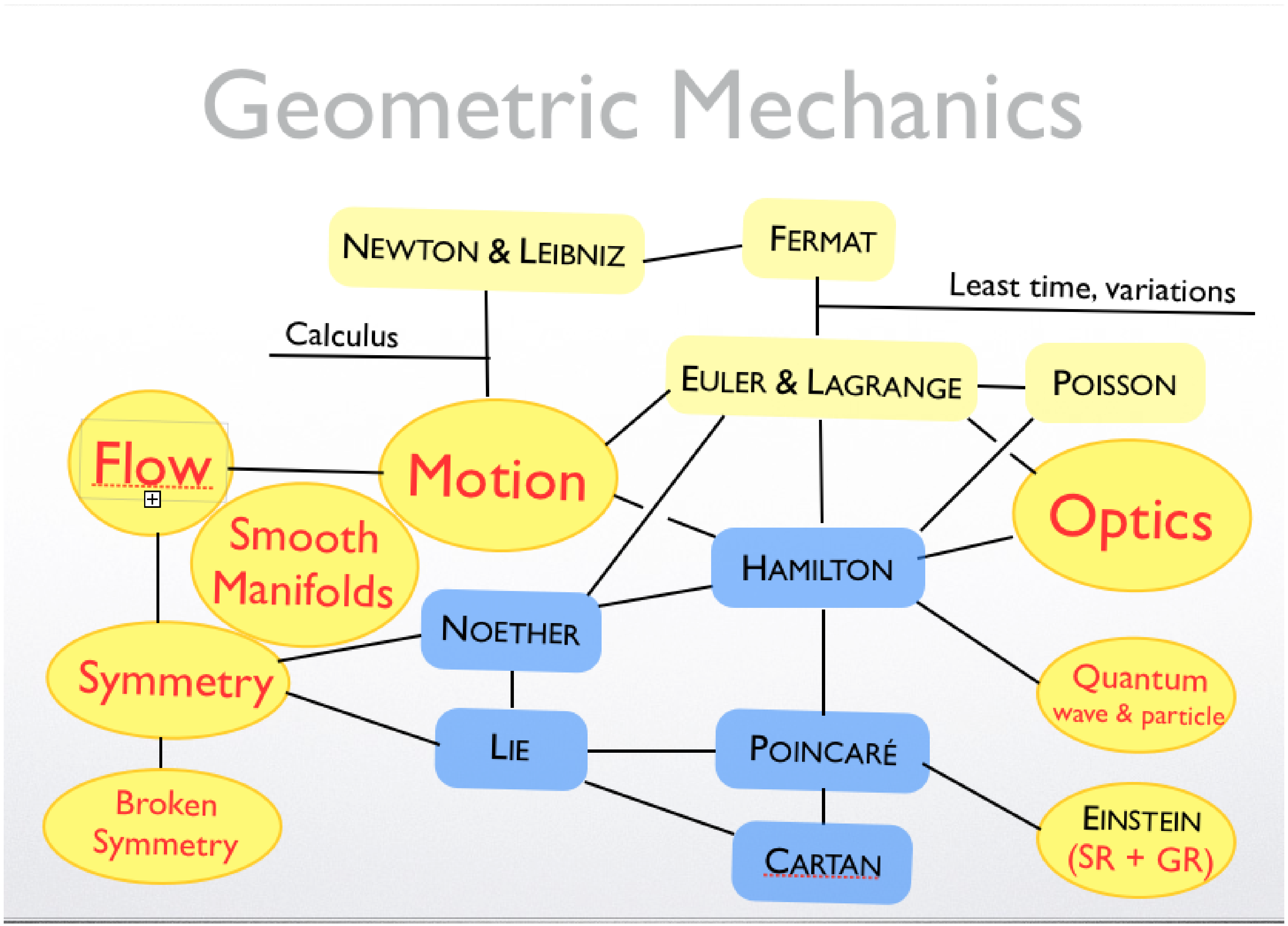


Figure 2: Geometric Mechanics has involved many great mathematicians!

3 Introduction

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

Background reading: Chapter 2, [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 $G = SE(3) \simeq SO(3)Ⓢ\mathbb{R}^3$ 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,

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

leads to the *Euler-Lagrange equations*

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

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.

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 2 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.

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

4.1 Poincaré's work in 1901 was based on earlier work of Lie in 1870's

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* is a set of elements with an associative binary product that has a unique inverse and identity element.
- 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, so it is an interesting arena for geometric mechanics.
- Choose the manifold M for mechanics as discussed above to be the Lie group G and denote the *identity element* as the point e. The identity element e satisfies $eg=g=ge$ for all $g\in G$, where the group product denoted by concatenation.
- 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 $a,b,c\in\mathfrak{g}$,

  $$[a,[b,c]]+[b,[c,a]]+[c,[a,b]]=0\,.$$

  The bracket operation among the *basis vectors* $e_k\in\mathfrak{g}$ with $k=1,2,\dots,\dim(\mathfrak{g})$ 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 requirement of skew-symmetry and the Jacobi condition put constraints on the structure constants. These constraints are

  – skew-symmetry

  $$c^k_{ji}=-\,c^k_{ij}\,, \tag{4.1}$$

– Jacobi identity

$$c^k_{ij}c^m_{lk} + c^k_{li}c^m_{jk} + c^k_{jl}c^m_{ik} = 0\,. \tag{4.2}$$

Conversely, any set of constants c^k_{ij} that satisfy relations (4.1)–(4.2) defines a Lie algebra $\mathfrak{g}$.

Exercise. Prove that the Jacobi identity requires the relation (4.2).
Hint: the Jacobi identity involves summing three terms of the form

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

★

4.2 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.

4.2.1 ADjoint, Adjoint and adjoint 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**,

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

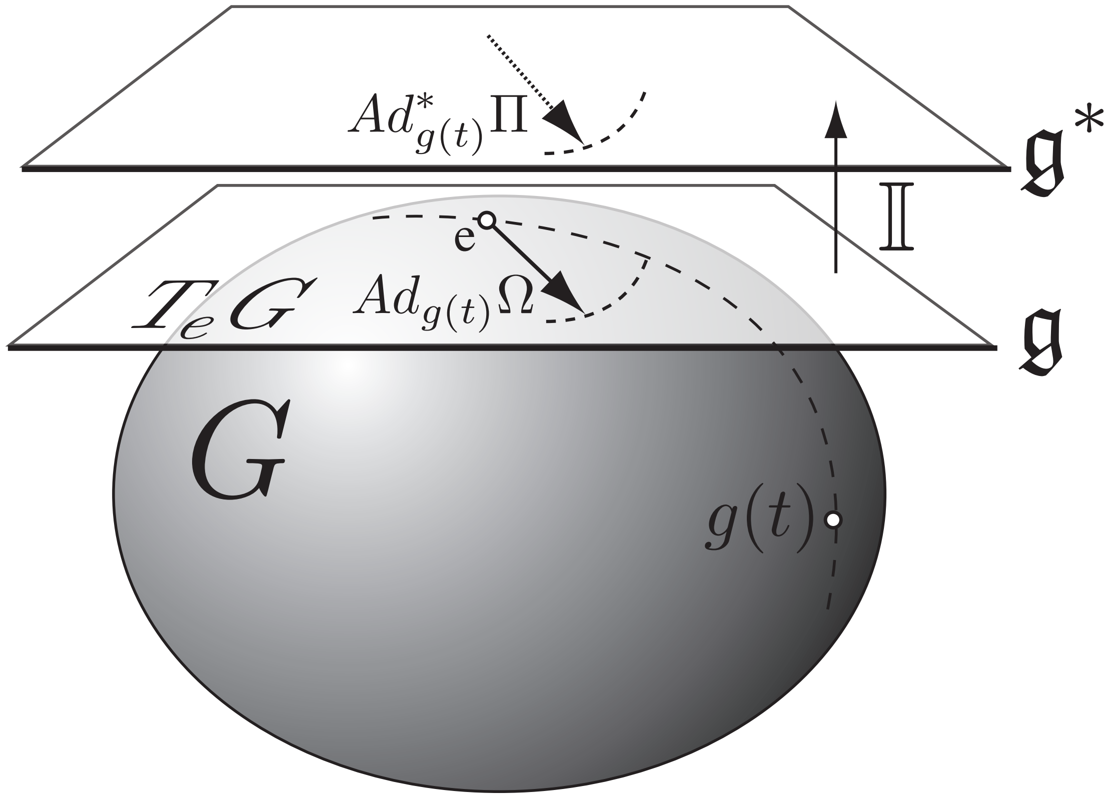
  Explicitly, one computes the ad operation by differentiating the Ad operation directly as

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{g(t)} \, \eta &= \frac{d}{dt}\Big|_{t=0} \Big(g(t)\eta g^{-1}(t) \Big) \\ &= \dot{g}(0)\eta g^{-1}(0) - g(0)\eta g^{-1}(0)\dot{g}(0)g^{-1}(0) \\ &= \xi \, \eta - \eta \, \xi = [\xi, \eta] = \mathrm{ad}_\xi \, \eta , \end{aligned} \tag{4.3}$$

  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

4.1 (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

4.2 (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 (4.3), for a curve $\eta(t) \in \mathfrak{g}$,

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=t_0} \mathrm{Ad}_{g(t)}\, \eta(t) &= \frac{d}{dt}\Big|_{t=t_0}\Big(g(t)\eta(t)g^{-1}(t)\Big) \\
&= g(t_0)\Big(\dot{\eta}(t_0) + g^{-1}(t_0)\dot{g}(t_0)\eta(t_0) \\
&\quad - \eta(t_0)g^{-1}(t_0)\dot{g}(t_0)\Big)g^{-1}(t_0) \\
&= \left[\mathrm{Ad}_{g(t)}\left(\frac{d\eta}{dt} + \mathrm{ad}_\xi\eta\right)\right]_{t=t_0}\,. \qquad (4.4)
\end{aligned}$$

□

Exercise. (Inverse Adjoint motion relation) Verify that

$$\frac{d}{dt}\mathsf{Ad}_{g(t)^{-1}}\eta = -\mathsf{ad}_\xi \mathsf{Ad}_{g(t)^{-1}}\eta\,, \tag{4.5}$$

for any fixed $\eta \in \mathfrak{g}$. Note the placement of $\mathsf{Ad}_{g(t)^{-1}}$ and compare with Exercise on page 17. ★

4.2.2 Compute the coAdjoint and coadjoint operations by taking duals

The pairing

$$\langle\,\cdot\,,\,\cdot\,\rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R} \tag{4.6}$$

(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,

  $$\mathrm{Ad}^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* : \quad \langle\, \mathrm{Ad}^*_g\,\mu\,,\,\eta\,\rangle := \langle\,\mu\,,\,\mathrm{Ad}_g\,\eta\,\rangle\,, \tag{4.7}$$

  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,

  $$\mathrm{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g}^* : \quad \langle\, \mathrm{ad}^*_\xi\,\mu\,,\,\eta\,\rangle := \langle\,\mu\,,\,\mathrm{ad}_\xi\,\eta\,\rangle\,, \tag{4.8}$$

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

Definition

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

$$\Phi^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* \quad \textit{given by} \quad (g,\mu) \mapsto \mathrm{Ad}^*_{g^{-1}}\mu \tag{4.9}$$

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

Remark

4.4 (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 4.3 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

$$\mathrm{Ad}^*_{g^{-1}} \mathrm{Ad}^*_{h^{-1}} = \mathrm{Ad}^*_{h^{-1}g^{-1}} = \mathrm{Ad}^*_{(gh)^{-1}} \,. \tag{4.10}$$

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

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

Proposition

4.5 (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*

$$\frac{d}{dt} Ad^*_{g(t)^{-1}} \mu(t) = Ad^*_{g(t)^{-1}} \left[\frac{d\mu}{dt} - ad^*_{\xi(t)} \mu(t) \right] , \tag{4.11}$$

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

Proof. The Exercise on page 18 introduces the inverse Adjoint motion relation (4.5) 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 (4.5) may be proven by the following computation,

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

in which for the last step one recalls

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

Relation (4.5) plays a key role in demonstrating relation (4.11) 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

$$\begin{aligned}
\left\langle \frac{d}{dt}\mathrm{Ad}^*_{g(t)^{-1}}\mu(t),\eta\right\rangle &= \frac{d}{dt}\left\langle \mathrm{Ad}^*_{g(t)^{-1}}\mu(t),\eta\right\rangle\\
\text{by (4.7)}\quad &= \frac{d}{dt}\left\langle \mu(t),\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle\\
&= \left\langle \frac{d\mu}{dt},\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle + \left\langle \mu(t),\frac{d}{dt}\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle\\
\text{by (4.5)}\quad &= \left\langle \frac{d\mu}{dt},\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle + \left\langle \mu(t),-\mathrm{ad}_{\xi(t)}\left(\mathrm{Ad}_{g(t)^{-1}}\eta\right)\right\rangle\\
\text{by (4.8)}\quad &= \left\langle \frac{d\mu}{dt},\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle - \left\langle \mathrm{ad}^*_{\xi(t)}\mu(t),\mathrm{Ad}_{g(t)^{-1}}\eta\right\rangle\\
\text{by (4.7)}\quad &= \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\frac{d\mu}{dt},\eta\right\rangle - \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\mathrm{ad}^*_{\xi(t)}\mu(t),\eta\right\rangle\\
&= \left\langle \mathrm{Ad}^*_{g(t)^{-1}}\left[\frac{d\mu}{dt}-\mathrm{ad}^*_{\xi(t)}\mu(t)\right],\eta\right\rangle.
\end{aligned}$$

This concludes the proof.

$\square$

Corollary

4.6. *The* ***coadjoint orbit relation***

$$\mu(t) = Ad^*_{g(t)}\mu(0) \tag{4.12}$$

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

$$\frac{d\mu}{dt} - ad^*_{\xi(t)}\mu(t) = 0\,. \tag{4.13}$$

Proof. Substituting Equation (4.13) into Equation (4.11) yields

$$\mathrm{Ad}^*_{g(t)^{-1}}\mu(t) = \mu(0)\,. \tag{4.14}$$

Operating on this equation with $\mathrm{Ad}^*_{g(t)}$ and recalling the composition rule for Ad^* from Remark 4.4 yields the result (4.12). □

Remark

4.7. *As it turns out, the equations in Poincaré (1901) for which we have been preparing describe coadjoint motion! Moreover, by equation* (4.14) *in the proof, coadjoint motion implies that* $\mathrm{Ad}^*_{g(t)^{-1}}\mu(t)$ *is a conserved quantity.*

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.

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

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

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

4.8. *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,*

$$\textit{by identifying the composition} \quad g.q = \frac{aq+b}{cq+d} \quad \textit{with the } SL(2,\mathbb{R}) \textit{ matrices} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \begin{bmatrix} 1 & \epsilon_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} e^{-\epsilon_2/2} & 0 \\ 0 & e^{\epsilon_2/2} \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \epsilon_3 & 1 \end{bmatrix}. \tag{4.15}$$

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

$$X_\alpha(q)\partial_q X_\beta(q) - X_\beta(q)\partial_q X_\alpha(q) =: \big[X_\alpha(q)\partial_q\,,\, X_\beta(q)\partial_q\big] =: [v_\alpha,\, v_\beta] = \begin{array}{|c|ccc|} \hline [\,\cdot\,,\,\cdot\,] & v_1 & v_2 & v_3 \\ \hline v_1 & 0 & -v_1 & 2v_2 \\ v_2 & v_1 & 0 & -v_3 \\ v_3 & -2v_2 & v_3 & 0 \\ \hline \end{array} \tag{4.16}$$

or, in index notation,

$$[v_\alpha\,,\, v_\beta]^i = v^j_\alpha \frac{\partial v^i_\beta}{\partial q^j} - v^j_\beta \frac{\partial v^i_\alpha}{\partial q^j} = c^\gamma_{\alpha\beta} v^i_\gamma\,, \tag{4.17}$$

or, upon suppressing Latin indices

$$\big[v_\alpha\,,\, v_\beta\big] = v_\alpha \frac{\partial v_\beta}{\partial q} - v_\beta \frac{\partial v_\alpha}{\partial q} = c^\gamma_{\alpha\beta} v_\gamma\,, \tag{4.18}$$

with

$$c^1_{12} = c^3_{23} = -1 = -c^1_{21} = -c^3_{32}, \qquad c^2_{13} = 2 = -c^2_{31}\,, \tag{4.19}$$

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

Anti-homomorphism. We note that *minus* the same commutator table (4.16) 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 (4.15),

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

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

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

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

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

1. Define a **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 **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 **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 $k=1,2,\dots,\dim(\mathfrak{g})$.

- **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

$$\delta\xi=\dot\eta+[\xi,\eta]\quad\text{with}\quad \xi,\eta\in\mathfrak{g}\,. \tag{4.20}$$

  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

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

- **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,

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

On taking the difference, the terms with cross derivatives cancel and one finds the variational formula (4.20),

$$\xi\,' - \dot\eta = [\,\xi\,,\,\eta\,] \quad \text{with} \quad [\,\xi\,,\,\eta\,] := \xi\,\eta - \eta\,\xi = \mathrm{ad}_\xi\,\eta\,. \tag{4.21}$$

(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

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

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 $\Big\langle - \mathrm{ad}^*_\xi \frac{\partial l}{\partial \xi},\, \eta \Big\rangle := \frac{\partial l}{\partial \xi^k} \xi^j c_{ij}{}^k \eta^i = \frac{\partial l}{\partial \xi^k} [e_i \eta^i,\, e_j \xi^j]^k =: \Big\langle \frac{\partial l}{\partial \xi},\, - \mathrm{ad}_\xi \eta \Big\rangle$.

- **Exercise:** Write Noether's theorem for the Euler-Poincaré theory.
- **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 $l(\xi)$.

- **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 [1]

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

  (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.

[1] (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 $\epsilon_{ij}{}^k$.

4.4 Euler-Poincaré variational principle for the rigid body

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

$$\mathbb{I}\dot{\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega} \times \boldsymbol{\Omega}\,, \tag{4.22}$$

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

Definition

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

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

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

4.10 (Euler's rigid-body equations)**.** *Euler's rigid-body equations (4.22) arise from Hamilton's principle,*

$$\delta S(\boldsymbol{\Omega}) = \delta \int_a^b l(\boldsymbol{\Omega})\, dt = 0, \tag{4.23}$$

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

$$l(\boldsymbol{\Omega}) = \frac{1}{2}\langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle = \frac{1}{2}\,\mathbb{I}\boldsymbol{\Omega}\cdot\boldsymbol{\Omega} = \frac{1}{2}\,(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)\,, \tag{4.24}$$

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

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega}\times\boldsymbol{\Xi}\,, \tag{4.25}$$

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

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

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

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

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

Since $\boldsymbol{\Xi}$ is otherwise arbitrary, (4.23) is equivalent to

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

which recovers Euler's equations (4.22) in vector form. □

Proposition

4.11 (Derivation of the restricted variation)**.** *The restricted variation in (4.25) 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, $O^{\dot{}\,\prime} = d^2O/dtds = O^{\prime\,\dot{}}$.

(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 (4.25),

$$
\widehat{\Omega}\,' - \widehat{\Xi}^{\,\cdot} = \Big[\widehat{\Omega}\,,\,\widehat{\Xi}\Big] \quad \text{with} \quad [\,\widehat{\Omega}\,,\,\widehat{\Xi}\,] := \widehat{\Omega}\widehat{\Xi} - \widehat{\Xi}\widehat{\Omega}\,. \tag{4.26}
$$

Under the bracket relation

$$
[\,\widehat{\Omega}\,,\,\widehat{\Xi}\,] = (\boldsymbol{\Omega}\times\boldsymbol{\Xi})\,\widehat{}
\tag{4.27}
$$

for the hat map, this equation recovers the vector relation (4.25) in the form

$$
\boldsymbol{\Omega}\,' - \dot{\boldsymbol{\Xi}} = \boldsymbol{\Omega}\times\boldsymbol{\Xi}\,. \tag{4.28}
$$

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

$$
\mathbb{I}\dot{\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega}\times\boldsymbol{\Omega}\,, \tag{4.29}
$$

follow from the variational principle (4.23) with variations of the form (4.25) derived from the definition of body angular velocity $\widehat{\Omega}$.

$\square$

Exercise. What conservation law does Noether's theorem imply for the rigid-body equations (4.22). Hint, is the Lagrangian in (4.24) invariant under rotations around Ξ? ★

4.5 Clebsch variational principle for the rigid body

Proposition

4.12 (Clebsch variational principle)**.**
The Euler rigid-body equations on $T\mathbb{R}^3$ *given in equation* (4.22) *as*

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

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

$$\delta S(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P}) = \delta \int_a^b l(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}; \mathbf{P})\, dt = 0, \tag{4.30}$$

for a ***constrained action integral***

$$\begin{aligned} S(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}}) &= \int_a^b l(\mathbf{\Omega}, \mathbf{Q}, \dot{\mathbf{Q}})\, dt \\ &= \int_a^b \frac{1}{2}\mathbf{\Omega}\cdot\mathbb{I}\mathbf{\Omega} + \mathbf{P}\cdot\Big(\dot{\mathbf{Q}} + \mathbf{\Omega}\times\mathbf{Q}\Big)\, dt\,. \end{aligned} \tag{4.31}$$

Remark

4.13 (Reconstruction as constraint)**.**

- *The first term in the Lagrangian (4.31),*
$$l(\mathbf{\Omega}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2) = \frac{1}{2}\mathbf{\Omega}^T\mathbb{I}\mathbf{\Omega}\,, \tag{4.32}$$
is the (rotational) kinetic energy of the rigid body.

- *The second term in the Lagrangian (4.31) 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

$$
\begin{aligned}
\dot{\mathbf{Q}}(t) &= -\,(O^{-1}\dot{O})O^{-1}(t)\mathbf{Q}(0) \\
&= -\,\hat{\Omega}(t)\mathbf{Q}(t) = -\,\boldsymbol{\Omega}(t)\times\mathbf{Q}(t)\,.
\end{aligned}
\tag{4.33}
$$

Proof. The variations of S are given by

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

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

$$
\mathbb{I}\boldsymbol{\Omega} = \mathbf{P}\times\mathbf{Q}\,,\quad \dot{\mathbf{Q}} = -\,\boldsymbol{\Omega}\times\mathbf{Q}\,,\quad \dot{\mathbf{P}} = -\,\boldsymbol{\Omega}\times\mathbf{P}\,.
\tag{4.34}
$$

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

4.14. *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

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

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 (4.34) are

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

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

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

Consequently, the derivative property of the canonical Poisson bracket yields

$$\{f, h\}(\mathbf{\Pi}) = \frac{\partial f}{\partial \Pi_a}\{\Pi_a, \Pi_i\}\frac{\partial h}{\partial \Pi_b} = -\,\epsilon_{abc}\Pi_c \frac{\partial f}{\partial \Pi_a}\frac{\partial h}{\partial \Pi_b} = -\mathbf{\Pi}\cdot\frac{\partial f}{\partial \mathbf{\Pi}} \times \frac{\partial h}{\partial \mathbf{\Pi}}\,, \tag{4.35}$$

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

4.15. *This exercise proves that the map $T^*\mathbb{R}^3 \to \mathbb{R}^3$ given by $\mathbf{\Pi} = \mathbf{P} \times \mathbf{Q} \in \mathbb{R}^3$ in (4.34) 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

$$\frac{dM}{dt} = [\, M \,,\, \Omega \,] \,, \tag{4.36}$$

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

$$S(\Omega, Q, P) = \int_a^b l(\Omega) + \mathrm{tr}\Big(P^T\left(\dot{Q} - Q\Omega\right)\Big)\, dt \,, \tag{4.37}$$

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 (4.34),

$$\dot{Q} = Q\Omega \quad \text{and} \quad \dot{P} = P\Omega \,, \tag{4.38}$$

as a result of the constraints.

(b) How does equation (4.36) for the $SO(n)$ rigid body dynamics change, if the Lagrangian $l(\Omega)$ in (4.37) 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

$$
\frac{dM}{dt} = [\, M \,,\, \Omega\,] + \frac{1}{2}\left(Q^T \frac{\partial l}{\partial Q} - \frac{\partial l}{\partial Q}^T Q \right) , \qquad \frac{dQ}{dt} = Q\Omega \tag{4.39}
$$

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 $SE(3)$, this Lie-Poisson bracket becomes, via the hat map,

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

▲

5 Integrability of 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

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

for $\delta\Omega$ now leads to

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

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

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

Finally, invoking stationarity for arbitrary Ξ implies the commutator form (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].

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 matrix Lie group G.

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

Remark

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

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

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 $so(3) \times so(3)$. ★

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 (4.12). ★

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 $\langle \frac{\delta l}{\delta\xi}\,,\,\eta\rangle|_a^b$ 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. □

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*

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

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,

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

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

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

The constancy of this quantity under the Euler–Poincaré dynamics in (6.2) is verified, upon taking the time derivative and using the coadjoint motion relation (4.11) in Proposition 4.5. □

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 ($g^{\cdot\,\prime} = d^2g/dtds = g\,'^{\,\cdot}$).

Namely, one computes,

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

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). $\square$

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 $\mathfrak{g}^*\times V$,

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

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

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

in the action (6.16) to the Hamiltonian,

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

whose variations are given by

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

These variational derivatives recover Equations (6.17) in canonical Hamiltonian form,

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

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 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.

7.1 Wave equations

Wave equations are evolutionary equations for time dependent curves in a space of smooth maps $C^\infty(\mathbb{R}^n, V)$ for solutions, $u \in V$, a vector space V.

$$\partial_t u = f(u)\,, \quad \text{or} \quad \partial_t u_i(\mathbf{x},t) = f_i(u_i,\, u_{i,j},\, u_{i,jk},\, u_{i,jkl},\, \dots)\,. \tag{7.1}$$

Typically, V is $\mathbb{R}$ or $\mathbb{C}$, $n = 1$. We are interested in the Cauchy problem. Namely, solve (7.1) for $u(x,t)$, given the initial condition $u(x,0)$ and boundary conditions $u(x|_{\partial D}, t)$.

Travelling waves. The simplest wave solution is called a travelling wave. This solution is a function u of the form

$$u(x,t) = F(x - ct)\,,$$

where $F : \mathbb{R} \to V$ is a function defining the wave shape, and c is a real number defining the propagation speed of the wave. Thus, travelling waves preserve their shape and simply translate to the right at constant speed, c.

Exercise. Find the travelling wave solutions of $u_t + cu_x = 0$ and $u_{tt} - c^2 u_{xx}$. Hint: notice that $(\partial_t - c\partial_x)(\partial_t + c\partial_x) = (\partial_t^2 - c^2\partial_x^2)$. Upon introducing independent variables $\xi = x - ct$ and $\eta = x + ct$ the second equation becomes $u_{\xi\eta} = 0$ so that $u(\xi,\eta) = F(\xi) + G(\eta)$. ★

Plane waves. A complex-valued travelling wave, called a plane wave, plays a fundamental role in the theory of linear wave equations. The general form of a plane wave is

$$u(x,t) = \Re e(Ae^{i(kx - \omega t)},$$

where $|A|$ is the wave amplitude, k is wave number, ω is wave frequency, and $c_p = \omega/k$ is the speed along the oscillating wave form.

Exercise. Find dispersion relations $\omega = \omega(k)$ and phase velocities $c_p(k) := \omega(k)/k$ for the plane wave solutions of $u_t + cu_x = 0$, $u_{tt} - c^2 u_{xx} = 0$, $u_t + \gamma u_{xxx} = 0$, and $u_t = -\partial_x P(\partial_x^2) u$, where $P(\,\cdot\,)$ is a real polynomial of its argument.

Why are these called dispersion relations? (Hint: consider the initial Fourier k-spectrum.) How do they differ? What is the importance of the relative signs? ★

7.2 Conservation laws

Conservation laws for evolutionary equations of the form $u_t = f(u)$ satisfy

$$\frac{d}{dt}\int F(u)\,dx = \int \frac{\delta F}{\delta u} u_t\,dx = \int \frac{\delta F}{\delta u} f(u)\,dx = \int dG(u) = 0\,,$$

for some functions F and G of u and its derivatives, and for suitable boundary conditions.

For example, the inviscid Burgers equation

$$u_t + uu_x = 0\,, \tag{7.2}$$

has an infinite number of conservation laws, given by

$$\frac{d}{dt}\int \frac{u^n}{n}\,dx = \int u^{n-1}u_t\,dx = -\int u^n u_x\,dx = -\int \frac{1}{n+1}\partial_x u^{n+1}\,dx = -\frac{1}{n+1}\int d(u^{n+1}) = 0\,, \tag{7.3}$$

for homogeneous boundary conditions and any integer n.

Even so, the solutions of the inviscid Burgers equation carry the seeds of their own destruction, since they exhibit wave breaking in finite time. That is, they develop negative vertical slope in finite time. This is shown in the proof of the following Lemma.

Lemma

7.1 (Steepening Lemma for the inviscid Burgers equation)**.**
Suppose the initial profile of velocity $u(0,x)$ for the inviscid Burgers equation (7.2) has an inflection point of negative slope $u_x(0,\overline{x}(0)) < 0$ located at $x = \overline{x}(0)$ to the right of its maximum, and otherwise it decays to zero in each direction sufficiently rapidly for all of its conservation laws in equation (7.3) 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, defined by $s(t) = u_x(\overline{x}(t),t)$. Then the inviscid Burgers equation (7.2) yields an evolution equation for the slope, $s(t)$. Namely, using $u_{xx}(\overline{x}(t),t) = 0$ the spatial derivative of equation (7.2) leads to

$$\frac{ds}{dt} = -\,s^2 \quad \Longrightarrow \quad s(t) = \frac{s(0)}{1+s(0)t}\,. \tag{7.4}$$

Thus, if $s(0) < 0$, the slope at the inflection point $s(t)$ will become increasingly more negative, until it becomes vertical at time $t = -1/s(0)$. □

7.3 Survey of weakly nonlinear water wave equations: KdV and CH

The derivation of weakly nonlinear water wave equations starts with Laplace's equation for the velocity potential of an inviscid, incompressible, and irrotational fluid moving in a vertical plane under gravity with an upper free surface, as, e.g., in [1]. The equations are then expanded in the small parameters $\epsilon_1 = a/h$ and $\epsilon_2 = h^2/l^2$. Here $\epsilon_1 \geq \epsilon_2 > \epsilon_1^2$ and a, h, and l denote the wave amplitude, the mean water depth, and a typical horizontal length scale (e.g., a wavelength), respectively. Length is measured in terms of l, height in h and time in l/c_0. The elevation η is scaled with a and fluid velocity u is scaled with $c_0 a/h$. Here, $c_0 = \sqrt{gh}$ is the linear wave speed for undisturbed water at rest at spatial infinity, where u and its derivatives u_x and u_{xx} are taken to vanish.

The result of the expansion to quadratic order in ϵ_1 and ϵ_2 is the equation for the surface elevation η [1], p. 466, while higher order terms (HOT) can e.g. be found, e.g., in [11],

$$0 = \eta_t + \eta_x + \frac{3}{2}\epsilon_1\,\eta\,\eta_x + \frac{1}{6}\epsilon_2\,\eta_{xxx} - \frac{3}{8}\epsilon_1^2\,\eta^2\,\eta_x + \epsilon_1\epsilon_2\left(\frac{23}{24}\eta_x\,\eta_{xx} + \frac{5}{12}\eta\,\eta_{xxx}\right) + \epsilon_2^2\frac{19}{360}\eta_{xxxxx} + HOT, \tag{7.5}$$

where partial derivatives are denoted by subscripts, Next, following Kodama [9, 10] one applies the near-identity transformation, $\eta = u + \epsilon_1\,f(u) + \epsilon_2\,g(u)$, to the $\eta-$equation (7.5) and seeks functionals $f(u)$ and $g(u)$ that consolidate the terms of order $O(\epsilon_1^2)$ and $O(\epsilon_1\epsilon_2)$ in (7.5) into one order $O(\epsilon_2^2)$ term under normal form transformations. This procedure produces the following 1+1 quadratically nonlinear equation for unidirectional water waves with fluid velocity, $u\,(x,t)$ and momentum $m = u - \alpha^2 u_{xx}$, with constant $\alpha^{\,2} = (19/60)\epsilon_2$, see [8],

$$m_t + u_x + \frac{\epsilon_1}{2}(um_x + 2mu_x) + \epsilon_2\frac{3}{20}u_{xxx} = 0\,,\ , \tag{7.6}$$

After these normal form transformations, equation (7.6) is equivalent to the shallow water wave equation (7.5) up to, and including, terms of order $\mathcal{O}(\epsilon_2^2)$.

Equation (7.6) restricts to two separately integrable soliton equations for water waves. After setting $\alpha^{\,2} \to 0$ and rescaling, this equation becomes the classic Korteweg-de Vries (KdV) equation,

$$u_t + c_0 u_x + 3u\,u_x + \gamma\,u_{xxx} = 0\,, \tag{7.7}$$

which for $c_0 = 0$ has the famous soliton solution $u(x,t) = u_0\,\mathrm{sech}^2((x-ct)\sqrt{u_0/\gamma}/2)$, $c = c_0 + u_0$ see, e.g., [2].

Instead, choosing the Galilean frame for which removes the u_{xxx} term in equation (7.6) implies the Camassa-Holm (CH) equation,

$$u_t + c_0 u_x + 3u\,u_x - \alpha^{\,2}\,u_{xxt} = \alpha^{\,2}\Big(2u_x u_{xx} + u u_{xxx}\Big)\,, \tag{7.8}$$

which for $c_0 = 0$ has the "peakon" soliton solutions $u(x,t) = ce^{-|x-ct|}$ discovered and analyzed in [3], [4].

For the remainder of these notes, we will investigate the KdV and CH soliton equations (7.7) and (7.8) in the context of geometric mechanics.

References

[1] G.B. Whitham, *Linear and Nonlinear Waves*, Wiley Interscience:New York (1974).

[2] M.J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*, SIAM: Philadelphia (1981).

[3] R. Camassa and D.D. Holm, Phys. Rev. Lett. **71**, 1661 (1993).

[4] Equation (7.6) also arises as the Euler-Poincaré equation for an averaged Lagrangian [5]. Its solutions describe geodesic motion with respect to the H_1 metric of u on the Bott-Virasoro Lie group [6]. The KdV equation arises the same way for the L_2 metric, see [7].

[5] D.D. Holm, J.E. Marsden and T.S. Ratiu, Phys. Rev. Lett., **80**, 4173 (1998). D.D. Holm, J.E. Marsden, T.S. Ratiu, Adv. in Math. **137**, 1 (1998).

[6] G. Misiolek, J. Geom. Phys. **24**, 203 (1998).

[7] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry*, 2nd Edition, Springer:New York (1999).

[8] H.R. Dullin, G. Gottwald, D.D. Holm, in preparation.

[9] Y. Kodama, Phys. Lett. A **107**, 245, **112**, 193 (1985); **123**, 276 (1987).

[10] Y. Kodama and A. V. Mikhailov, in *Algebraic Aspects of Integrable Systems: In Memory of Irene Dorfman*, edited by A. S. Fokas and I. M. Gelfand, Birkhäuser, Boston, (1996) pp 173-204.

[11] Li Zhi and N. R. Sibgatullin, J. Appl. Maths. Mechs. **61** 177 (1997)

7.4 Introduction to Lax equations and Isospectrality Principles for KdV and CH

Lax equation Suppose we have a smooth one-parameter family $U(t)$ (with $U(0) = I$) of unitary transformations of a Hilbert space H, so that $\psi(t) = U(t)\psi(0)$ for $\psi(0) \in H$. The time derivative $U_t(t)$ of $U(t)$, is a tangent vector at $U(t)$ of the group $U(H)$ of unitary transformations of H, satisfying $U^{-1} = U^*$, where superscript U^* denotes Hermitian adjoint of U. Consider the right-invariant Lie algebra operator $B(t) \in \mathfrak{u}$ given by

$$B(t) = U_t(t)U(t)^{-1} = U_t(t)U(t)^* \in \mathfrak{u}\,,$$

which is a tangent vector to the curve $U(H)$ at the identity, I. Differentiating $UU^* = I$ shows that $B = U_tU^* = -(U_tU^*)^* = -B^*$ is a skew-Hermitian operator on H.

Now suppose that $L(0)$ is a self-adjoint operator on H, and define a family of unitarily equivalent operators $L(t)$ by

$$L(t) = U(t)L(0)U(t)^{-1}\,, \quad \text{so} \quad L(0) = U(t)^*L(t)U(t)\,.$$

Differentiating the latter equation with respect to t yields,

$$0 = U_t^*LU + U^*L_tU + U^*LU_t = U^*(-BL + L_t + LB)U\,.$$

Hence, writing $[B, L] = BL - LB$ as usual for the commutator of B and L, we see that $L(t)$ satisfies the so-called *Lax Equation*,

$$L_t = [B, L]\,.$$

For any $\psi(0)$ in H, recall that $\psi(t) = U(t)\psi(0)$. Since $U(t)L(0) = L(t)U(t)$, it follows that if $\psi(0)$ satisfies the eigenvalue equation $L(0)\psi(0) = \lambda\psi(0)$, then $L(t)\psi(t) = \lambda\psi(t)$, so that $\psi(t)$ remains an function of $L(t)$ belonging to the same eigenvalue λ. In addition, since $\psi(0) = U(t)^{-1}\psi(t)$,

$$\psi_t = U_t(t)\psi(0) = B\psi(t)\,,$$

so $\psi(t)$ solves this linear evolutionary ODE with initial value $\psi(0)$.

Isospectral Principle Let $L(t)$ and $B(t)$ be smooth one-parameter families of self-adjoint and skew-adjoint operators respectively on a Hilbert space H, satisfying the Lax Equation $L_t = [B, L]$, and let $\psi(t)$ be a curve in H that is a solution of the time-dependent linear ODE $\psi_t = B\psi(t)$.

If the initial value, $\psi(x, 0)$, is an eigenvector of $L(0)$ belonging to an eigenvalue λ, then $\psi(x, t)$ is an eigenvector of $L(t)$ belonging to the same eigenvalue λ. In the next remark, we will explain the geometrical meaning of the Lax equation as a Zero Curvature Condition (ZCC). Then we will discuss the Isospectral Principle for the celebrated KdV equation, for which $H = L^2(\mathbb{R})$.

Remark

7.2 (Zero Curvature Condition (ZCC)). *The differential geometric meaning of the Lax equation $L_t = [B, L]$ can be seen by rewriting $L = \partial_x - A$ for an operator A and computing the commutation relations*

$$\begin{aligned} L_t &= \big[B\,,\,L\big] = \big[B\,,\,\partial_x - A\big] = -B_x - \big[B\,,\,A\big] \\ &= -\big[L\,,\,\partial_t\big] = -\big[\partial_x - A\,,\,\partial_t\big] = -\big[\partial_x\,,\,\partial_t\big] + \big[A\,,\,\partial_t\big] = -A_t\,. \end{aligned}$$

Provided we have equality of cross derivative, $\psi_{xt} = \psi_{tx}$, the previous calculation produces the result

$$\big[\partial_x - A\,,\,\partial_t - B\big] = A_t - B_x + \big[A\,,\,B\big] = 0\,.$$

As we will now discuss, the previous relation may be interpreted as the Zero Curvature Condition (ZCC) for a flat connection on the trivial principal bundle $\mathbb{R}^2 \times G$, where G is a matrix Lie group. On this bundle, we may write the flat connection as $\nabla = d - \varpi$, where ϖ is a 1-form on $\mathbb{R}^2$ with values in the matrix Lie algebra $\mathfrak{g}$ of G. In coordinates $(x,t) \in \mathbb{R}^2$, we may write the connection 1-form as $\varpi = Adx + Bdt$ where A and B are smooth maps of $\mathbb{R}^2$ into the matrix Lie algebra $\mathfrak{g}$.

For a vector field X on $\mathbb{R}^2$, the covariant derivative operator in the direction X is given by $\nabla_X = \partial_X - \varpi(X)$. In particular, the covariant derivatives in the coordinate directions ∂_x and ∂_t for the connection 1-form $\varpi = Adx + Bdt$ are

$$\nabla_{\partial_x} = \partial_x - A\,, \quad \textit{and} \quad \nabla_{\partial_t} = \partial_t - B\,.$$

The flatness of the connection ∇ may be expressed in several equivalent ways. First, the curvature 2-form vanishes. This shown by computing

$$d\varpi - \varpi \wedge \varpi = 0\,.$$

Equivalently, the covariant derivative operators in the ∂_x and ∂_t directions commute, i.e.,

$$\big[\nabla_{\partial_x}\,,\,\nabla_{\partial_t}\big] = 0\,,$$

and finally, cross-derivatives of ψ are equal, $(A\psi)_t = \psi_{xt} = \psi_{tx} = (B\psi)_x$.

In the first paragraph of this remark, we showed directly that equality of cross-derivatives of ψ is sufficient for the Lax pair relation to imply the ZCC. The last remark shows that the converse also holds. That is, the ZCC implies the Lax pair relation.

Remark. In the next theorem, we will see that if $u(x,t)$ satisfies the KdV equation in the form, $u_t + c_0 u_x + 3uu_x + \gamma\, u_{xxx} = 0$, then the family of Schroedinger operators $L(t) = 4\gamma\partial_x^2 + 2u(x,t) + c_0$ on H satisfies the Lax Equation,

$$L_t = [B, L]\,, \quad \text{with} \quad L\psi = \Big(4\gamma\partial^2 + 2u(x,t) + c_0\Big)\psi = \lambda\psi \quad \text{and} \quad B\psi = -\Big(c_0\partial + 4\gamma\partial^3 + \frac{3}{2}(u\partial + \partial u)\Big)\psi\,, \tag{7.9}$$

where $L^* = L$ is a self-adjoint operator on $H = L^2(\mathbb{R})$, and $B^* = -B$ is a skew-adjoint operator. This Lax Equation is the basis for the complete integrability of the KdV equation as a Hamiltonian PDE.

Exercise. Check that the Lax Equation with the L and B operators in equation (7.9) produces the KdV equation, provided $d\lambda/dt = 0$.
★

7.5 KdV Isospectrality Theorem

Theorem

7.3 (KdV Isospectrality Theorem)**.** *Suppose $u(x,t)$ is a solution of the KdV equation for weakly nonlinear surface waves propagating in shallow water of depth h in a reference frame moving with the linear wave speed $c_0 = \sqrt{gh}$, where g is gravitational acceleration,*

$$u_t + c_0 u_x + 3uu_x + \gamma\, u_{xxx} = 0\,. \tag{7.10}$$

Suppose the initial value $u(x,0)$ lies in the Schwartz space $S(R)$, and that $\psi(x)$ is an eigenfunction of the Schroedinger Equation with potential $u(x,0)$ and eigenvalue λ:

$$L(0)\psi(x) = 4\gamma\psi_{xx}(x) + 2u(x,0)\psi(x) + c_0\psi(x) = \lambda\psi(x)\,. \tag{7.11}$$

Let $\psi(x,t)$ be the solution of the evolution equation $\psi_t = B\psi$, with

$$\psi_t = B\psi = -\Big(c_0\partial + 4\gamma\partial^3 + \frac{3}{2}(u\partial + \partial u)\Big)\psi(x,t) \tag{7.12}$$

with the initial value $\psi(x,0) = \psi(x)$. Then $\psi(x,t)$ is an eigenfunction for the Schroedinger Equation with potential $2u(x,t) + c_0$ and the same eigenvalue λ:

$$4\gamma\psi_{xx}(x,t) + \big(2u(x,t) + c_0\big)\psi(x,t) = \lambda\psi(x,t)\,. \tag{7.13}$$

Moreover, if $\psi(x)$ is in L^2, then the L^2 norm of $\psi(\cdot,t)$ is independent of time t. Finally, $\psi(x,t)$ also satisfies the first-order evolution equation

$$\psi_t + (\lambda + u)\psi_x - \frac{1}{2}u_x\psi = 0. \tag{7.14}$$

Proof. Except for the final statement, this is an immediate application of the Isospectrality Principle. Differentiating the eigenvalue equation for $\psi(x,t)$ with respect to x gives

$$-4\gamma\psi_{xxx} = u_x\psi + (u - \lambda)\psi_x\,,$$

and substituting this into the assumed evolution equation $\psi_t = B\psi(t)$ for ψ gives the asserted first-order equation (7.14) for ψ. □

Corollary

7.4 (KdV Compatibility relation)**.** *Requiring compatibility $\psi_{xxt} = \psi_{txx}$ between equations* (7.13) *and* (7.14) *in the KdV Isospectrality Theorem 7.5 yields the KdV equation* (7.10)*, as the condition for $d\lambda/dt = 0$. (The proof is a direct calculation.)*

7.6 The CH equation

The Camassa-Holm (CH) equation (7.8) for unidirectional water waves with fluid velocity $u\,(x,t)$ may be written compactly as [3],

$$m_t + c_0 u_x + u\,m_x + 2m\,u_x = -\,\gamma\,u_{xxx}\,. \tag{7.15}$$

Here $m = u - \alpha^2 u_{xx}$ is a momentum variable, partial derivatives are denoted by subscripts, the constants α^2 and γ/c_0 are squares of length scales, and $c_0 = \sqrt{gh}$ is the linear wave speed for undisturbed water at rest at spatial infinity, where u and m are taken to vanish. (Any constant value $u = u_0$ is also a solution.)

The interplay between the local and nonlocal linear dispersion in this equation is evident in its phase velocity relation,

$$c_p(k) = \frac{\omega}{k} = \frac{c_0 - \gamma\,k^2}{1 + \alpha^2 k^2}\,, \tag{7.16}$$

for waves with frequency ω and wave number k linearized around $u = 0$. For $\gamma/c_0 < 0$, short waves and long waves travel in the same direction. Long waves travel faster than short ones (as required in shallow water) provided $\gamma/c_0 > -\,\alpha^2$. In that case, the phase velocity $c_p = \omega/k$ lies in the interval, $\omega/k \in (-\,\gamma/\alpha^2, c_0]$.

Exercise. Verify the phase velocity relation for CH claimed in equation (7.16). ★

The CH water wave equation in (7.15) is not Galilean invariant. Upon shifting the velocity variable by u_0 and moving into a Galilean frame $\xi = x - ct$ with velocity c, so that $u(x,t) = \tilde{u}(\xi,t) + c + u_0$, this equation transforms to

$$\tilde{m}_t + \tilde{u}\,\tilde{m}_\xi + 2\tilde{m}\,\tilde{u}_\xi = -\tilde{c}_0\,\tilde{u}_\xi - \tilde{\gamma}\,\tilde{u}_{\xi\xi\xi}\,, \tag{7.17}$$

with $\tilde{c_0} = (c_0 + 2c + 3u_0)$, $\tilde{\gamma} = (\gamma - u_0\alpha^2)$ and appropriately altered boundary conditions at spatial infinity. Hence, we must regard equation (7.6) as a *family* of equations whose linear dispersion parameters c_0, γ depend on the appropriate choice of Galilean frame and boundary conditions. The parameters c_0 and γ may even be *removed* by making such a choice. For example, as we have seen in the CH equation in (7.8), $\tilde{\gamma}$ may be removed in (7.17) by choosing $u_0 = \gamma/\alpha^2$ in the boundary condition at spatial infinity.

Below, we shall identify how the dispersion coefficients for the linearized water waves appear as *parameters* in the isospectral problem for CH.

7.7 CH Compatibility Theorem

Theorem

7.5 (CH Compatibility Theorem). *The nonlinear equation* (7.15) *arises as a compatibility condition for two linear equations, namely, the isospectral eigenvalue problem,*

$$\lambda\left(\frac{1}{4} - \alpha^2\partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi\,, \tag{7.18}$$

and the evolution equation for the eigenfunction ψ,

$$\psi_t = -(u+\lambda)\,\psi_x + \frac{1}{2}u_x\,\psi\,. \tag{7.19}$$

Compatibility of these two linear equations ($\psi_{xxt} = \psi_{txx}$) *and isospectrality* ($d\lambda/dt = 0$) *imply equation (7.15).*

Exercise. Verify the CH Compatibility Theorem 7.5 by direct computation. ★

Proof. Direct computation. □

Corollary

7.6 (CH Isospectrality Theorem). *The nonlinear equation* (7.15) *preserves the spectrum of the operator*

$$L = \left(\frac{1}{4} - \alpha^2 \partial_x^2\right)^{-1} * \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right). \tag{7.20}$$

Consequently, the nonlinear water wave equation (7.15) admits the IST method for the solution of its initial value problem, just as the KdV and CH equations do. In fact, the isospectral problem for equation (7.15) restricts to the isospectral problem for KdV (i.e., the Schrödinger equation) when $\alpha^2 \to 0$ *and it restricts to the isospectral problem for CH discovered in [3] when* $\gamma \to 0$.

Remark

7.7. *The spectral problem on the real line can be transformed to the string density problem on a finite interval. In the special case in which* $c_0 = 0 = \gamma$*, the inverse problem was solved by Stieltjes (1894).*

For more discussion of this relationship between the isospectral problem for the CH equation and the classical inverse problem for a plucked string, see

Beals, R., Sattinger, D.H. and Szmigielski, J. [2007] The string density problem and the Camassa–Holm equation. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 365(1858), pp.2299-2312.

7.8 The Euler-Poincaré equation, $\mathrm{EPDiff}H^1(\mathbb{R})$, aka the CH equation for $c_0 = 0 = \gamma$

The CH equation for the case $c_0 = 0 = \gamma$ reduces to the following equation, called $\mathrm{EPDiff}H^1(\mathbb{R})$,

$$m_t + um_x + 2mu_x = 0\,, \quad \text{where} \quad m = u - \alpha^2 u_{xx}\,, \quad \text{and} \quad \lim_{|x|\to\infty}(u, u_x, u_{xx}) \to 0\,. \tag{7.21}$$

where subscripts denote partial derivatives in x and t, and we have assumed that $u(x)$ and its first two derivatives vanish as $|x| \to \infty$. This is the EPDiff evolution equation on the real line ($\mathbb{R}$) for a Lagrangian $\ell(u) = \frac{1}{2}\|u\|^2_{H^1}$, written in terms of its velocity u and its momentum $m = \delta l/\delta u = u - \alpha^2 u_{xx}$.

Exercise.

Verify that the $\mathrm{EPDiff}H^1(\mathbb{R})$ equation (7.21) arises as an Euler-Poincaré equation, when the Lagrangian is chosen to be half the square of the H^1 norm $\|u\|_{H^1}$ of the vector field of velocity $u = \dot{g}g^{-1} \in \mathfrak{X}(\mathbb{R})$ on the real line $\mathbb{R}$ with $g \in \mathrm{Diff}(\mathbb{R})$. That is, the Lagrangian is chosen as

$$\ell(u) = \frac{1}{2}\|u\|^2_{H^1} = \frac{1}{2}\int_{-\infty}^{\infty} u^2 + \alpha^2 u_x^2 \, dx\,. \tag{7.22}$$

Explain the geometric meaning of the solutions of the initial value problem for the $\mathrm{EPDiff}H^1(\mathbb{R})$ evolution equation. Hint: consider geodesic motion. Discuss the issue of existence of geodesics. ★

Comparisons of the $\mathrm{EPDiff}H^1(\mathbb{R})$ equation for geodesic motion with respect to the H^1 norm $\|u\|_{H^1}$ with Burgers and KdV.

- When $\alpha^2 \to 0$, the Lagrangian in (7.22) becomes

$$\ell(u) = \frac{1}{2}\|u\|^2_{L^2} = \frac{1}{2}\int_{-\infty}^{\infty} u^2 \, dx\,, \tag{7.23}$$

and the corresponding $\mathrm{EPDiff}L^2(\mathbb{R})$ equation becomes the **inviscid Burgers equation**, written as

$$u_t + 3uu_x = 0\,. \tag{7.24}$$

Exercise. Write the Burgers equation as an Euler-Poincaré equation. Explain the geometric meaning of the solutions of the initial value problem for the Burgers equation. Hint: consider geodesic motion. Discuss the issue of existence of geodesics connecting pairs of different L^2 functions. ★

- **The KdV equation**

$$u_t + c_0 u_x + 3uu_x + \gamma\, u_{xxx} = 0\,, \tag{7.25}$$

models the propagation of weakly nonlinear surface waves in shallow water of depth h, in a reference frame moving with the linear wave speed $c_0 = \sqrt{gh}$, where g is gravitational acceleration. According to the KdV Isospectrality Theorem 7.5, if $u(x,t)$ satisfies the KdV equation (7.25), then the family of Schroedinger operators

$$L(t) = 4\gamma\partial_x^2 + 2u(x,t) + c_0$$

on $H(\mathbb{R})$ satisfies the Lax Equation

$$L_t = [B, L]\,, \quad \text{so that} \quad L(t) = U(t)l(0)U^*(t)\,,$$

where

$$\begin{aligned}\psi_t(x,t) &= U_t\psi(x,0) = U_tU^*\psi(x,0)\\ &= B(x,t)\psi(x) = -\Big(c_0\partial + 4\gamma\partial^3 + \frac{3}{2}\big(u(x,t)\partial + \partial u(x,t)\big)\Big)\psi\,,\end{aligned}$$

or, as a skew-adjoint operator, $B^* = -B = c_0\partial + 4\gamma\partial^3 + \frac{3}{2}(u\partial + \partial u)$. This is the basis for the complete integrability of the KdV equation as a Hamiltonian PDE, since KdV is the compatibility condition for

$$L\psi = \lambda\psi\,, \quad \psi_t = B\psi \quad \text{and} \quad d\lambda/dt = 0\,,$$

so that the eigenvalue spectrum of KdV provides an infinity of conserved quantities.

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

$$m_t + um_x + 2mu_x = 0\,, \quad \text{where} \quad m = u - u_{xx}\,. \tag{7.26}$$

This $\mathrm{EPDiff}H^1(\mathbb{R})$ equation will turn out to be an integrable soliton equation and, thus, will share all of the fundamental properties of the famous KdV equation.

Remark

7.8 (Solution behaviour of EPDiff$H^1(\mathbb{R})$). *The peakon-train solutions of* EPDiff$H^1(\mathbb{R})$ *are an* ***emergent phenomenon****. A wave train of peakons emerges in solving the initial-value problem for the* EPDiff$H^1(\mathbb{R})$ *equation (7.26) for essentially any spatially confined initial condition. A numerical simulation of the solution behaviour for* EPDiff$H^1(\mathbb{R})$ *given in Figure 4 shows the emergence of a wave train of peakons from a Gaussian initial condition.*

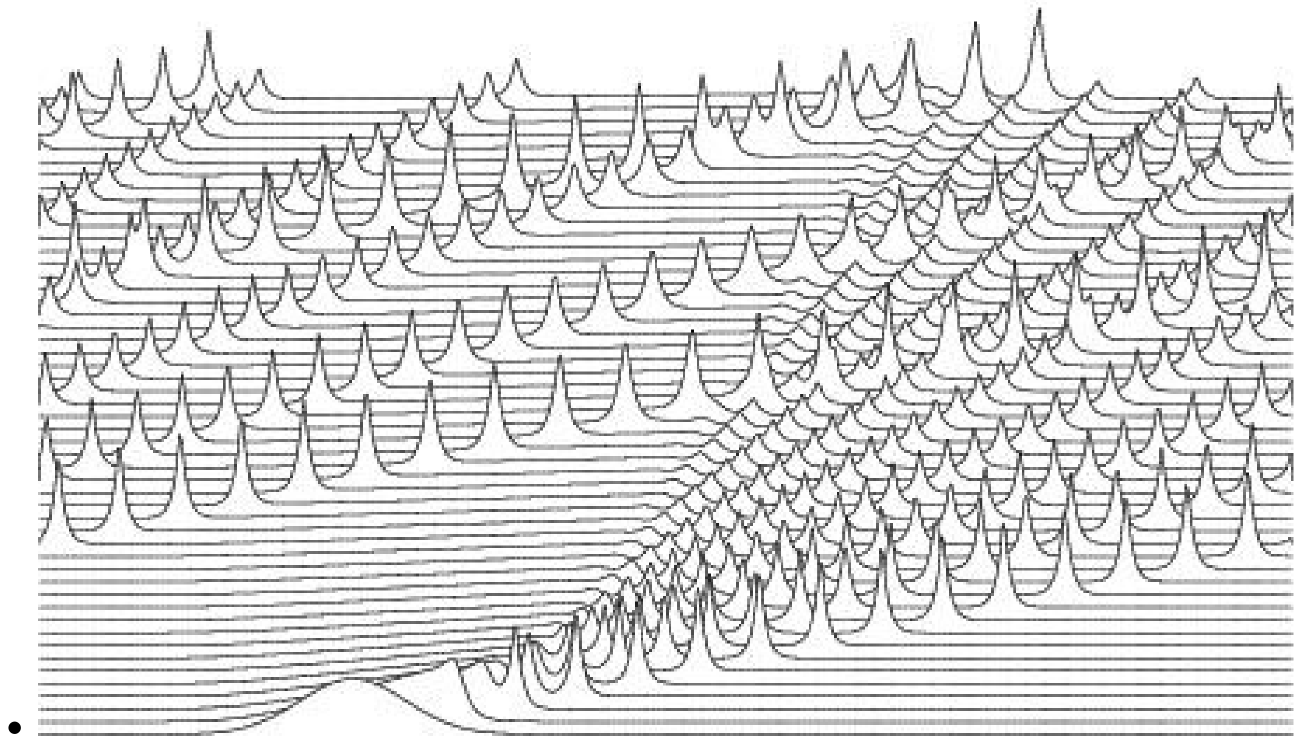


Figure 4: Under the evolution of the EPDiff$H^1(\mathbb{R})$ equation (7.26), an ordered ***wave train of peakons*** emerges from a smooth localized initial condition (a Gaussian). The spatial profiles at successive times are offset in the vertical to show the evolution. The peakon wave train eventually 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. (Properties of the EPDiff$H^1(\mathbb{R})$ equation)

(A) Obtain the EPDiff equation (7.21) by computing

$$\frac{d}{dt}\Big(\mathrm{Ad}^*_{g^{-1}(t)}\mu(x,t)\Big)=0 \quad \text{with} \quad \mu(x,t)=m(x,t)\,(dx)^2\,.$$

Answer.

One computes directly that

$$0=\frac{d}{dt}\Big(\mathrm{Ad}^*_{g^{-1}(t)}\mu(x,t)\Big)=\mathrm{Ad}^*_{g^{-1}(t)}\Big(\partial_t\mu(x,t)+\mathrm{ad}^*_{\dot g_t g_t^{-1}}\mu(x,t)\Big)=\mathrm{Ad}^*_{g^{-1}(t)}\Big(\partial_t\mu(x,t)+\pounds_{\dot g_t g_t^{-1}}\mu(x,t)\Big)\,,$$

where the coadjoint operation $\mathrm{ad}^*_{\dot g_t g_t^{-1}}$ is the Lie derivative $\pounds_{\dot g_t g_t^{-1}}$ for 1-form densities such as $\mu(x,t)=m(x,t)(dx)^2$. Finally, one substitutes $\dot g_t g_t^{-1}=u$, which completes the comparison with the EPDiff equation.

▲

(B) Regard the momentum 1-form density μ as the pullback $g_t l$ of a Lagrangian label l by the flow g_t; namely,

$$\mu(g_t l,t)=g^*(t)(m(l,0)(dl)^2)=m(g_t l,t)(d(g_t l))^2\,,$$

where $x(l,t)=g_t l$ is the Lagrangian trajectory of fluid parcels in the flow of g_t, labelled by their initial positions $x(l,0)=g_0 l=l$. Prove by direct computation that the EPDiff equation (7.21) is equivalent to the following relation,

$$\frac{d}{dt}\Big(m(x(l,t),t)(dx(l,t))^2\Big)=0 \quad \text{along} \quad \frac{dx(l,t)}{dt}=u(x(l,t),t)\,, \quad \text{arising from} \quad \dot g_t=u\circ g_t\,.$$

$$0=\frac{d}{dt}\Big(m(x(l,t),t)(dx(l,t))^2\Big)=\Big[\Big(\partial_t m(x,t)+um_x+2mu_x\Big)(dx)^2\Big]_{x=x(l,t)}\,,$$

since $\frac{dx(l,t)}{dt}=u(x(l,t),t)$ and $\frac{d}{dt}dx(l,t)=d\frac{dx}{dt}=du=u_x\,dx$ along $x=x(l,t)$.

Answer.

$$0 = \frac{d}{dt}\Big(m(x(l,t),t)(dx(l,t))^2 \Big) = \Big[\Big(\partial_t m(x,t) + u m_x + 2 m u_x \Big) (dx)^2 \Big]_{x=x(l,t)} \,,$$

since $\frac{dx(l,t)}{dt} = u(x(l,t),t)$ and $\frac{d}{dt}dx(l,t) = d\frac{dx}{dt} = du = u_x\,dx$ along $x = x(l,t)$.

▲

(C) By using the Euler-Poincaré approach with Lagrangian $\ell(u) = \frac{1}{2}\|u\|^2_{H^1}$, derive the EPDiff equation on the real line (7.21) in terms of its velocity u and its momentum $m = \delta l/\delta u = u - u_{xx}$ in one Eulerian spatial dimension for this Lagrangian.
Hint: Prove a Lemma first, that $u = \dot{g}_t g_t^{-1}$ implies $\delta u = \partial_t \eta - \mathrm{ad}_u \eta$ with $\eta = \delta g g^{-1}(t)$.

Answer.

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,

$$u' - \dot{\eta} = \nabla u \cdot \eta - \nabla \eta \cdot u = -\left[u, \eta \right] =: -\mathrm{ad}_u \eta \,,$$

so that

$$u' = \dot{\eta} - \mathrm{ad}_u \eta\,.$$

In 3D, this becomes

$$\delta \mathbf{u} = \dot{\mathbf{v}} - \mathrm{ad}_{\mathbf{u}} \mathbf{v}\,. \tag{7.27}$$

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 (7.27), since

$$\mathbf{u}' - \dot{\mathbf{v}} = \nabla \mathbf{u} \cdot \mathbf{v} - \nabla \mathbf{v} \cdot \mathbf{u} = -\,[\mathbf{u}, \mathbf{v}] = -\,\mathrm{ad}_{\mathbf{u}} \mathbf{v}\,.$$

□

The EPDiff(H^1) **equation on** $\mathbb{R}$**.** The EPDiff(H^1) equation is written on the real line in terms of its velocity u and its momentum $m = \delta l/\delta u$ in one spatial dimension as

$$m_t + u m_x + 2 m u_x = 0\,, \quad \text{where} \quad m = u - u_{xx} \tag{7.28}$$

where subscripts denote partial derivatives in x and t.

Proof. This equation is derived from the variational principle with $l(u) = \frac{1}{2}\|u\|^2_{H^1}$ as follows.

$$
\begin{aligned}
0 = \delta S &= \delta \int l(u)dt = \frac{1}{2}\,\delta\iint u^2 + u_x^2\,dx\,dt \\
&= \iint (u - u_{xx})\,\delta u\,dx\,dt =: \iint m\,\delta u\,dx\,dt \\
&= \iint m\,(\eta_t - \mathrm{ad}_u\eta)\,dx\,dt \\
&= \iint m\,(\eta_t + u\eta_x - \eta u_x)\,dx\,dt \\
&= -\iint (m_t + (um)_x + mu_x)\,\eta\,dx\,dt \\
&= -\iint (m_t + \mathrm{ad}^*_u m)\,\eta\,dx\,dt\,,
\end{aligned}
$$

where $u = \dot{g}g^{-1}$ implies $\delta u = \eta_t - \mathrm{ad}_u\eta$ with $\eta = \delta g g^{-1}$. □

▲

(D) Re-derive the EPDiff equation (7.21), by using the Hamilton-Pontryagin approach, by using the constrained action

$$
S(u,m,g) = \int \ell(u)\,dt + \int m(x,t)\big(\dot{g}_t g_t^{-1}x - u(x,t)\big)\,dx\,dt\,.
$$

Answer.

The Hamilton-Pontryagin principle is given by

$$
0 = \delta S(u,m,g) = \delta\int \ell(u)\,dt + \delta\int m(x,t)\big(\dot{g}_t g_t^{-1}x - u(x,t)\big)\,dx\,dt\,.
$$

Taking variations yields

$$\delta u : \quad \frac{\delta \ell}{\delta u} - m(x,t) = 0\,, \quad \delta g : \quad \delta\big(\dot g_t g_t^{-1}\big) = (\dot\eta - \mathrm{ad}_{\dot g_t g_t^{-1}}\eta)\,, \quad \text{with} \quad \eta = \delta g_t g_t^{-1}\,.$$

Integrations by parts in time and space in the 2nd integral, while using the definition of ad^* then yields the EPDiff equation for $m = \delta\ell/\delta u$.

▲

(E) Derive the EPDiff equation (7.21) once more, by using the Clebsch approach based on the inverse map $g_t^{-1}x = l(x,t)$, where $l(x,t)$ is the initial label of the Lagrangian parcel that is occupying Eulerian spatial position x at a given time t during the flow. The label is carried by the flow, so it satisfies the scalar advection law, $\partial_t l + u(x,t)l_x = 0$. In terms of these variables, the Clebsch-constrained Hamilton's principle is given explicitly by,

$$0 = \delta S(u,l,\pi) = \delta \int \ell(u)\,dt + \delta \int \pi(x,t)\big(\partial_t l + u(x,t)l_x\big)dx\,dt\,.$$

Answer.

The Clebsch Hamilton principle is given by,

$$0 = \delta S(u,l,\pi) = \delta \int \ell(u)\,dt + \delta \int \pi(x,t)\big(\partial_t l + u(x,t)l_x\big)dx\,dt\,.$$

Taking variations yields

$$\delta u : \quad \frac{\delta \ell}{\delta u} + \pi l_x = 0\,, \quad \delta\pi : \quad \partial_t l + u(x,t)l_x = 0\,, \quad \delta l : \quad \partial_t \pi + \partial_x(u(x,t)\pi) = 0\,.$$

Substituting the 2nd and 3rd equations into the partial time derivative of the 1st equation now yields the EPDiff equation for $m = \delta\ell/\delta u$.

▲

(F) Determine the cotangent-lift momentum map $J_R(l,\pi) := -\pi l_x dx$ from the Clebsch Hamilton principle and show via the canonical Poisson bracket, $\{l(x,t),\pi(x',t)\} = \delta(x-x')$, that $\int J(l,\pi)$ generates an infinitesimal Eulerian spatial shift in both the Lagrangian label or inverse map $l(x,t) = g_t^{-1}x$ and in its canonically conjugate momentum in phase space, $\pi(x,t)$.

Answer.

The momentum map is given by the variation in velocity, u, as $\frac{\delta \ell}{\delta u} + \pi l_x = 0$. The canonical Poisson bracket $\{l(x,t),\pi(x',t)\} = \delta(x-x')$ with $\int J(l,\pi) = -\int \pi l_x dx$ yields

$$\delta l = \{l(x,t)\,,\, J\} = -\,l_x \quad \text{and} \quad \delta \pi = \{\pi(x,t)\,,\, J\} = -\,\pi_x$$

▲

(G) Use the Clebsch-constrained Hamilton's principle

$$S(u,p,q) = \int \ell(u)\,dt + \sum_{a=1}^{N}\int p_a(t)\big(\dot q_a(t) - u(q_a(t),t)\big)dt$$

to derive the peakon singular solution of EPDiff

$$m(x,t) = \frac{\delta l}{\delta u} = \sum_{a=1}^{N} p_a(t)\delta(x-q_a(t)) =: J_L(q,p) \tag{7.29}$$

as a cotangent-lift momentum map in terms of canonically conjugate variables $q_a(t)$ and $p_a(t)$, with $a = 1,2,\dots,N$.

Answer.

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

$$\begin{aligned} 0 = \delta S \quad &= \int \left(\frac{\delta l}{\delta u} - \sum_{a=1}^{N} p_a \delta(x - q_a(t)) \right) \delta u \, dx\, dt \\ &\quad - \int \left(\dot p_a(t) + \frac{\partial u}{\partial q_a}\, p_a(t) \right) \delta q_a - \delta p_a \Big(\dot q_a(t) - u(q_a(t),t) \Big) dt \,. \end{aligned}$$

The singular momentum solution $m(x,t)$ of EPDiff is thus obtained as the cotangent-lift momentum map in (7.29); namely,

$$m(x,t) = \delta l/\delta u = \sum_{a=1}^{N} p_a(t)\delta(x - q_a(t)) \tag{7.30}$$

The two momentum maps $\mathbf{J}_R(l,\pi)$ and $\mathbf{J}_L(q,p)$ may be assembled into a single figure as follows:

$$\begin{tikzcd} & T^*\mathrm{Diff}(\mathbb{R}) \arrow[dl, "\mathbf{J}_L(q,p)"'] \arrow[dr, "\mathbf{J}_R(l,\pi)"] & \\ \mathfrak{X}(\mathbb{R})^* & & \mathfrak{X}(\mathbb{R})^* \end{tikzcd}$$

▲

(H) Use the momentum map $m = -\,\pi l_x$ to transform from the canonical Poisson bracket in the variables $(l(x,t), \pi(x,t))$ to the associated Lie-Poisson bracket in the variable $m(x,t)$ for the Hamiltonian form of EPDiff($\mathbb{R}$).

Answer.

The momentum map $m = -\pi l_x$ allows one to transform from the canonical Poisson bracket in the variables $(l(x,t), \pi(x,t))$ to the associated Lie-Poisson bracket in the variable $m(x,t)$ for the Hamiltonian form of EPDiff($\mathbb{R}$), as follows.

$$
\begin{aligned}
B_2 = JB_1J^\dagger &= \left[\tfrac{\delta m}{\delta l} = -\pi\partial, \quad \tfrac{\delta m}{\delta \pi} = -l_x\right] \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \partial\pi \\ -l_x \end{bmatrix} \\
&= \pi\partial l_x + l_x\pi\partial\pi = \partial(\pi l_x) + (\pi l_x)\partial = -\big(m\partial_x + \partial_x m\big)
\end{aligned}
$$

▲

(I) (a) Legendre transform the Lagrangian $\ell(u)$ in equation (7.22) and determine its corresponding Hamiltonian, $h(m)$.

(b) Then use the EPDiff equation and the fibre derivative to compute the associated Lie-Poisson bracket for the Hamiltonian form of EPDiff($\mathbb{R}$). Is this the same bracket as in previous part?

(c) Determine the Lie algebra to which the Lie-Poisson bracket is dual.

(d) Find the Casimir for this Lie-Poisson bracket.

Answer.

(a) The Legendre transform $h(m) =< m, u > -\ell(u)$ of the Lagrangian $\ell(u)$ in equation (7.22) implies the corresponding Hamiltonian,

$$
h(m) = \frac{1}{2}\Big\langle m\,,\,(1-\alpha^2\partial^2)^{-1} * m\Big\rangle = \frac{1}{4}\int_{-\infty}^{\infty} m(x)e^{-|x-y|/\alpha}m(y)\,dx\,dy,.
$$

(b) The Lie-Poisson bracket for the Hamiltonian form of EPDiff($\mathbb{R}$) is

$$
\{f,h\} = -\int \frac{\delta f}{\delta m}\big(m\partial_x + \partial_x m\big)\frac{\delta h}{\delta m}\,dx
$$

Is this the same bracket as in previous part? Yes, including the minus sign!

(c) The Lie-Poisson bracket is defined on $\mathfrak{X}^*(\mathbb{R}) \simeq \Lambda^1 \otimes dens(\mathbb{R})$, the 1-form densities, which comprise the dual of the Lie algebra of vector fields on the real line, $\mathfrak{X}(\mathbb{R})$ with respect to the L^2 pairing.

(d) The Casimir for this Lie-Poisson bracket is $C = \int \sqrt{m}\,dx$, as is easily shown.

Summary of the Lie-Poisson Hamiltonian form of EPDiff. In terms of m, the conserved energy Hamiltonian for the EPDiff equation (7.26) 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 (7.26) and the role of $K(x,y)$ in the kinetic energy metric on the Hamiltonian side.

The corresponding ***Lie-Poisson bracket*** for EPDiff as a Hamiltonian evolution equation is given by,

$$\partial_t m = \big\{m,h\big\} = -\mathrm{ad}^*_{\delta h/\delta m}m = -\left(\partial_x m + m\partial_x\right)\frac{\delta h}{\delta m} \quad \text{and} \quad \frac{\delta h}{\delta m} = u\,,$$

which recovers the starting equation and indicates some of its connections with fluid equations on the Hamiltonian side. For any two smooth functionals f,h of m in the space for which the solutions of EPDiff exist, this Lie-Poisson bracket may be expressed as,

$$\big\{f,h\big\} = -\int\frac{\delta f}{\delta m}(\partial_x m + m\partial_x)\frac{\delta h}{\delta m}dx = -\int m\left[\frac{\delta f}{\delta m}\,,\,\frac{\delta h}{\delta m}\right]dx$$

where $[\cdot\,,\,\cdot]$ denotes the Lie algebra bracket of vector fields. That is,

$$\left[\frac{\delta f}{\delta m}\,,\,\frac{\delta h}{\delta m}\right] = \frac{\delta f}{\delta m}\partial_x\frac{\delta h}{\delta m} - \frac{\delta h}{\delta m}\partial_x\frac{\delta f}{\delta m}\,.$$

▲

(J) For the Lagrangian in equation (7.22), use the momentum map (7.29) to determine the Hamiltonian and canonical equations of motion for the finite-dimensional system of conjugate variables $q_a(t)$ and $p_a(t)$.

Answer.

Inserting the momentum map (7.29) into the Legendre transform

$$h(m) = \langle\, m, u \,\rangle - l(u)$$

yields the conserved energy as the Hamiltonian in canonical variables,

$$e = \frac{1}{2}\int m(x,t)u(x,t)\,dx = \frac{1}{2}\sum_{a=1}^{N}\int p_a(t)\delta(x-q_a(t))u(x,t)\,dx = \frac{1}{2}\sum_{a=1}^{N} p_a(t)u(q_a(t),t)\,. \tag{7.31}$$

Consequently, the variables (q_a, p_a) satisfy the canonical Hamiltonian equations,

$$\dot{q}_a(t) = u(q_a(t),t)\,, \qquad \dot{p}_a(t) = -\frac{\partial u}{\partial q_a}\,p_a(t)\,, \tag{7.32}$$

with the pulse-train solution for velocity

$$u(q_a,t) = \sum_{b=1}^{N} p_b K(q_a,q_b) = \frac{1}{2}\sum_{b=1}^{N} p_b e^{-|q_a-q_b|} \tag{7.33}$$

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 (7.33) has a sharp peak. For that reason, these pulses are called *peakons*. In fact, equations (7.32) are Hamilton's canonical equations with Hamiltonian obtained from equations (7.31) for energy and (7.33) for velocity, as given in [CaHo1993],

$$H_N(q,p) = \frac{1}{2}\sum_{a,b=1}^{N} p_a\,p_b\,K(q_a,q_b) = \frac{1}{4}\sum_{a,b=1}^{N} p_a\,p_b\,\mathrm{e}^{-|q_a-q_b|}\,. \tag{7.34}$$

The first canonical equation in eqn (7.32) implies that the peaks at the positions $x = q^a(t)$ in the peakon-train solution (7.33) 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 (7.29) is the map from Lagrangian coordinates to Eulerian coordinates (that is, the ***Lagrange-to-Euler map***) for the momentum.

▲

(K) Determine the group actions responsible for the two momentum maps $J_R(l, \pi)$ and $J_L(q, p)$.

Answer.

The group actions responsible for the two momentum maps $J_R(l, \pi)$ and $J_L(q, p)$ are obvious from the Clebsch constraints. The equation $l_t + ul_x = 0$ for the scalar label $l(x, t)$ has the solution $l(x, t) = l(x, 0)g_t^{-1}$, which is a right action by g_t^{-1}. The Lie-algebra action $\dot{(q)} = u(q(t), t)$ is a left action, since it corresponds to $q(t) = g_t q(0)$ for which $\dot{q}(t) = \dot{g}_t q(0) = \dot{g}_t g_t^{-1} g_t q(0) = \dot{g}_t g_t^{-1} q(t) = u(q(t), t)$.

▲

(L) Use any variational method you like to derive the new EP equation(s) obtained for a Lagrangian $\ell(u, \rho)$ with $\rho := l_x$.

Answer.

The equation $l_t + ul_x = 0$ for the scalar label $l(x, t)$ implies for $\rho := l_x$ that $\rho_t + \partial_x(\rho u) = 0$. Thus, $\rho(x, t)dx$ is a density, satisfying

$$\rho_t + \partial_x(\rho u) = 0\,, \quad \text{or equivalently} \quad (\partial_t + \pounds_u)(\rho(x, t)dx) = 0\,.$$

Correspondingly, the variation of ρ is given by $\delta(\rho dx) = -\pounds_\eta(\rho dx)$ and $\delta\rho = -(\rho\eta)_x$. Consequently, the Euler-Poincaré variations will yield

$$
\begin{aligned}
0 = \delta S &= \int \Big(- \big(m_t + m\partial_x + \partial_x m\big)\eta + \frac{\delta \ell}{\delta \rho}\delta\rho\Big)\, dxdt \\
&= \int \Big(- \big(m_t + m\partial_x u + \partial_x (mu)\big)\eta + \frac{\delta \ell}{\delta \rho}(-(\rho\eta)_x)\Big)\, dxdt \\
&= \int -\Big(m_t + m\partial_x u + \partial_x (mu) - \rho\partial_x \frac{\delta \ell}{\delta \rho}\Big)\eta\, dxdt
\end{aligned}
$$

Thus the full EP system for a Lagrangian $\ell(u,\rho)$ will be

$$
m_t + m\partial_x u + \partial_x(mu) + \rho\partial_x \frac{\delta \ell}{\delta \rho} = 0\,, \quad \text{and} \quad \rho_t + \partial_x(\rho u) = 0\,.
$$

▲

(M) Derive the Lie-Poisson bracket for the full Euler-Poincaré system arising from the Lagrangian $\ell(u,\rho)$ and identify the Lie algebra to which it is dual.

Answer.

The corresponding Lie-Poisson bracket may be used to express the Hamiltonian equations as

$$
\partial_t \begin{bmatrix} m \\ \rho \end{bmatrix} = - \begin{bmatrix} m\partial_x + \partial_x m & \rho\partial_x \\ \partial_x \rho & 0 \end{bmatrix} \begin{bmatrix} \frac{\delta h}{\delta m} = u \\ \frac{\delta h}{\delta \rho} = - \frac{\delta \ell}{\delta \rho} \end{bmatrix}
$$

This Lie-Poisson bracket is dual to the semidirect-product Lie algebra $\mathfrak{X}\circledS\Lambda^0$, of vector fields $X \in \mathfrak{X}$ acting scalar functions $\phi \in \Lambda^0$, with commutator

$$
\Big[\big(X, \phi\big)\,,\, \big(\bar{X}, \bar{\phi}\big)\Big] = \Big(\big[X, \bar{X}\big]\,,\, X(\bar{\phi}) - \bar{X}(\phi)\Big) = \Big(X\bar{X}_x - \bar{X}X_x\,,\, X\bar{\phi}_x - \bar{X}\phi_x\Big)
$$

The dual coordinates are $m \in \mathfrak{X}^*$ and $\rho \in \Lambda^1$.

▲

★

7.9 The CH equation is bi-Hamiltonian

The completely integrable CH equation for unidirectional shallow water waves first derived in [CaHo1993],

$$m_t + um_x + 2mu_x = \underbrace{-c_0 u_x + \gamma u_{xxx}}_{\text{Linear Dispersion}}, \qquad m = u - \alpha^2 u_{xx}, \qquad u = K * m \quad \text{with} \quad K(x,y) = \frac{1}{2} e^{-|x-y|}. \tag{7.35}$$

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

$$m_t = -B_2 \frac{\delta H_1}{\delta m} = -B_1 \frac{\delta H_2}{\delta m} \tag{7.36}$$

with

$$\begin{aligned} H_1 &= \frac{1}{2} \int (u^2 + \alpha^2 u_x^2)\, dx\,, \quad \text{and} \quad B_2 = \partial_x m + m\partial_x + c_0 \partial_x + \gamma\, \partial_x^3 \\ H_2 &= \frac{1}{2} \int u^3 + \alpha^2 u\, u_x^2 + c_0 u^2 - \gamma\, u_x^2 \, dx\,, \quad \text{and} \quad B_1 = \partial_x - \alpha^2 \partial_x^3\,. \end{aligned} \tag{7.37}$$

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,

$$\{f, h\}(m) = -\int \frac{\delta f}{\delta m} (c_1 B_1 + c_2 B_2) \frac{\delta h}{\delta m}\, dx\,, \tag{7.38}$$

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.

7.10 Magri's theorem

As we shall see, because equation (7.35) 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

$$\frac{\delta H_n}{\delta m} = \mathcal{R}^T \frac{\delta H_{n-1}}{\delta m}\,, \quad n = -1, 0, 1, 2, \ldots. \tag{7.39}$$

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 $n = 0, 1, 2, \ldots$,

$$m_t^{(n+1)} = K_{n+1}[m] = -B_1 \frac{\delta H_n}{\delta m} = -\, B_2 \frac{\delta H_{n-1}}{\delta m} = B_2 B_1^{-1} K_n[m]\,. \tag{7.40}$$

The first three flows in the "positive hierarchy" when $c_0, \gamma \to 0$ are

$$m_t^{(1)} = 0\,, \quad m_t^{(2)} = -\, m_x\,, \quad m_t^{(3)} = -\,(m\partial + \partial m)u\,, \tag{7.41}$$

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 $n = 0, 1, 2, \ldots$.

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

$$H_0 = \int_{-\infty}^{\infty} m\, dx\,, \qquad H_{-1} = \int_{-\infty}^{\infty} \sqrt{m}\, dx\,, \tag{7.42}$$

and

$$H_{-2} = \frac{1}{2}\int_{-\infty}^{\infty} \left[\frac{\alpha^2}{4} \frac{m_x^2}{m^{5/2}} - \frac{2}{\sqrt{m}} \right]. \tag{7.43}$$

The flow defined by (7.40) for these is thus,

$$m_t^{(0)} = -B_1 \frac{\delta H_{-1}}{\delta m} = -\, B_2 \frac{\delta H_{-2}}{\delta m} = -(\partial - \alpha^2\partial^3) \left(\frac{1}{2\sqrt{m}} \right). \tag{7.44}$$

This flow is similar to the Dym equation,

$$u_{xxt} = \partial^3 \left(\frac{1}{2\sqrt{u_{xx}}} \right). \tag{7.45}$$

7.11 Proof by Magri's Theorem that the CH equation (7.35) is isospectral

The isospectral eigenvalue problem associated with equation (7.35) 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

(7.40), then summing yields

$$B_1 \sum_{n=0}^{\infty} \lambda^n \frac{\delta H_n}{\delta m} = \lambda B_2 \sum_{n=0}^{\infty} \lambda^{(n-1)} \frac{\delta H_{n-1}}{\delta m}\,, \tag{7.46}$$

or, by introducing the **squared eigenfunction relation**,

$$\psi^2(x,t;\lambda) := \sum_{n=-1}^{\infty} \lambda^n \frac{\delta H_n}{\delta m}\,, \tag{7.47}$$

one finds that, formally,

$$B_1 \psi^2(x,t;\lambda) = \lambda B_2 \psi^2(x,t;\lambda)\,. \tag{7.48}$$

This is a *third order* eigenvalue problem for the squared-eigenfunction ψ^2, which turns out to be equivalent to the following *second order* Sturm-Liouville eigenvalue problem,

$$\lambda\left(\frac{1}{4} - \alpha^2 \partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi\,, \tag{7.49}$$

Exercise.

Verify that if ψ satisfies the eigenvalue equation (7.49) then ψ^2 is a solution of (7.48) with

$$B_1 = \partial_x - \alpha^2 \partial_x^3 \quad \text{and} \quad B_2 = \partial_x m + m\partial_x + c_0 \partial_x + \gamma\,\partial_x^3\,,$$

which are the two Hamiltonian structures for CH, as given in equation (7.37). ★

Now, assuming λ will be independent of time, we seek, in analogy with the KdV equation, an evolution equation for ψ of the form,

$$\psi_t = a\psi_x + b\psi\,, \tag{7.50}$$

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 (7.49) and (7.50) implies (7.35). Cross differentiation shows

$$b = -\frac{1}{2} a_x\,, \quad \text{and} \quad a = -\left(\lambda + u\right). \tag{7.51}$$

Consequently,

$$\psi_t = -\,(\lambda + u)\psi_x + \frac{1}{2}u_x\psi\,, \tag{7.52}$$

is the desired evolution equation for ψ.

Summary of the isospectral property of equation (7.35) 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 (7.35) arises as a compatibility condition for two linear equations. These are the *isospectral eigenvalue problem*,

$$\lambda\left(\frac{1}{4} - \alpha^2\partial_x^2\right)\psi = \left(\frac{c_0}{4} + \frac{m(x,t)}{2} + \gamma\,\partial_x^2\right)\psi\,, \tag{7.53}$$

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 (7.35). Consequently, the nonlinear water wave equation (7.35) 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 (7.35) 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 (7.35) becomes the 1D geodesic equation EPDiff(H^1) in (7.26)

$$m_t + um_x + 2mu_x = 0\,,\qquad m = u - \alpha^2 u_{xx}\,, \tag{7.54}$$

and the spectrum of its eigenvalue problem (7.53) becomes *purely disctrete*. The traveling wave solutions of 1D EPDiff (7.54) in this dispersionless case are the "peakons," described by the reduced, or collective, solutions (7.32) for EPDiff equation (7.26) with traveling waves

$$u(x,t) = c\,K(x-ct) = c\,e^{-|x-ct|/\alpha}\,.$$

In this case, the EPDiff equation (7.26) may also be written as a conservation law for momentum,

$$\partial_t m = -\partial_x\Big(um + \frac{1}{2}u^2 - \frac{\alpha^2}{2}u_x^2\Big)\,. \tag{7.55}$$

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 (7.53) 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,

$$u(x,t) = \sum_{i=1}^{N} p_i(t) e^{-|x-q_i(t)|/\alpha} . \tag{7.56}$$

Here $p_i(t)$ and $q_i(t)$ satisfy the finite dimensional geodesic motion equations obtained as canonical Hamiltonian equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} , \tag{7.57}$$

when the Hamiltonian is given by,

$$H = \frac{1}{2} \sum_{i,j=1}^{N} p_i \, p_j \, e^{-|q_i - q_j|/\alpha} . \tag{7.58}$$

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 (7.56) (which produces the sum of delta functions in (7.29) for the momentum map m) into the isospectral problem (7.53). This substitution reduces (7.53) to an $N \times N$ matrix eigenvalue problem.

In fact, the canonical equations (7.57) for the peakon Hamiltonian (7.58) may be written directly in Lax matrix form,

$$\frac{dL}{dt} = [L, A] \quad \Longleftrightarrow \quad L(t) = U(t) L(0) U^{\dagger}(t) , \tag{7.59}$$

with $A = \dot{U} U^{\dagger}(t)$ and $UU^{\dagger} = Id$. Explicitly, L and A are $N \times N$ matrices with entries

$$L_{jk} = \sqrt{p_j p_k} \, \phi(q_j - q_k) , \quad A_{jk} = -2 \sqrt{p_j p_k} \, \phi'(q_j - q_k) . \tag{7.60}$$

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 (7.60) evolves by time-dependent unitary transformations, which leave its spectrum invariant. Isospectrality then implies that the traces $\operatorname{tr} L^n$, $n = 1, 2, \ldots, N$ 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 (7.57) is integrable.

Exercise. Show that the peakon Hamiltonian H_N in (7.58) is expressed as a function of the invariants of the matrix L, as

$$H_N = -\mathrm{tr}\, L^2 + 2(\mathrm{tr}\, L)^2 \,. \tag{7.61}$$

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. ★

7.12 Steepening Lemma: themechanism underlying 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,

$$\partial_t m + u m_x + 2u_x m = 0 \quad \text{with} \quad m = u - \alpha^2 u_{xx} \,, \tag{7.62}$$

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

7.9 (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 (7.37) 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 (7.62), rewritten as,

$$(1 - \alpha^2\partial^2)(u_t + u u_x) = -\,\partial\Big(u^2 + \frac{\alpha^2}{2}u_x^2\Big)\,, \tag{7.63}$$

yields an equation for the evolution of s. Namely, using $u_{xx}(\overline{x}(t),t) = 0$ leads to

$$\frac{ds}{dt} = -\frac{1}{2}s^2 + \frac{1}{2}\int_{-\infty}^{\infty} \operatorname{sgn}(\overline{x}-y)e^{-|\overline{x}-y|}\partial_y\left(u^2 + \frac{1}{2}u_y^2\right) dy\,. \tag{7.64}$$

Integrating by parts and using the inequality $a^2 + b^2 \geq 2ab$, for any two real numbers a and b, leads to

$$\begin{aligned} \frac{ds}{dt} &= -\frac{1}{2}s^2 - \frac{1}{2}\int_{-\infty}^{\infty} e^{-|\overline{x}-y|}\left(u^2 + \frac{1}{2}u_y^2\right) dy + u^2(\overline{x}(t),t) \\ &\leq -\frac{1}{2}s^2 + 2u^2(\overline{x}(t),t)\,. \end{aligned} \tag{7.65}$$

Then, provided $u^2(\overline{x}(t),t)$ remains finite, say less than a number $M/4$, we have

$$\frac{ds}{dt} = -\frac{1}{2}s^2 + \frac{M}{2}\,, \tag{7.66}$$

which implies, for negative slope initially $s \leq -\sqrt{M}$, that

$$s \leq \sqrt{M}\coth\left(\sigma + \frac{t}{2}\sqrt{M}\right)\,, \tag{7.67}$$

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 (7.66) exists is verified in general by a Sobolev inequality. In fact, $M = 8H_1$, since

$$\max_{x\in\mathbb{R}} u^2(x,t) \leq \int_{-\infty}^{\infty} (u^2 + u_x^2)\, dx = 2H_1 = const\,. \tag{7.68}$$

$\square$

Remark

7.10. 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 (7.35). 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 (7.62) 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 (7.53) 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.

8 The Euler-Poincaré framework: fluid dynamics à la [HoMaRa1998a]

The basic idea for the description of fluid dynamics by the action of diffeomorphisms is sketched in Fig 5.

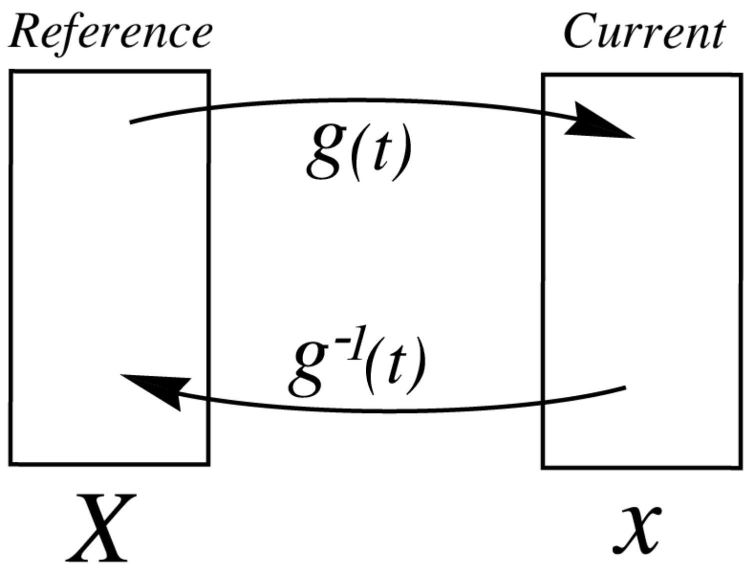


Figure 5: The forward and inverse group actions $g(t)$ and $g^{-1}(t)$ that represent ideal fluid flow are sketched here.

The forward and inverse maps sketched in Fig 5 represent ideal fluid flow by *left* group action of $g_t \in$ Diff on reference $(X \in M)$ and current $(x \in M)$ coordinates. They are denoted as,

$$g_t : x(t,X) = g_t X \quad \text{and} \quad g_t^{-1} : X(t,x) = g_t^{-1} x \,, \tag{8.1}$$

so that taking time derivatives yields

$$\dot{x}(t,X) = \dot{g}_t X = (\dot{g}_t g_t^{-1}) x = \pounds_u x =: u(x,t) = u_t \circ g_t X \,, \tag{8.2}$$

and

$$\dot{X}(t,x) = (T_x g_t^{-1})(\dot{g}_t g_t^{-1} x) = T_x X \cdot u = \pounds_u X =: V(X,t) = V_t \circ g_t^{-1} x \,. \tag{8.3}$$

Here $u = \dot{g}_t g_t^{-1}$ is called the *Eulerian* velocity, and $V = \mathrm{Ad}_{g_t^{-1}} u$ is called the *convective* velocity. For $O_t \in SO(3)$, these correspond to the *spatial* angular velocity $\omega = \dot{O}_t O_t^{-1}$ and the *body* angular velocity $\Omega = \mathrm{Ad}_{O_t^{-1}} \omega = O_t^{-1} \dot{O}_t$. We shall mainly deal with the Eulerian fluid velocity in these notes.

Exercise. Use the Clebsch method to compute the momentum maps for the left group actions in (8.1). ★

8.1 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.[2]
- 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.

[2]For fluid dynamics, right G–invariance of the Lagrangian function L is traditionally called "particle relabeling symmetry."

- For a curve $g(t) \in G$, let $u(t) := \dot{g}(t)g(t)^{-1}$ and define the curve $a(t)$ as the unique solution of the linear differential equation with time dependent coefficients $\dot{a}(t) = -\,a(t)u(t) = \pounds_u a(t)$, where the right action of an element of the Lie algebra $u \in \mathfrak{g}$ on an advected quantity $a \in V^*$ is denoted by concatenation from the right. The solution with initial condition $a(0) = a_0 \in V^*$ can be written as $a(t) = a_0 g(t)^{-1}$.

Notation for reduction of Hamilton's Principle by symmetries

- Let $\mathfrak{g}(\mathcal{D})$ denote the space of vector fields on $\mathcal{D}$ of some fixed differentiability class. These vector fields are endowed with the **Lie bracket** given in components by (summing on repeated indices)

$$[\mathbf{u}, \mathbf{v}]^i = u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j} =: -\,(\mathrm{ad}_{\mathbf{u}}\, \mathbf{v})^i \,. \tag{8.4}$$

  The notation $\mathrm{ad}_{\mathbf{u}}\, \mathbf{v} := -[\mathbf{u},\, \mathbf{v}]$ formally denotes the adjoint action of the *right* Lie algebra of $\mathrm{Diff}(\mathcal{D})$ on itself. This Lie algebra is given by the smooth right-invariant vector fields, $\mathfrak{g} = \mathfrak{X}$.

- Identify the Lie algebra of vector fields $\mathfrak{g}$ with its dual $\mathfrak{g}^*$ by using the L^2 pairing

$$\langle \mathbf{u}, \mathbf{v} \rangle = \int_{\mathcal{D}} \mathbf{u} \cdot \mathbf{v}\, dV \,. \tag{8.5}$$

- Let $\mathfrak{g}(\mathcal{D})^*$ denote the geometric dual space of $\mathfrak{g}(\mathcal{D})$, that is, $\mathfrak{g}(\mathcal{D})^* := \Lambda^1(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$. This is the space of one–form densities on $\mathcal{D}$. If $\mathbf{m} \otimes dV \in \Lambda^1(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, then the pairing of $\mathbf{m} \otimes dV$ with $\mathbf{u} \in \mathfrak{g}(\mathcal{D})$ is given by the L^2 pairing,

$$\langle \mathbf{m} \otimes dV, \mathbf{u} \rangle = \int_{\mathcal{D}} \mathbf{m} \cdot \mathbf{u}\, dV \tag{8.6}$$

  where $\mathbf{m} \cdot \mathbf{u}$ is the standard contraction of a one–form $\mathbf{m}$ with a vector field $\mathbf{u}$.

- For $\mathbf{u} \in \mathfrak{g}(\mathcal{D})$ and $\mathbf{m} \otimes dV \in \mathfrak{g}(\mathcal{D})^*$, the dual of the adjoint representation is defined by

$$\langle \mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV), \mathbf{v} \rangle = \int_{\mathcal{D}} \mathbf{m} \cdot \mathrm{ad}_{\mathbf{u}} \mathbf{v}\, dV = -\int_{\mathcal{D}} \mathbf{m} \cdot [\mathbf{u}, \mathbf{v}]\, dV \tag{8.7}$$

  and its expression is

$$\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV) = (\pounds_{\mathbf{u}} \mathbf{m} + (\mathrm{div}_{dV}\, \mathbf{u})\mathbf{m}) \otimes dV = \pounds_{\mathbf{u}}(\mathbf{m} \otimes dV)\,, \tag{8.8}$$

  where $\mathrm{div}_{dV} \mathbf{u}$ is the divergence of $\mathbf{u}$ relative to the measure dV, that is, $\pounds_{\mathbf{u}} dV = (\mathrm{div}_{dV} \mathbf{u}) dV$. Hence, $\mathrm{ad}^*_{\mathbf{u}}$ coincides with the Lie-derivative $\pounds_{\mathbf{u}}$ for one-form densities.

- If $\mathbf{u} = u^j \partial/\partial x^j$, $\mathbf{m} = m_i dx^i$, then the one–form factor in the preceding formula for $\mathrm{ad}^*_{\mathbf{u}}(\mathbf{m} \otimes dV)$ has the **coordinate expression**

$$\Big(\mathrm{ad}^*_{\mathbf{u}}\, \mathbf{m}\Big)_i dx^i = \left(u^j \frac{\partial m_i}{\partial x^j} + m_j \frac{\partial u^j}{\partial x^i} + (\mathrm{div}_{dV}\, \mathbf{u}) m_i \right) dx^i = \left(\frac{\partial}{\partial x^j}(u^j m_i) + m_j \frac{\partial u^j}{\partial x^i} \right) dx^i \,. \tag{8.9}$$

The last equality assumes that the divergence is taken relative to the standard measure $dV = d^n\mathbf{x}$ in $\mathbb{R}^n$. (On a Riemannian manifold the metric divergence needs to be used.)

Definition

8.1. The ***representation space*** V^* of $\mathrm{Diff}(\mathcal{D})$ in continuum mechanics is often some subspace of the tensor field densities on $\mathcal{D}$, denoted as $\mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, and the representation is given by pull back. It is thus a *right* representation of $\mathrm{Diff}(\mathcal{D})$ on $\mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$. The right action of the Lie algebra $\mathfrak{g}(\mathcal{D})$ on V^* is denoted as ***concatenation from the right***. That is, we denote

$$a\mathbf{u} := \pounds_{\mathbf{u}} a \,,$$

which is the Lie derivative of the tensor field density a along the vector field $\mathbf{u}$.

Definition

8.2. The ***Lagrangian of a continuum mechanical system*** is a function

$$L : T\,\mathrm{Diff}(\mathcal{D}) \times V^* \to \mathbb{R} \,,$$

which is right invariant relative to the tangent lift of right translation of $\mathrm{Diff}(\mathcal{D})$ on itself and pull back on the tensor field densities. Invariance of the Lagrangian L induces a function $\ell : \mathfrak{g}(\mathcal{D}) \times V^* \to \mathbb{R}$ given by

$$\ell(\mathbf{u}, a) = L(\mathbf{u} \circ \eta, \eta^* a) = L(\mathbf{U}, a_0) \,,$$

where $\mathbf{u} \in \mathfrak{g}(\mathcal{D})$ and $a \in V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$, and where $\eta^* a$ denotes the pull back of a by the diffeomorphism η and $\mathbf{u}$ is the Eulerian velocity. That is,

$$\mathbf{U} = \mathbf{u} \circ \eta \quad \text{and} \quad a_0 = \eta^* a \,. \tag{8.10}$$

The evolution of a is by right action, given by the equation

$$\dot{a} = -\,\pounds_{\mathbf{u}}\, a = -\, a\mathbf{u}. \tag{8.11}$$

The solution of this equation, for the initial condition a_0, is

$$a(t) = \eta_{t*} a_0 = a_0 g^{-1}(t)\,, \tag{8.12}$$

where the lower star denotes the push forward operation and η_t is the flow of $\mathbf{u} = \dot{g}g^{-1}(t)$.

Definition

8.3. ***Advected Eulerian quantities*** are defined in continuum mechanics to be those variables which are Lie transported by the flow of the Eulerian velocity field. Using this standard terminology, equation (8.11), or its solution (8.12) states that the tensor field density $a(t)$ (which may include mass density and other Eulerian quantities) is advected.

Remark

8.4 (Dual tensors)**.** As we mentioned, typically $V^* \subset \mathfrak{T}(\mathcal{D}) \otimes \mathrm{Den}(\mathcal{D})$ for continuum mechanics. On a general manifold, tensors of a given type have natural duals. For example, symmetric covariant tensors are dual to symmetric contravariant tensor densities, the pairing being given by the integration of the natural contraction of these tensors. Likewise, k–forms are naturally dual to $(n-k)$–forms, the pairing being given by taking the integral of their wedge product.

Definition

8.5. The ***diamond operation*** $\diamond$ between elements of V and V^* produces an element of the dual Lie algebra $\mathfrak{g}(\mathcal{D})^*$ and is defined as

$$\langle b \diamond a, \mathbf{w} \rangle = -\int_{\mathcal{D}} b \cdot \pounds_{\mathbf{w}}\, a\,, \tag{8.13}$$

where $b \cdot \pounds_{\mathbf{w}}\, a$ denotes the contraction, as described above, of elements of V and elements of V^* and $\mathbf{w} \in \mathfrak{g}(\mathcal{D})$. (These operations do *not* depend on a Riemannian structure.)

For a path $\eta_t \in \mathrm{Diff}(\mathcal{D})$, let $\mathbf{u}(x,t)$ be its Eulerian velocity and consider the curve $a(t)$ with initial condition a_0 given by the equation

$$\dot{a} + \pounds_{\mathbf{u}} a = 0. \tag{8.14}$$

Let the Lagrangian $L_{a_0}(\mathbf{U}) := L(\mathbf{U}, a_0)$ be right-invariant under $\mathrm{Diff}(\mathcal{D})$. We can now state the Euler–Poincaré Theorem for Continua of [HoMaRa1998a].

Theorem

8.6 (Euler–Poincaré Theorem for Continua.). *Consider a path η_t in $\mathrm{Diff}(\mathcal{D})$ with Lagrangian velocity $\mathbf{U}$ and Eulerian velocity $\mathbf{u}$. The following are equivalent:*

i *Hamilton's variational principle*

$$\delta \int_{t_1}^{t_2} L\left(X, \mathbf{U}_t(X), a_0(X)\right)\, dt = 0 \tag{8.15}$$

holds, for variations $\delta\eta_t$ vanishing at the endpoints.

ii *η_t satisfies the Euler–Lagrange equations for L_{a_0} on $\mathrm{Diff}(\mathcal{D})$.*

iii *The constrained variational principle in Eulerian coordinates*

$$\delta \int_{t_1}^{t_2} \ell(\mathbf{u}, a)\, dt = 0 \tag{8.16}$$

holds on $\mathfrak{g}(\mathcal{D}) \times V^*$, using variations of the form

$$\delta \mathbf{u} = \frac{\partial \mathbf{w}}{\partial t} + [\mathbf{u}, \mathbf{w}] = \frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\,\mathbf{u}}\mathbf{w}\,, \qquad \delta a = -\pounds_{\mathbf{w}}\, a, \tag{8.17}$$

where $\mathbf{w}_t = \delta\eta_t \circ \eta_t^{-1}$ vanishes at the endpoints.

iv *The Euler–Poincaré equations for continua*

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \mathbf{u}} = -\,\mathrm{ad}^*_{\mathbf{u}}\frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a = -\pounds_{\mathbf{u}}\frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta a} \diamond a\,, \tag{8.18}$$

hold, with auxiliary equations $(\partial_t + \pounds_{\mathbf{u}})a = 0$ for each advected quantity $a(t)$. The $\diamond$ operation defined in (8.13) needs to be determined on a case by case basis, depending on the nature of the tensor $a(t)$. The variation $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ is a one–form density and we have used relation (8.8) in the last step of equation (8.18).

We refer to [HoMaRa1998a] for the proof of this theorem in the abstract setting. We shall see some of the features of this result in the concrete setting of continuum mechanics shortly.

Discussion of the Euler-Poincaré equations

The following string of equalities shows *directly* that **iii** is equivalent to **iv**:

$$
\begin{aligned}
0 &= \delta \int_{t_1}^{t_2} l(\mathbf{u}, a)dt = \int_{t_1}^{t_2} \left(\frac{\delta l}{\delta \mathbf{u}} \cdot \delta \mathbf{u} + \frac{\delta l}{\delta a} \cdot \delta a \right) dt \\
&= \int_{t_1}^{t_2} \left[\frac{\delta l}{\delta \mathbf{u}} \cdot \left(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\, \mathbf{w} \right) - \frac{\delta l}{\delta a} \cdot \pounds_{\mathbf{w}}\, a \right] dt \\
&= \int_{t_1}^{t_2} \mathbf{w} \cdot \left[- \frac{\partial}{\partial t} \frac{\delta l}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}} \frac{\delta l}{\delta \mathbf{u}} + \frac{\delta l}{\delta a} \diamond a \right] dt\,.
\end{aligned}
\tag{8.19}
$$

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[3] can be written in the EP form in equation (8.18),

$$
\frac{\partial}{\partial t} \frac{\delta \ell}{\delta \mathbf{u}} + \mathrm{ad}^*_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} = \frac{\delta \ell}{\delta a} \diamond a\,, \quad \text{with} \quad (\partial_t + \pounds_{\mathbf{u}})a = 0\,. \tag{8.20}
$$

Equation (8.20) is **Newton's Law**: The Eulerian time derivative of the momentum density $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ (a one-form density dual to the velocity $\mathbf{u}$) is equal to the force density $(\delta\ell/\delta a) \diamond a$, with the $\diamond$ operation defined in (8.13). Thus, Newton's Law is written in the Eulerian fluid representation as,[4]

$$
\frac{d}{dt}\Big|_{Lag} \mathbf{m} := (\partial_t + \pounds_{\mathbf{u}})\mathbf{m} = \frac{\delta \ell}{\delta a} \diamond a\,, \quad \text{with} \quad \frac{d}{dt}\Big|_{Lag} a := (\partial_t + \pounds_{\mathbf{u}})a = 0\,. \tag{8.21}
$$

- The left side of the EP equation in (8.21) describes the fluid's dynamics due to its kinetic energy. A fluid's kinetic energy typically defines a norm for the Eulerian fluid velocity, $KE = \frac{1}{2}\|\mathbf{u}\|^2$. The left side of the EP equation is the **geodesic** part of its evolution, with respect to this norm.

[3]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.

[4] In coordinates, a one-form density takes the form $\mathbf{m} \cdot d\mathbf{x} \otimes dV$ and the EP equation (8.18) is given neumonically by

$$
\frac{d}{dt}\Big|_{Lag} (\mathbf{m} \cdot d\mathbf{x} \otimes dV) = \underbrace{\frac{d\mathbf{m}}{dt}\Big|_{Lag} \cdot\, d\mathbf{x} \otimes dV}_{\text{Advection}} + \underbrace{\mathbf{m} \cdot d\mathbf{u} \otimes dV}_{\text{Stretching}} + \underbrace{\mathbf{m} \cdot d\mathbf{x} \otimes (\nabla \cdot \mathbf{u})dV}_{\text{Expansion}} = \frac{\delta \ell}{\delta a} \diamond a
$$

with $\frac{d}{dt}\big|_{Lag} d\mathbf{x} := (\partial_t + \pounds_{\mathbf{u}})d\mathbf{x} = d\mathbf{u} = \mathbf{u}_{,j} dx^j$, 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 (8.9).

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 (8.21) modifies the geodesic motion. Naturally, the right side of the EP equation is also a geometrical quantity. The diamond operation $\diamond$ represents the dual of the Lie algebra action of vectors fields on the tensor a. Here $\delta\ell/\delta a$ is the dual tensor, under the natural pairing (usually, L^2 pairing) $\langle\,\cdot\,,\,\cdot\,\rangle$ that is induced by the variational derivative of the Lagrangian $\ell(\mathbf{u}, a)$. The diamond operation $\diamond$ is defined in terms of this pairing in (8.13). For the L^2 pairing, this is integration by parts of (minus) the Lie derivative in (8.13).

- The quantity a is typically a tensor (e.g., a density, a scalar, or a differential form) and we shall sum over the various types of tensors a that are involved in the fluid description. The second equation in (8.21) states that each tensor a is carried along by the Eulerian fluid velocity $\mathbf{u}$. Thus, a is for fluid "attribute," and its Eulerian evolution is given by minus its Lie derivative, $-\pounds_{\mathbf{u}} a$. That is, a stands for the set of fluid attributes that each Lagrangian fluid parcel carries around (advects), such as its buoyancy, which is determined by its individual salt, or heat content, in ocean circulation.

- Many examples of how equation (8.21) arises in the dynamics of continuous media are given in [HoMaRa1998a]. The EP form of the Eulerian fluid description in (8.21) is analogous to the classical dynamics of rigid bodies (and tops, under gravity) in body coordinates. Rigid bodies and tops are also governed by Euler-Poincaré equations, as Poincaré showed in a two-page paper with no references, over a century ago [Po1901]. For modern discussions of the EP theory, see, e.g., [MaRa1994], or [HoMaRa1998a].

Exercise. For what types of tensors a_0 can one recast the EP equations for continua (8.18) as geodesic motion, perhaps by using a version of the Kaluza-Klein construction? ★

Exercise. State the EP theorem and write the EP equations for the *convective* velocity. ★

8.2 Corollary of the EP theorem: the Kelvin-Noether circulation theorem

Corollary

8.7 (Kelvin-Noether Circulation Theorem.)**.** Assume $\mathbf{u}(x,t)$ satisfies the Euler–Poincaré equations for continua:

$$\frac{\partial}{\partial t}\left(\frac{\delta \ell}{\delta \mathbf{u}}\right) = -\pounds_{\mathbf{u}}\left(\frac{\delta \ell}{\delta \mathbf{u}}\right) + \frac{\delta \ell}{\delta a}\diamond a$$

and the quantity a satisfies the ***advection relation***

$$\frac{\partial a}{\partial t} + \pounds_{\mathbf{u}} a = 0. \tag{8.22}$$

Let η_t be the flow of the Eulerian velocity field $\mathbf{u}$, that is, $\mathbf{u} = (d\eta_t/dt)\circ\eta_t^{-1}$. Define the advected fluid loop $\gamma_t := \eta_t \circ \gamma_0$ and the circulation map $I(t)$ by

$$I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta \mathbf{u}}\,. \tag{8.23}$$

In the circulation map $I(t)$ the advected mass density D_t satisfies the push forward relation $D_t = \eta_* D_0$. This implies the advection relation (8.22) with $a = D$, namely, the continuity equation,

$$\partial_t D + \operatorname{div} D\mathbf{u} = 0\,.$$

Then the map $I(t)$ satisfies the ***Kelvin circulation relation***,

$$\frac{d}{dt}I(t) = \oint_{\gamma_t} \frac{1}{D}\frac{\delta \ell}{\delta a}\diamond a\,. \tag{8.24}$$

Both an abstract proof of the Kelvin-Noether Circulation Theorem and a proof tailored for the case of continuum mechanical systems are given in [HoMaRa1998a]. We provide a version of the latter below.

Proof. First we change variables in the expression for $I(t)$:

$$I(t) = \oint_{\gamma_t} \frac{1}{D_t}\frac{\delta l}{\delta \mathbf{u}} = \oint_{\gamma_0} \eta_t^*\left[\frac{1}{D_t}\frac{\delta l}{\delta \mathbf{u}}\right] = \oint_{\gamma_0} \frac{1}{D_0}\eta_t^*\left[\frac{\delta l}{\delta \mathbf{u}}\right].$$

Next, we use the Lie derivative formula, namely

$$\frac{d}{dt}\left(\eta_t^*\alpha_t\right) = \eta_t^*\left(\frac{\partial}{\partial t}\alpha_t + \pounds_{\mathbf{u}}\alpha_t\right) \,,$$

applied to a one–form density α_t. This formula gives

$$\begin{aligned}
\frac{d}{dt}I(t) &= \frac{d}{dt}\oint_{\gamma_0}\frac{1}{D_0}\eta_t^*\left[\frac{\delta l}{\delta \mathbf{u}}\right] \\
&= \oint_{\gamma_0}\frac{1}{D_0}\frac{d}{dt}\left(\eta_t^*\left[\frac{\delta l}{\delta \mathbf{u}}\right]\right) \\
&= \oint_{\gamma_0}\frac{1}{D_0}\eta_t^*\left[\frac{\partial}{\partial t}\left(\frac{\delta l}{\delta \mathbf{u}}\right) + \pounds_{\mathbf{u}}\left(\frac{\delta l}{\delta \mathbf{u}}\right)\right].
\end{aligned}$$

By the Euler–Poincaré equations (8.18), this becomes

$$\frac{d}{dt}I(t) = \oint_{\gamma_0}\frac{1}{D_0}\eta_t^*\left[\frac{\delta l}{\delta a}\diamond a\right] = \oint_{\gamma_t}\frac{1}{D_t}\left[\frac{\delta l}{\delta a}\diamond a\right],$$

again by the change of variables formula. □

Corollary

8.8. Since the last expression holds for every loop γ_t, we may write it as

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right)\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} = \frac{1}{D}\frac{\delta l}{\delta a}\diamond a\,. \tag{8.25}$$

Remark

8.9. The Kelvin-Noether theorem is called so here because its derivation relies on the invariance of the Lagrangian L under the particle relabeling symmetry, and Noether's theorem is associated with this symmetry. However, the result (8.24) is the ***Kelvin circulation theorem***: the circulation integral $I(t)$ around any fluid loop (γ_t, moving with the velocity of the fluid parcels $\mathbf{u}$) is invariant under the fluid motion. These two statements are equivalent. We note that ***two velocities*** appear in the integrand $I(t)$: the fluid velocity $\mathbf{u}$ and $D^{-1}\delta\ell/\delta\mathbf{u}$. The latter velocity is the momentum density $\mathbf{m} = \delta\ell/\delta\mathbf{u}$ divided by the mass density D. These two velocities are the basic ingredients for performing modeling and analysis in any ideal fluid problem. One simply needs to put these ingredients together in the Euler-Poincaré theorem and its corollary, the Kelvin-Noether theorem.

8.3 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

$$h(m,a)=\Big\langle m\,,\,u\Big\rangle - l(u,a)\,. \tag{8.26}$$

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 (8.18) solely in terms of momentum m and advected quantities a as

$$\begin{aligned}
\partial_t m &= -\,\mathrm{ad}^*_{\delta h/\delta m}\,m - \frac{\delta h}{\delta a}\diamond a\,,\\
\partial_t a &= -\,\pounds_{\delta h/\delta m}\,a\,.
\end{aligned} \tag{8.27}$$

Hamiltonian equations The corresponding Hamiltonian equation for any functional of $f(m,a)$ is then

$$\begin{aligned}
\frac{d}{dt}f(m,a) &= \Big\langle\,\partial_t m\,,\,\frac{\delta f}{\delta m}\,\Big\rangle + \Big\langle\,\partial_t a\,,\,\frac{\delta f}{\delta a}\,\Big\rangle\\
&= -\Big\langle\,\mathrm{ad}^*_{\delta h/\delta m}\,m + \frac{\delta h}{\delta a}\diamond a\,,\,\frac{\delta f}{\delta m}\,\Big\rangle - \Big\langle\,\pounds_{\delta h/\delta m}\,a\,,\,\frac{\delta f}{\delta a}\,\Big\rangle\\
&= -\Big\langle\,m\,,\,\Big[\frac{\delta f}{\delta m},\,\frac{\delta h}{\delta m}\Big]\,\Big\rangle + \Big\langle\,a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\,\Big\rangle\\
&=: \ \{f\,,\,h\}(m,a)\,,
\end{aligned} \tag{8.28}$$

which is plainly antisymmetric under the exchange $f\leftrightarrow h$. Assembling these equations into Hamiltonian form gives, symbolically,

$$\frac{\partial}{\partial t}\begin{bmatrix} m\\ a\end{bmatrix} = -\begin{bmatrix}\mathrm{ad}^*_{\square}m & \square\diamond a\\ \pounds_{\square}a & 0\end{bmatrix}\begin{bmatrix}\delta h/\delta m\\ \delta h/\delta a\end{bmatrix} \tag{8.29}$$

The boxes $\square$ in Equation (8.29) indicate how the various operations are applied in the matrix multiplication. For example,

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

so each matrix entry acts on its corresponding vector component.

Remark

8.10. *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 (8.28) *defines the Lie-Poisson bracket on the dual to the semidirect-product Lie algebra* $\mathfrak{X}\circledS V^*$ *with Lie bracket*

$$\mathrm{ad}_{(u,\alpha)}\,(\overline{u},\,\overline{\alpha}) = \big(\mathrm{ad}_u\overline{u},\,\pounds^T_u\overline{\alpha} - \pounds^T_{\overline{u}}\alpha\big)$$

The coordinates are velocity vector field $u\in\mathfrak{X}$ *dual to momentum density* $m\in\mathfrak{X}^*$ *and* $\alpha\in V^*$ *dual to the vector space of advected quantities* $a\in V$.

Proof. We check that

$$\begin{aligned}
\frac{df}{dt}(m,a) = \{f\,,\,h\}(m,a) &= \left\langle m\,,\,\mathrm{ad}_{\frac{\delta f}{\delta m}}\frac{\delta h}{\delta m}\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a} - \pounds^T_{\delta h/\delta m}\frac{\delta f}{\delta a}\right\rangle\\
&= -\left\langle \mathrm{ad}^*_{\frac{\delta h}{\delta m}} m\,,\,\frac{\delta f}{\delta m}\right\rangle + \left\langle a\,,\,\pounds^T_{\delta f/\delta m}\frac{\delta h}{\delta a}\right\rangle + \left\langle -\pounds_{\delta h/\delta m}a\,,\,\frac{\delta f}{\delta a}\right\rangle\\
&= -\left\langle \mathrm{ad}^*_{\frac{\delta h}{\delta m}} m + \frac{\delta h}{\delta a}\diamond a\,,\,\frac{\delta f}{\delta m}\right\rangle - \left\langle \pounds_{\delta h/\delta m}a\,,\,\frac{\delta f}{\delta a}\right\rangle
\end{aligned}$$

Note that the angle brackets refer to different types of pairings. This should cause no confusion. $\square$

9 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.

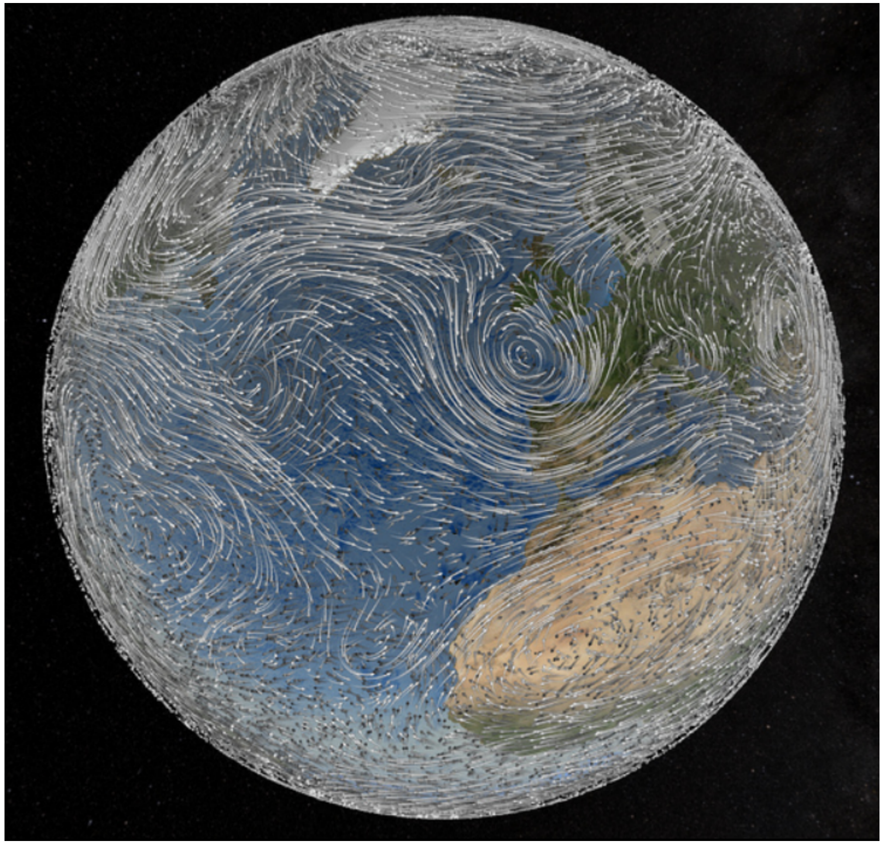


Figure 6: Atmospheric flows on Earth (wind currents) are driven by the Sun and its interaction with the surface and they are damped primarily by friction with the surface.

9.1 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

$$a \in \{b, D\, d^3x\}\,. \tag{9.1}$$

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,

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) b = 0, \quad \text{or} \quad \frac{\partial b}{\partial t} = -\ \mathbf{u}\cdot\nabla\, b\,, \tag{9.2}$$

$$\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right) D\, d^3x = 0, \quad \text{or} \quad \frac{\partial D}{\partial t} = -\ \nabla\cdot(D\mathbf{u})\,. \tag{9.3}$$

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 8.6, equation (8.16), the variations of the tensor functions a at fixed $\mathbf{x}$ and t are also given by Lie derivatives, namely $\delta a = -\,\pounds_{\mathbf{w}}\, a$, or

$$\begin{aligned} \delta b &= -\,\pounds_{\mathbf{w}}\, b = -\,\mathbf{w}\cdot\nabla\, b\,, \\ \delta D\ d^3x &= -\,\pounds_{\mathbf{w}}\,(D\, d^3x) = -\,\nabla\cdot(D\mathbf{w})\ d^3x\,. \end{aligned} \tag{9.4}$$

Hence, Hamilton's principle (8.16) with this dependence yields

$$\begin{aligned} 0 &= \delta\int dt\ \ell(\mathbf{u}; b, D) \\ &= \int dt\ \left[\frac{\delta \ell}{\delta \mathbf{u}}\cdot\delta\mathbf{u} + \frac{\delta \ell}{\delta b}\ \delta b + \frac{\delta \ell}{\delta D}\ \delta D\right] \\ &= \int dt\ \left[\frac{\delta \ell}{\delta \mathbf{u}}\cdot\left(\frac{\partial \mathbf{w}}{\partial t} - \mathrm{ad}_{\mathbf{u}}\,\mathbf{w}\right) - \frac{\delta \ell}{\delta b}\ \mathbf{w}\cdot\nabla\, b - \frac{\delta \ell}{\delta D}\ \Big(\nabla\cdot(D\mathbf{w})\Big)\right] \\ &= \int dt\ \mathbf{w}\cdot\left[-\,\frac{\partial}{\partial t}\frac{\delta \ell}{\delta \mathbf{u}} - \mathrm{ad}^*_{\mathbf{u}}\ \frac{\delta \ell}{\delta \mathbf{u}} - \frac{\delta \ell}{\delta b}\ \nabla\, b + D\ \nabla\frac{\delta \ell}{\delta D}\right] \\ &= -\int dt\ \mathbf{w}\cdot\left[\left(\frac{\partial}{\partial t} + \pounds_{\mathbf{u}}\right)\frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta b}\ \nabla\, b - D\ \nabla\frac{\delta \ell}{\delta D}\right], \end{aligned} \tag{9.5}$$

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.

9.2 Euler–Poincaré framework for GFD

The Euler–Poincaré equations for continua (8.18) may now be summarized in vector form for advected Eulerian variables a in the set (9.1). We adopt the notational convention of the circulation map I in equations (8.23) and (8.24) that a one form density can be made into a one form (no longer a density) by dividing it by the mass density D and we use the Lie-derivative relation for the continuity equation $(\partial/\partial t+\pounds_{\mathbf{u}})Dd^3x=0$. Then, the Euclidean components of the Euler–Poincaré equations for continua in equation (9.5) are expressed in Kelvin theorem form (8.25) with a slight abuse of notation as

$$\Big(\frac{\partial}{\partial t}+\pounds_{\mathbf{u}}\Big)\Big(\frac{1}{D}\frac{\delta\ell}{\delta\mathbf{u}}\cdot d\mathbf{x}\Big)+\frac{1}{D}\frac{\delta\ell}{\delta b}\nabla b\cdot d\mathbf{x}-\nabla\Big(\frac{\delta\ell}{\delta D}\Big)\cdot d\mathbf{x}=0\,, \tag{9.6}$$

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 (9.6) is the Kelvin–Noether form of the equation of motion for ideal continua. Hence, we have the explicit Kelvin theorem expression, cf. equations (8.23) and (8.24),

$$\frac{d}{dt}\oint_{\gamma_t(\mathbf{u})}\frac{1}{D}\frac{\delta\ell}{\delta\mathbf{u}}\cdot d\mathbf{x}=-\oint_{\gamma_t(\mathbf{u})}\frac{1}{D}\frac{\delta\ell}{\delta b}\nabla b\cdot d\mathbf{x}\,, \tag{9.7}$$

where the curve $\gamma_t(\mathbf{u})$ moves with the fluid velocity $\mathbf{u}$. Then, by Stokes' theorem, the Euler equations generate circulation of $\mathbf{v}:=(D^{-1}\delta l/\delta\mathbf{u})$ whenever the gradients ∇b and $\nabla(D^{-1}\delta l/\delta b)$ are not collinear. The corresponding **conservation of potential vorticity** q on fluid parcels is given by

$$\frac{\partial q}{\partial t}+\mathbf{u}\cdot\nabla q=0\,,\quad\text{where}\quad q=\frac{1}{D}\nabla b\cdot\operatorname{curl}\left(\frac{1}{D}\frac{\delta\ell}{\delta\mathbf{u}}\right). \tag{9.8}$$

This is also called **PV convection**. Equations (9.6-9.8) embody most of the panoply of equations for GFD. The vector form of equation (9.6) is,

$$\underbrace{\Big(\frac{\partial}{\partial t}+\mathbf{u}\cdot\nabla\Big)\Big(\frac{1}{D}\frac{\delta l}{\delta\mathbf{u}}\Big)+\frac{1}{D}\frac{\delta l}{\delta u^j}\nabla u^j}_{\textbf{Geodesic Nonlinearity: Kinetic energy}}\quad=\quad\underbrace{\nabla\frac{\delta l}{\delta D}-\frac{1}{D}\frac{\delta l}{\delta b}\nabla b}_{\textbf{Potential energy}} \tag{9.9}$$

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

$$I=\oint_{\gamma(t)}\frac{1}{D}\frac{\delta l}{\delta\mathbf{u}}\cdot d\mathbf{x}\,.$$

9.3 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

$$l(\mathbf{u}, b, D) = \int \rho_0 D(1+b)\Big(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - gz\Big) - p(D-1)\, d^3x\,, \tag{9.10}$$

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,

$$\begin{aligned}\frac{1}{D}\frac{\delta l}{\delta \mathbf{u}} &= \rho_0(1+b)(\mathbf{u}+\mathbf{R})\,, \quad \frac{\delta l}{\delta b} = \rho_0 D\Big(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\Big)\,,\\ \frac{\delta l}{\delta D} &= \rho_0(1+b)\Big(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R} - gz\Big) - p\,, \quad \frac{\delta l}{\delta p} = -\,(D-1)\,.\end{aligned} \tag{9.11}$$

Hence, from the Euclidean component formula (9.9) for Hamilton principles of this type and the fundamental vector identity,

$$(\mathbf{b}\cdot\nabla)\mathbf{a} + a_j\nabla b^j = -\,\mathbf{b}\times(\nabla\times\mathbf{a}) + \nabla(\mathbf{b}\cdot\mathbf{a})\,, \tag{9.12}$$

we find the motion equation for an Euler fluid in three dimensions,

$$\frac{d\mathbf{u}}{dt} - \mathbf{u}\times\operatorname{curl}\mathbf{R} + g\hat{\mathbf{z}} + \frac{1}{\rho_0(1+b)}\nabla p = 0\,, \tag{9.13}$$

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 (9.2). The pressure p is determined by requiring preservation of the constraint $D = 1$, for which the continuity equation (9.3) implies $\operatorname{div}\mathbf{u} = 0$. The Euler motion equation (9.13) 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

$$\frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q = 0\,, \quad \text{where, on using } D = 1\,, \quad q = \nabla b\cdot\operatorname{curl}\big(\mathbf{u}+\mathbf{R}\big). \tag{9.14}$$

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

$$l(\mathbf{u}, b, D) = \int D\Big(\frac{1}{2}|\mathbf{u}|^2 + \mathbf{u}\cdot\mathbf{R}(\mathbf{x}) - e(D, b)\Big)\, d^3x\,, \tag{9.15}$$

   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.

[CaHo1993] Camassa, R. and Holm, D. D. [1993] An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.* **71**, 1661-64.

[Clebsch1859] A. Clebsch [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.

[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.

[HoKu1988] Holm, D. D. and Kupershmidt, B. A. [1988] The analogy between spin glasses and Yang–Mills fluids. *J. Math Phys.* **29**, 21–30.

[Ho2005] Holm, D. D. [2005] The Euler–Poincaré variational framework for modeling fluid dynamics. In *Geometric Mechanics and Symmetry: The Peyresq Lectures*, edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306. Cambridge: Cambridge University Press.

[Ho2011GM1] Holm, D. D. [2011]
Geometric Mechanics I: Dynamics and Symmetry, Second edition, World Scientific: Imperial College Press, Singapore, .

[Ho2011GM2] Holm, D. D. [2011]
Geometric Mechanics II: Rotating, Translating & Rolling, Second edition, World Scientific: Imperial College Press, Singapore, .

[Ho2011] Holm, D. D. [2011]
Applications of Poisson geometry to physical problems, *Geometry & Topology Monographs* **17**, 221–384.

[HoMaRa1998a] Holm, D. D., Marsden, J. E. and Ratiu, T. S. [1998a] The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. in Math.* **137**, 1–81.

[HoSmSt2009] Holm, D. D., Schmah, T. and Stoica, C. [2009]
Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions,
Oxford University Press.

[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., On Henri Poincarés note: sur une forme nouvelle des equations de la mécanique. *J. Geom. Symmetry Phys.* 29: 1–38 (2013).

[MaRa1994] Marsden, J. E. and Ratiu, T. S. [1994]
Introduction to Mechanics and Symmetry. Texts in Applied Mathematics, Vol. 75. New York: Springer-Verlag.

[Ol2000] P. J. Olver [2000] *Applications of Lie Groups to Differential Equations*. Springer: New York.

[Po1901] H. Poincaré, Sur une forme nouvelle des équations de la méchanique,
C.R. Acad. Sci. **132** (1901) 369-371.
English translation in [Ho2011GM2], Appendix D.

[RaTuSbSoTe2005] Ratiu, T. S., Tudoran, R., Sbano, L., Sousa Dias, E. and Terra, G. [2005]
A crash course in geometric mechanics.
In *Geometric Mechanics and Symmetry: The Peyresq Lectures*,
edited by J. Montaldi and T. Ratiu. London Mathematical Society Lecture Notes Series 306.
Cambridge: Cambridge University Press.